

# Global sensitivity analysis for a numerical model of radionuclide migration from the RRC “Kurchatov Institute” radwaste disposal site

E. Volkova · B. Iooss · F. Van Dorpe

Published online: 23 November 2006  
© Springer-Verlag 2006

**Abstract** Today, in different countries, there exist sites with contaminated groundwater formed as a result of inappropriate handling or disposal of hazardous materials or wastes. Numerical modeling of such sites is an important tool for a correct prediction of contamination plume spreading and an assessment of environmental risks associated with the site. Many uncertainties are associated with a part of the parameters and the initial conditions of such environmental numerical models. Statistical techniques are useful to deal with these uncertainties. This paper describes the methods of uncertainty propagation and global sensitivity analysis that are applied to a numerical model of radionuclide migration in a sandy aquifer in the area of the RRC “Kurchatov Institute” radwaste disposal site in Moscow, Russia. We consider 20 uncertain input parameters of the model and 20 output variables (contaminant concentration in the observation wells predicted by the model for the end of 2010). Monte Carlo simulations allow calculating uncertainty in the output values and analyzing the linearity and the monotony of the relations between input and output variables. For the non monotonic relations, sensitivity analyses are classically done with the Sobol sensitivity

indices. The originality of this study is the use of modern surrogate models (called response surfaces), the boosting regression trees, constructed for each output variable, to calculate the Sobol indices by the Monte Carlo method. It is thus shown that the most influential parameters of the model are distribution coefficients and infiltration rate in the zone of strong pipe leaks on the site. Improvement of these parameters would considerably reduce the model prediction uncertainty.

**Keywords** Radionuclide migration · Groundwater transport · Uncertainty · Sensitivity analysis · Response surface · Computer codes

## 1 Introduction

Today, in different countries, there exist sites with contaminated groundwater formed as a result of inappropriate handling or disposal of hazardous materials or wastes. Numerical modeling of such sites is an important tool for a correct prediction of contamination plume spreading and an assessment of environmental risks associated with the site. It was recognized that in such complex numerical models many input variables are largely uncertain and that a rigorous procedure is necessary to arrive at realistic uncertainty distributions. Numerical transport models also require some initial (contaminant concentrations, aquifer level, etc.) and boundary conditions, which are often unknown or have been only partially recognized.

To deal with all these uncertainties, statistical techniques are useful. In many applications of numerical models simulating physical phenomena, the same

---

E. Volkova (✉)  
RRC “Kurchatov Institute”, Nuclear Safety Institute,  
pl. Kurchatova 1, 123182 Moscow, Russia  
e-mail: elena227@nsi.kiae.ru

B. Iooss  
CEA Cadarache, DEN/DER/SESI/LCFR, Bât. 212,  
13108 Saint-Paul-lez Durance cedex, France

F. Van Dorpe  
CEA Cadarache, DEN/DTN/SMTM/LMTE, Bât. 307,  
13108 Saint-Paul-lez Durance cedex, France

difficulties arise: computer codes are more and more complex, often very cpu time expensive, taking as inputs some large number of parameters and giving as a result many variables of interest. This kind of problems is treated in the general context of the “statistical analysis of computer experiments” (Sacks et al. 1989; Kleijnen 1997; Saltelli et al. 2000; Helton and Davis 2003). The uncertainty analysis is used to evaluate the confidence interval or the density probability distributions of the model responses. The global sensitivity analysis is used to quantify the influence of the uncertainties of input parameters on the uncertainties of output variables. Recent studies have applied different methods of uncertainty and sensitivity analysis to environmental models (Helton 1993; Campolongo and Saltelli 1997; Maddalena et al. 2001; Tarantola et al. 2002; Ma 2002; Iooss et al. 2006).

This paper describes the methods of uncertainty propagation and global sensitivity analysis that are applied to a numerical model of radionuclide migration in a sandy aquifer in the area of the RRC “Kurchatov Institute” (KI) radwaste disposal site in Moscow, Russia. Combined numerical model of groundwater flow and radionuclide transport on the site is considered. It is created in a hydrogeological program package called MARTHE (developed by the “Bureau de Recherches Géologiques et Minières”, France). This package allows numerical modeling of flow and transport in three-dimensional porous media. MARTHE is adapted for dealing with various problems of subsurface hydrodynamics (water balance assessment, hydrodynamic influence of pumping, drainage and other types of interference, groundwater contaminant spreading prediction, etc.).

In our case, the package was used to construct a combined three-dimensional flow and transport model of Sr-90 radionuclide spreading. Our scenario considers a period of 8 years, starting from an initial contamination plume observed in groundwater in 2002 and predicting its spreading for the end of 2010. Radionuclide contamination has spread into groundwater from a series of temporary waste storages on the territory of the RRC “KI” that is situated close to the city residential area. That is why it is important to make an accurate prediction of the plume spreading in order to assess the potential risk and to be able to guide the future works on the site. Prediction made by the model is associated with some uncertainty, coming from uncertainty in initial conditions and input parameters values. Uncertainty propagation and sensitivity analysis for the constructed model will allow evaluating this uncertainty and distinguishing parameters that contribute mostly to prediction uncertainty. Improvement

of these parameters would considerably reduce the model uncertainty.

The following section presents the four steps of the general statistical methodology: uncertainty propagation, linear sensitivity analysis between input and output variables, calculations of non linear sensitivity measures (Sobol indices), construction of response surfaces (surrogate or simplified models requiring negligible computation times) to replace the computer codes for the Monte Carlo simulations. In the third section, we develop and explain the numerical model of radionuclide migration from the RRC “KI” radwaste disposal site. A statistical analysis of this model is detailed in Sect. 4. A discussion concludes the paper.

## 2 Global sensitivity analysis methodology

The sensitivity analysis aims to investigate how a given computational model responds to variations in its inputs. Such knowledge is useful for determining the degree of resemblance of a model and a real system, distinguishing the factors that mostly influence the output variability and those that are insignificant, revealing interactions among input parameters and correlations among output variables, etc. A detailed description of sensitivity analysis methods can be found in Saltelli et al. (2000).

In the global sensitivity analysis strategy, the emphasis is put on apportioning the output uncertainty to the uncertainty in the input factors, given by their uncertainty ranges and probability distributions. Most of the used methods are based on variance analysis.

Assume that the model under consideration is given by  $Y = f(X)$ , where  $X = (X_1, \dots, X_p)$  are independent input random values,  $f$  is a deterministic function and  $Y$  is a scalar model output. Model function  $f$  can represent a system of differential equations, a program code or any other deterministic correspondence between  $X$  and  $Y$  values that can be calculated for a finite period of time.

Let  $X^*$  be some initial set of model input data. In many works partial derivative  $\left. \frac{\partial f}{\partial X_k} \right|_{X=X^*}$  is referred to by the term “sensitivity” of response  $Y^* = f(X^*)$  to the input parameter  $X_k$ . Such an approach is often called local sensitivity analysis.

Global sensitivity analysis (called “global” in opposition to the local one) does not distinguish any initial set of model input values but considers function  $f$  in the entire domain of possible input parameter variations. Thus, the global sensitivity analysis is an instrument to study a mathematical model as a whole rather than its particular solution.

Schematically the global sensitivity analysis algorithm can be presented as follows (Kleijnen 1997, 2005; Campolongo and Saltelli 1997; Iooss et al. 2006):

1. uncertainty propagation: statistical modeling of input parameters, calculation and statistical analysis of a set of output values;
2. correlation analysis of obtained data: calculation of sensitivity indices under linearity or monotony assumptions;
3. nonlinear sensitivity analysis: construction (if needed) of response surfaces (simplified models approximating the computer code) and calculation of Sobol sensitivity indices based on complete variance decomposition.

### 2.1 Uncertainty propagation

The uncertainty analysis of numerical model aims to evaluate the output parameter  $Y$  uncertainty that corresponds to uncertainties in inputs  $X_i$ . Each input parameter  $X_i$  is considered as a random value. Thus, the first stage is the statistical modeling of input parameters by choosing their probability distribution functions. This choice can be based on an expert judgment of parameter physical nature, on some general information on possible value intervals, on statistical analysis of measurements, etc. (Devictor and Bolado Lavin 2006). An example of this uncertainty modeling phase for environmental models can be found in Helton (1993).

Having a large number of experimental or observation data, one can conduct statistical analysis in order to adjust a probability distribution law. The choice of the distribution in this case can be based both on visual data analysis and on classical statistical tests (Chi-square, Kolmogorov–Smirnov, Anderson-Darling, etc.). When the amount of available data is not sufficient for a distribution function construction, expert judgment is used to obtain some modeling hypothesis based on preliminary information about the parameter, its physical properties, etc. In case when preliminary information is insufficient or unsuitable for choosing a probability distribution law, the uniform distribution function can be chosen, with a sufficiently large interval of possible variation. Note that the procedure of probability distribution choice is often strongly subjective.

On the next stage of uncertainty analysis, a large number of model input value variants is generated using the previously chosen distribution functions. The most frequent is the simple Monte Carlo method, while other efficient techniques can be also used, as Latin

Hypercube Sampling (McKay et al. 1979; Helton et al. 2003). Monte Carlo and Latin Hypercube Sampling are becoming the most popular and accepted methods for performing probabilistic analyses on complex models.

For each variant of input values, the model code is executed. All calculated data are then gathered into one output sample for which elementary statistics are calculated. Analytical form of probability density functions can be adjusted using classical statistical tests. At this stage, an important issue to consider is the sample size adequacy. No general rules are known to determine how many samples are necessary to provide sufficient accuracy on mean, standard deviation, or empirical distributions of output variables. In practice, we can use some bootstrap techniques to determine confidence intervals on statistical estimates. Another interesting way to study the sample size influence is to resample the output values and look at the convergence of the different statistical estimates in function of the sample size (Devictor and Bolado Lavin 2006).

On this stage, correlation coefficients of different pairs of output parameters can be calculated. If correlation coefficient between two output parameters is close to  $\pm 1$ , this means that these two outputs are quasi-proportional. Therefore, they are influenced by the input parameters in an identical manner. Consequently, one of these outputs can be excluded from the analysis and calculation costs can be thus reduced.

### 2.2 Correlation analysis

Recall that the Pearson correlation coefficient between samples  $X$  and  $Y$  of size  $N$  that is the simplest sensitivity index, is defined as:

$$\rho_{X,Y} = \frac{\sum_{i=1}^N (X^i - \bar{X})(Y^i - \bar{Y})}{\sqrt{\sum_{i=1}^N (X^i - \bar{X})^2} \sqrt{\sum_{i=1}^N (Y^i - \bar{Y})^2}}, \tag{1}$$

where  $X^i$  and  $Y^i$  are separate values of samples  $X$  and  $Y$  and  $\bar{X}$  and  $\bar{Y}$  are their mean values. If  $\rho_{X,Y}$  is close to  $+1$  or  $-1$ , the relation between  $X$  and  $Y$  is close to linear.

If this is the case for all inputs  $X_i$  ( $i = 1, \dots, p$ ), the model is close to a linear relation that can be found using a linear regression:

$$Y = \beta_0 + \sum_{i=1}^p \beta_i X_i, \tag{2}$$

where  $\beta_i$  ( $i = 1, \dots, p$ ) are the regression coefficients. In this case, the degree of  $X_i$  variance contribution into

the model variance is expressed by the standardized regression coefficient (SRC):

$$\text{SRC}_i = \beta_i \frac{\sigma(X_i)}{\sigma(Y)}, \quad (3)$$

where  $\sigma(X_i)$  and  $\sigma(Y)$  are the standard deviations of the samples. Thus, the index SRC preserves the information of positivity or negativity of relation. Connection between standardized regression coefficient and correlation coefficient is expressed as follows:

$$\rho_{X_i, Y} = \frac{\text{cov}(X_i, Y)}{\sqrt{\sigma^2(X_i)\sigma^2(Y)}} = \frac{\beta_i \sigma^2(X_i)}{\sigma(X_i)\sigma(Y)} = \beta_i \frac{\sigma(X_i)}{\sigma(Y)}, \quad (4)$$

thus,  $\rho_{X_i, Y} = \text{SRC}_i$ .

However, quantifying sensitivity of output  $Y$  to input  $X_i$ , SRC index does not take into account the fact that correlation between  $X_i$  and  $Y$  can be a consequence of a third parameter influence. In order to evaluate correlation of  $Y$  to  $X_i$  without influence of any other parameters, the partial correlation coefficient (PCC) is used (Saltelli et al. 2000).

All these sensitivity indices are valid under the hypothesis of linear relation between the output  $Y$  and all the inputs  $X_i$ . In order to measure the degree of linearity of the relation between  $Y$  and  $X_i$  ( $i = 1, \dots, p$ ), the determination coefficient  $R^2$  is calculated:

$$R^2 = 1 - \frac{\sum_{i=1}^N (Y^i - \tilde{Y}^i)^2}{\sum_{i=1}^N (\bar{Y} - Y^i)^2}, \quad (5)$$

where  $\bar{Y}$  is the mean of  $Y^i$  and  $\tilde{Y}$  is the value of the linear model that is found by linear regression.

The more  $R^2$  is close to 1, the more the relation between  $Y$  and  $X_i$  is close to a linear relation. In this case, coefficient  $\text{SRC}_i$  presents a sensitivity index of the input  $X_i$ . Basing on these indices values, the input parameters can be classified by the degree of their contribution into the output variance. Otherwise (if  $R^2$  is not close to 1),  $\text{SRC}_i$  gives inadequate evaluation of relative input importance.

If the model is not linear, its monotony can be verified using a rank transform of the variables  $Y$  and  $X_i$  ( $i = 1, \dots, p$ ). The values of analogous indices can be calculated: Spearman correlation coefficient, standardized rank regression coefficient (SRRC) and partial rank correlation coefficient (PRCC).

However, in practice one can often encounter models that are neither linear, nor monotone. In this

case Sobol indices can be used to evaluate the model sensitivity to input parameters.

### 2.3 Sobol global sensitivity indices

This variance-based method uses the so-called ANOVA-representation of square integrable functions developed by Sobol (for example, Sobol 2001).

Assume that the model under investigation is given by a function  $Y = f(X)$ , where  $X = (X_1, \dots, X_p)$  are independent input variables and  $Y$  is a scalar output. In order to determine the importance of each input variable, the methods of variance analysis consider to what extent decreases the variance of the output variable when fixing variable  $X_i$  in its true value  $x_i^*$ :

$$V(Y|X_i = x_i^*), \quad (6)$$

where  $V(\cdot)$  denotes the variance function.

However, in general, the true value  $x_i^*$  of  $X_i$  is unknown, so instead of expression 6, the variance of mathematical expectation is considered:

$$E[V(Y|X_i)]. \quad (7)$$

Expression 7 determines, on average, how much  $Y$  varies, if  $X_i$  is fixed. This value decreases if  $X_i$  is an important factor, because  $Y$  will have a small variation.

Instead of using  $E[V(Y|X_i)]$ , we use  $V(E[Y|X_i])$ . The sum of these two terms is constant and equal to  $V(Y)$  by the variance decomposition formula.

Therefore, the global sensitivity index of the first order is defined by:

$$S_i = \frac{V(E[Y|X_i])}{V(Y)} = \frac{V_i}{V} \quad (8)$$

and is comprised between 0 (the case of no influence of  $X_i$  on  $Y$ ) and 1 (all the variability of  $Y$  is explained by  $X_i$ ); the more  $X_i$  is important the closer  $V(E[Y|X_i])$  is to  $V(Y)$  and  $S_i$  to one. Analogously indices of the second and higher orders can be defined by:

$$S_{ij} = \frac{V(E[Y|X_i, X_j])}{V(Y)} = \frac{V_{ij}}{V}, \dots, \quad (9)$$

$$S_{1, \dots, p} = \frac{V(E[Y|X_1, \dots, X_p])}{V(Y)} = \frac{V_{1, \dots, p}}{V}.$$

All  $S_{i_1, \dots, i_s}$  are nonnegative and their sum  $\sum_{s=1}^p \sum_{i_1 < \dots < i_s} S_{i_1, \dots, i_s}$  is equal to one.

Along with these indices, a notion of total sensitivity index has been introduced. It expresses the total

sensitivity of  $Y$  variance to the  $X_i$  input, including sensitivity to all interactions that involve  $X_i$ :

$$S_{T_i} = \sum_{k: i \in k} S_k. \quad (10)$$

In practice, the FAST or Monte Carlo methods are used to calculate the first and total Sobol indices (Saltelli et al. 2000). These methods demand a significant number of model calculations (about 10,000 for complex models) to evaluate one index for one input parameter. With the growth of the number  $p$  of input parameters, the number of first order and total sensitivity indices grows linearly and evaluation of Sobol indices quickly becomes impossible. In order to perform calculations in such cases when the model launching is time consuming (like in our study), the model can be replaced by another one with negligible computation time. Such models are called “response surfaces” or “surrogate models”.

#### 2.4 Response surfaces

The response surface methodology is frequently used to simulate the behavior of an experimental system or a long running computational code based on a certain number of output values. It was originally proposed as a statistical tool, to find the operating conditions of a process at which some responses were optimized (Box and Wilson 1951; Box and Draper 1987). Subsequent generalizations led to these methods being used to develop approximating functions of deterministic computer codes (Downing et al. 1985; Sacks et al. 1989; Kleijnen and Sargent 2000). It consists in generating a surrogate model that fits the initial data, which has good prediction capacities and demands negligible time for one calculation. It is thus efficient for uncertainty and sensitivity analyses requiring several thousands of model calculations (Iooss et al. 2006).

Multiple response surface families exist (Hastie et al. 2002): polynomials, splines, generalized linear models, or learning statistical models like neural networks, support vector machines, etc. Linear and quadratic functions are commonly considered as a first iteration. Knowledge on some input interaction types may be also introduced in polynomials (Jourdan and Zabalza-Mezghani 2004; Kleijnen 2005; Iooss et al. 2006). However, these kinds of models are not always efficient, especially in hydrogeological heterogeneous models which simulate complex and non linear phenomena. For such models, modern statistical learning algorithms can show much better ability to build

accurate models with strong predictive capabilities (Khalil et al. 2004; Marrel et al. 2006).

To construct a response surface, we need the following elements:

- the availability of a model  $f$  representing the phenomenon under investigation;
- a sample  $D$  consisting of  $N$  points  $(\mathbf{x}_i, y_i)$  where  $\mathbf{x}_i$  is a vector of inputs (of dimension  $p$ ),  $y_i = f(\mathbf{x}_i)$  is the response of the model  $f$  and  $i$  varies from 1 to  $N$ ;
- a family  $H$  of functions  $h(\mathbf{x}, \mathbf{c})$  where  $\mathbf{c}$  is a vector of parameters (parametric regression) or a vector of indices (nonparametric regression) that allows distinguishing among different elements of  $H$ .

In order to adjust the best element  $h_0(\mathbf{x}, \mathbf{c}_0)$  from the family  $H$ , the least squares method is often used. It involves the solution of an optimization problem, consisting in minimizing the following function by  $\mathbf{c}$ :

$$R(h) = \frac{1}{N} \sum_{i=1}^N (y_i - h(\mathbf{x}_i, \mathbf{c}))^2. \quad (11)$$

Necessary conditions on response surface include its good approximation and prediction quality. Different characteristics can be used to quantify these qualities: the regression determination coefficient (Eq. 5), statistics of the regression residuals (mean square error, test on the independence and normality, etc.), statistics on the relative residuals, etc. (Kleijnen and Sargent 2000; Iooss et al. 2006). The principal methods of evaluating these characteristics are the cross-validation and bootstrap techniques (Stone 1974; Hastie et al. 2002). The cross validation method allows a good estimation of the theoretical prediction error associated with the response surface, while the bootstrap method is especially useful when the size of the data sample is small. With a sufficient size of data sample, a simpler method can be chosen by:

1. the quality of approximation is given by statistical analysis carried out on the basis of the points used to build the surface (this set of points is called here a “training set”);
2. the quality of prediction is obtained by a statistical analysis carried out on the points not belonging to the building base (this set of points is called a “test base”).

In this study we have used several response surfaces like polynomials, neural networks, support vector machines, etc. In this paper, we present the model that gives the best results: the boosting of regression trees. This modern statistical learning method is based on the

sequential construction of weak models (here these are regression trees with a low interaction depth), which are then aggregated (Schapire 1990). A lot of algorithms are available to construct this model. We use the multiple additive regression trees (MART) algorithm that is one of the most popular and is described in Hastie et al. (2002). The boosting trees method is relatively complex, in sense of, as neural networks, it is a black box model, very efficient but quite difficult to interpret.

### 3 Numerical model of radionuclide migration from the RRC “KI” radwaste disposal site

Numerical model of Sr-90 transport in groundwater is developed for the RRC “KI” radwaste disposal site situated in the city of Moscow, Russia. It aims to provide a correct prediction of further contamination plume spreading till the end of the year 2010 that will show the risks associated with contamination and can serve as a base for future engineering decision-making.

#### 3.1 Description of the site

In the period between 1943 and 1974 radioactive wastes were buried in eleven temporary repositories built on a specially allocated site at the RRC “KI” territory in Moscow. The site used for the temporary radioactive waste repositories has an area of about 2 ha and is situated near the KI external perimeter in the immediate vicinity to the city residential area (Fig. 1) (Volkov et al. 2003, 2004). Radioactive survey of the site and its adjacent area performed in the late 1980s to early 1990s and in 2002 showed that radioac-

tive contamination is present not only at the surface but has a tendency of spreading into groundwater. The porous media of the site is represented principally by sands and clays that form several horizontal superposed aquifers. The uppermost moraine aquifer is unconfined, the others are confined. Seasonal fluctuations of groundwater level are small ( $\leq 30$  cm).

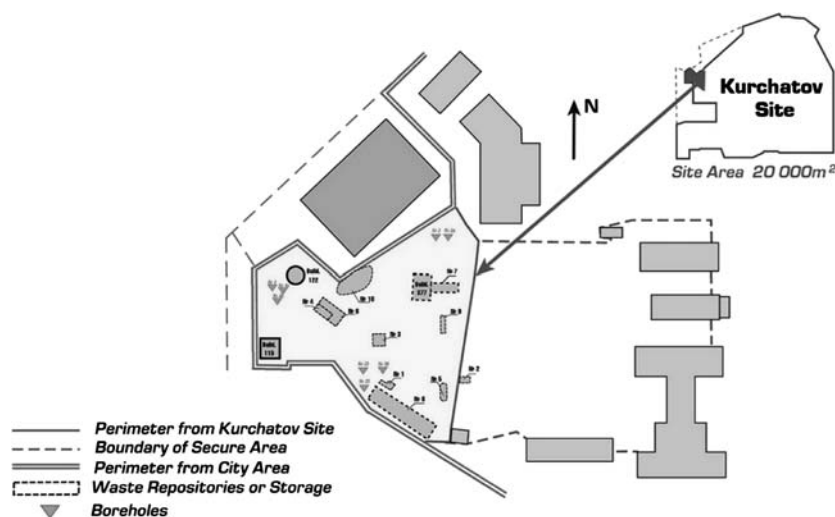
To analyze radioactive contamination of groundwater, about a hundred exploration wells were drilled on the site. As a result of the survey, it was discovered that contamination of the soil is mainly connected with Sr-90 and Cs-137, and contamination of groundwater—with Sr-90. Since the radiation survey results have demonstrated the necessity to clean up the site, rehabilitation activities on radwaste removal and liquidation of old repositories have been started at the site in 2002.

At present almost all old repositories have been disposed. Now groundwater conditions of the two upper aquifers are controlled by an observation network that consists of 20 observation wells for the upper moraine aquifer and nine for the second Jurassic aquifer; this network is used for a regular recording of groundwater levels, its chemical and radionuclide composition. The 20 observation wells of the first aquifer can be visualized in Fig. 2, where they are represented by black and white circles (each one being combined with an alpha-numeric code).

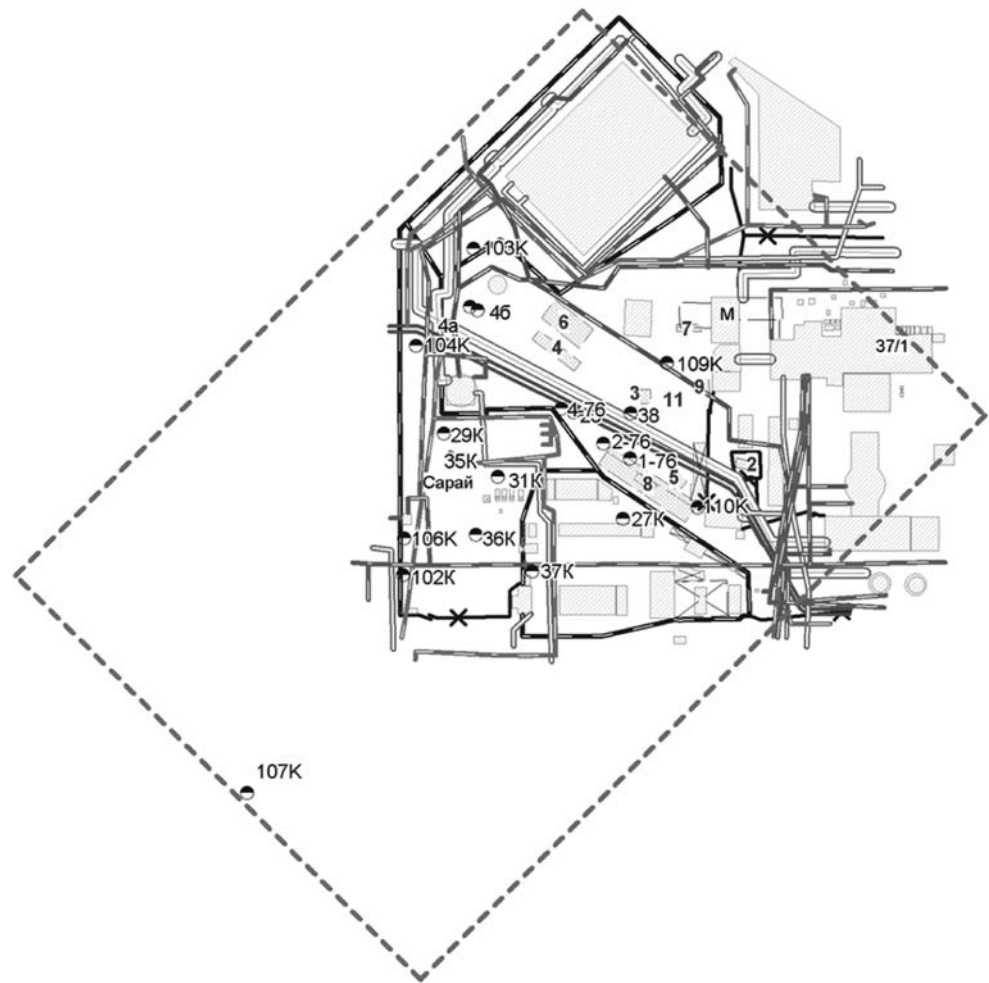
#### 3.2 Model development and calibration

A mathematical model of radionuclide contamination transport from the radwaste disposal site is constructed in the MARTHE hydrogeological program package (developed by BRGM, the French Geological Survey). It is a three-dimensional combined transient flow and

**Fig. 1** Radwaste Disposal Site on the territory of the RRC “Kurchatov Institute”, Moscow, Russia



**Fig. 2** Model domain borders (dashed line) relatively to the radwaste disposal site



transport convection-dispersion model taking into account sorption and radioactive decay. Since the Sr-90 contamination was found to be the most mobile on the site, the model considers only this radionuclide transport in groundwater of the upper moraine aquifer.

The model mesh is chosen oriented south-west along the groundwater flow direction. A uniform planar discretization has a step of  $5\text{ m} \times 5\text{ m}$  and the whole domain has an area of  $250\text{ m} \times 350\text{ m}$  (Fig. 2). Vertical discretization of the model was chosen in accordance with filtration and migration characteristics of the sands forming the upper moraine aquifer; three layers were singled out.

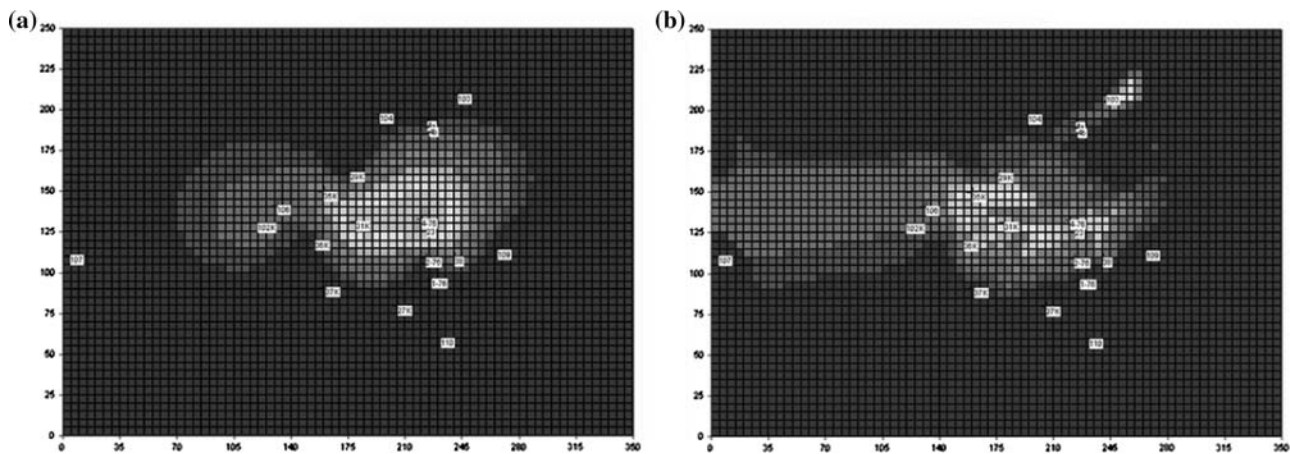
Input parameters of the flow and transport models—hydraulic conductivity of different layers, infiltration, possible locations of leakage from the pipes on the site, sorption, porosity and dispersivity parameters—were obtained as a result of the model verification against observations, field and laboratory experiments, as well as through the use of geostatistical analysis methods.

### 3.3 Modeling results

Prediction of the initial contamination plume spreading was made for the end of year 2010. Initial concentration plume and prediction made for the year 2010 are shown in Fig. 3.

As it can be seen in Fig. 3b contamination plume predicted for the year 2010 is not uniform and is more diffused than the initial one. This is due, above all, to the influence of intensive infiltration assigned in several zones of the model domain (see Fig. 4) that results in local dispersion of contamination plume.

One can thus assume that the form of predicted contamination plume depends essentially on the model input values. Moreover, all the model inputs are exposed to some uncertainty, since their values have been obtained through the model calibration and field and laboratory experiments that have some irreducible error. These errors and other sources of model uncertainty, such as parameterization, for example, lead to errors in model prediction. In order



**Fig. 3** Initial (a) and predicted (b) Sr-90 concentrations (*light color* represents higher levels of concentration); the *small white rectangles* (with alpha-numeric code inside) represent the location of the observation wells

to evaluate the degree of input influence on the resulting contamination plume form and concentration values predicted in observation wells, it seems interesting to perform statistical analysis of the constructed model. The uncertainty propagation method will allow evaluating prediction error, while the global sensitivity analysis will distinguish the most influent model parameters.

#### 4 Analysis for the numerical model of radionuclide migration from the RRC “KI” radwaste disposal site

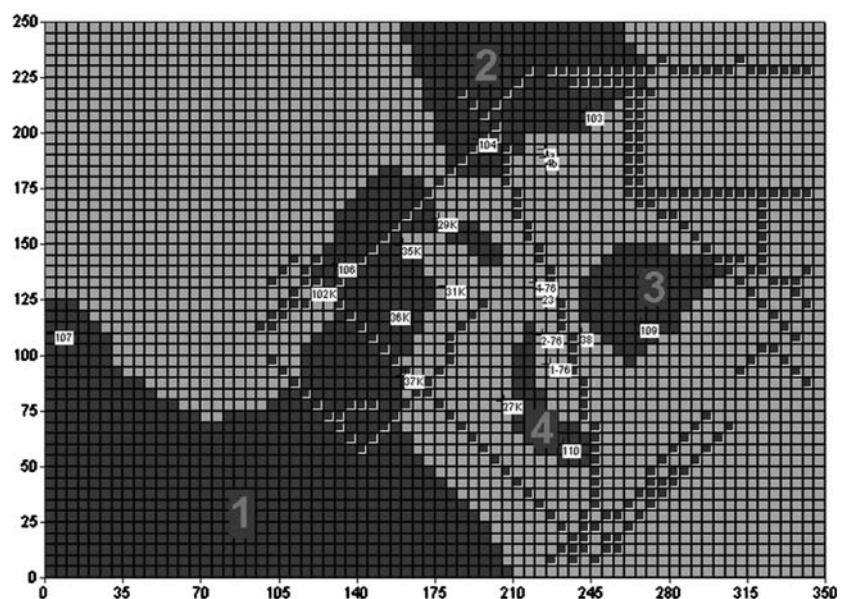
Global statistical analysis and uncertainty propagation techniques described above are applied to the numerical model of radionuclide migration from the RRC

“KI” radwaste disposal site. This section presents procedures of probability distribution choice, input parameters generation, uncertainty propagation and global sensitivity analysis phases including response surface construction and calculation results.

The whole study procedure can be summarized as follows:

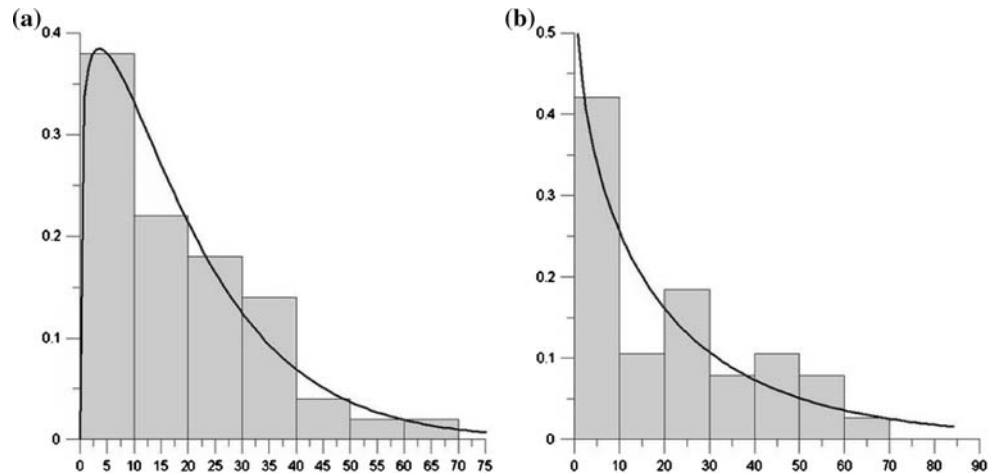
1. Intervals of possible variation and probability distributions on these intervals are chosen for all model inputs included in the analysis.
2. Three hundred variants of model input values are generated using the Latin Hypercube Sampling method. The MARTHE model code is launched for each sample. Thus, 300 variants of output parameters are obtained. They are organized into

**Fig. 4** Zones (numbered from 1 to 4) of low conductivity; lines present zones of high infiltration rates





**Fig. 5** Histograms for Sr-90 distribution coefficient experimental values for the first (a) and second (b) model layers and corresponding Weibull probability distribution density curves,  $m^3/m^3$



histograms and analyzed in order to obtain elementary statistics.

3. Input–output linearity and monotony are then tested and simple sensitivity factors are calculated.
4. To validate obtained sensitivity values for those outputs that display linear or monotonic behavior and to find sensitivity indices for those output that are neither linear nor monotonic, response surfaces are constructed for all model outputs. Afterwards, Sobol sensitivity indices are calculated using these surrogate models.

#### 4.1 Choice and generation of input parameters

Twenty input parameters of the MARTHE numerical model are chosen for uncertainty propagation and global sensitivity analysis. These parameters are the following: hydraulic conductivity of different zones of simulated area, longitudinal and transversal dispersivities of the geological layers, distribution coefficients, porosity and infiltration rates in different zones of the domain. Twenty outputs chosen for the analysis are Sr-90 concentration values calculated by the model at the 20 observation wells locations for the end of the year 2010.

Each input parameter is analyzed in order to choose an interval of possible variation and its probability distribution on this interval. These characteristics are chosen on the base of the analysis of possible parameter values. Such information is taken from different sources: results of field and laboratory experiments, statistical analysis of measurement results (in case of sufficient number of data) and reference literature. For parameters presenting hydraulic conductivity and Sr-90 distribution coefficients, the quantity of available experimental results allows a construction of data

histograms. Corresponding probability distribution laws are chosen using an automatic fitting procedure (examples of such histograms and corresponding distribution densities are shown in Fig. 5). Due to the lack of data, uniform probability distribution laws are taken for the other model inputs (porosity, longitudinal and transversal dispersivities and infiltration rates).

The Table 1 summarizes all the inputs included in sensitivity analysis with their initial (model) values, chosen type and parameters of probability distributions.

According to the chosen intervals and distribution laws, 300 variants of model input values are generated using the Latin Hypercube Sampling method (McKay et al. 1979; Helton and Davis 2003).

Note in Table 1 that certain inputs are generated in a dependent way while variance-based methods of sensitivity analysis demand independency of model input parameters. Correlation coefficients calculated for such pairs show that values of parameters denoted  $i_1$ ,  $i_2$  and  $i_3$  do not really correlate, while correlations in pairs  $dk-dtk$  ( $k = 1,2,3$ ) are significant. Nevertheless, the sensitivity analysis performed later for the same model with combined  $dk-dtk$  parameters showed that the independency assumption for these parameters does not influence the analysis results.

#### 4.2 Uncertainty propagation

A special program that automates cyclic launching of the MARTHE code is created. For each input variant the program reads a set of 20 input parameter values, generates corresponding MARTHE input files, starts the MARTHE code calculation and saves values of 20 output variables in a separate file. Therefore, we recover 300 variants of output variables that present

**Table 1** Input parameters for sensitivity analysis with initial values, distribution types and possible variation intervals

	Parameters	Notation	Model value	Distribution type	Interval or distribution parameters
1	Hydraulic conductivity layer 1	per1	8	Uniform	1–15
2	Hydraulic conductivity layer 2	per2	15	Uniform	5–20
3	Hydraulic conductivity layer 3	per3	8	Uniform	1–15
4	Hydraulic conductivity zone 1	perz1	8	Uniform	1–15
5	Hydraulic conductivity zone 2	perz2	8	Uniform	1–15
6	Hydraulic conductivity zone 3	perz3	8	Uniform	1–15
7	Hydraulic conductivity zone 4	perz4	8	Uniform	1–15
8	Longitudinal dispersivity layer 1	d1	0.8	Uniform	0.05–2
9	Longitudinal dispersivity layer 2	d2	0.8	Uniform	0.05–2
10	Longitudinal dispersivity layer 3	d3	0.8	Uniform	0.05–2
11	Transversal dispersivity layer 1	dt1	0.08	Uniform	0.01d1–0.1d1
12	Transversal dispersivity layer 2	dt2	0.08	Uniform	0.01d2–0.1d2
13	Transversal dispersivity layer 3	dt3	0.08	Uniform	0.01d3–0.1d3
14	Volumetric distribution coefficient l. 1	kd1	5.1	Weibull	1.1597, 19.9875
15	Volumetric distribution coefficient l. 2	kd2	0.34	Weibull	0.891597, 24.4455
16	Volumetric distribution coefficient l. 3	kd3	5.1	Weibull	1.27363, 22.4986
17	Porosity	poros	0.3	Uniform	0.3–0.37
18	Infiltration type 1	i1	0.0001	Uniform	0–0.0001
19	Infiltration type 2	i2	0.004	Uniform	i1–0.01
20	Infiltration type 3	i3	0.02	Uniform	i2–0.1

possible values of Sr-90 concentration in observation wells. Elementary statistics (mean, maximum, minimum, standard deviation, skewness, and kurtosis) are calculated for each output variable. Output histograms give uncertainty in predicted Sr-90 concentration in each of 20 observation wells associated with assumed uncertainties in model inputs. We can see in Fig. 6 that some response distributions are far from Gaussian behaviour.

Two output variables show negligible concentration values for all calculations. These values are not physically significant and it makes no sense to analyze these results. Therefore, we prefer to remove from our analysis these two output variables. Eighteen output variables are now to be considered. This example illustrates the interest of making the uncertainty propagation phase before the sensitivity analysis phase: the uncertainty analysis allows validating the output variables that are to be considered in the sensitivity analysis.

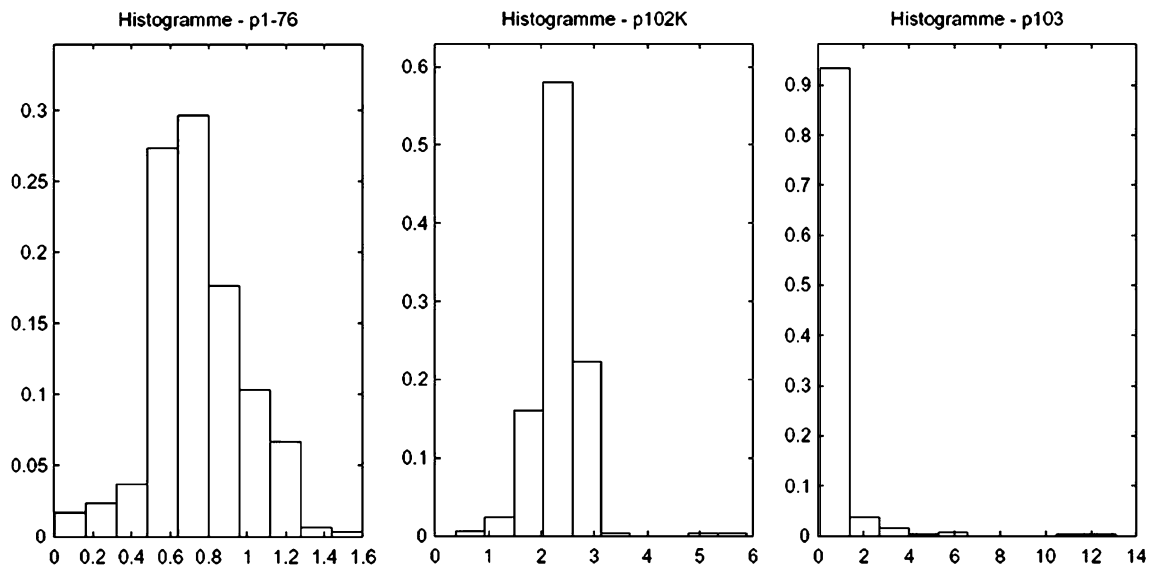
The representativeness of the base of 300 values for each output variable is controlled using a visual criterion of empirical statistical moments estimation with the growth of the number of output values. In this case we start with a sample of 20 values chosen randomly from the whole set of 300 results for the output under consideration. We increase this sample stepwise by adding a randomly chosen sample of 20 values until the whole set is exhausted. This procedure is repeated 200 times and thus for each sample size (from 20 to 300) we obtain the mean and standard deviation empirical

values together with their confidence intervals. The convergence criterion is purely visual: we consider a sample of size  $N$  representative, if empirical values of statistical values tend to corresponding values of the whole set and if the error bars do not fluctuate strongly in the neighborhood of  $N$ . In our case, this criterion gives a positive answer for the majority of output variables. We illustrate this for the first output variable (p1–76) in Fig. 7.

On this stage, correlation coefficients between output variables are also calculated. High values of correlation coefficients correspond to pairs of those observation wells that are located close to each other or in the zones of identical physical properties. Only two pairs of output variables are strongly correlated: p1–76 with p2–76 ( $\rho = 0.9$ ) and p23 with p4–76 ( $\rho = 0.96$ ). This is not sufficient to reduce drastically the number of studied output variables, and in the framework of this study we prefer keeping all of them.

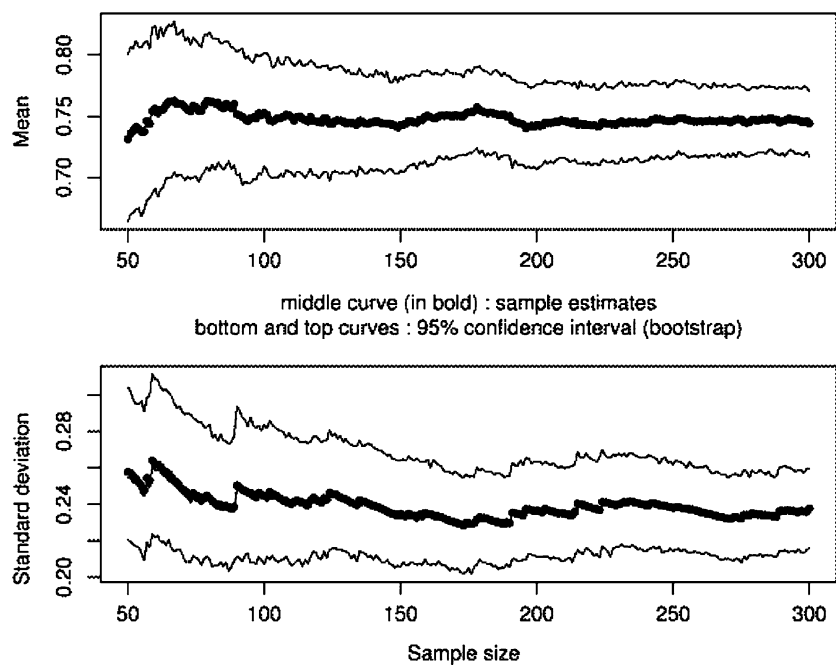
#### 4.3 Correlation analysis

On the first stage of the sensitivity analysis, linear regression is performed. It shows that for the majority of outputs the determination coefficient  $R^2$  of linear regression is not sufficiently close to one, which means that the dependences between inputs and outputs are not linear. Thus, linear sensitivity indices have no quantitative sense in this case. Nevertheless, they give a first idea for the list of the most influent model inputs,



**Fig. 6** Examples of histograms for output parameters p1–76, p102K and p103, Bq/l

**Fig. 7** Example of convergence visualization for the output parameter p1–76. We calculate the mean and standard deviation in function of the sample size (varying between 50 and 300). Confidence intervals are calculated by bootstrap resampling



among which we can see distribution coefficient of the first and second model layers, infiltration intensity of the zones of strong leakage and hydraulic conductivity of different model zones.

The next stage is to verify the monotony of input-output dependence. Determination coefficient  $R^{2*}$  of rank regression has a value greater than 70% for 12 out of 18 output variables (see Table 2). These are outputs p102K, p103, p104, p106, p107, p109, p23, p27K, p29K, p31K, p36K and p4–76. For these outputs, rank sensi-

tivity indices SRRC and PRCC have a quantitative sense. Table 2 gives only the SRRC values which are greater than 10%. The results based on the ranks show approximately the same list of the most influential parameters than the results based on the linear regression.

In order to justify these results and to find sensitivity indices for the outputs that are neither linear nor monotonic, response surfaces are constructed and Sobol sensitivity indices are then calculated.

**Table 2** Determination coefficients of rank regression and monotonic sensitivity indices (SRRC multiplied by  $10^2$ ) of the input parameters for each output variable

	Parameter	$R^{2*}$	kd1	kd2	kd3	i2	i3	per1	per2	per3	pz1	pz2	pz3	pz4
1	p1–76	0.32	30	–35					16			11		
2	p102K	0.90	33	83		–12	–34		–17					
3	p103	0.79	–65	–30	15		28			–11		35		
4	p104	0.91	89				–30	16						
5	p106	0.74	–63					–15	10		57			
6	p107	0.92	–80				19				51			
7	p109	0.70					73	–29	–14	–21			11	
8	p110	0.63	30	–28			66			–13				–14
9	p2–76	0.30	32	–30					16					
10	p23	0.90	32	78			–51							
11	p27K	0.81	–10	–53			73			–16				
12	p29K	0.83	–31	–74			32	–14	22	–14	21			
13	p31K	0.71	15	69			–46	15	–13	13				
14	p35K	0.68	–12	–56			49	–25	12	–17	16			
15	p36K	0.74	–68		–34					19	35			
16	p37K	0.68	15	–19	16		–66	38	15					
17	p38	0.47	48	39			–21	14	–14	–14				13
18	p4–76	0.95	35	84			–45		–11					
19	p4a	0.45				15	55			–20	–11	14		
20	p4b	0.59		44		13	44		–19	–22		17		

4.4 Response surfaces construction

Different types of response surfaces (polynomials, neural networks, support vector machines, etc.) have been tested for each output variable and the best results are obtained using the boosting trees algorithm. It uses 10,000 regression trees with interactions of order 3.

The statistical validity of each response surface is verified using the 10-fold cross-validation method. The procedure is the following:

- The 300 samples database is divided into 10 subsets (each subset contains 30 values).
- The following procedure is repeated ten times: construction of the response surface on nine subsets, prediction on the other subset. At each iteration the prediction subset is different.
- The approximation statistics are estimated on the residuals of the whole construction database (2,700 residual values due to the ten times repetition of the procedure).

**Table 3** Response surface determination coefficients and total Sobol sensitivity indices in percent ( $S_{Ti} \times 10^{-2}$ ) of the input parameters for each output variable

	Variable	$R^2$	kd1	kd2	kd3	i3	per1	per2	per3	perz1	perz2	perz3	d2
1	p1–76	0.58	41	69				1					
2	p102K	0.75	23	69		4		4					
3	p103	–0.09	78	21	1			1			3	1	2
4	p104	0.94	93			4	3						
5	p106	0.28	82							20			
7	p109	0.33	43	1		14	30	2	7	1	4	4	
9	p2–76	0.57	39	68		3		1					
10	p23	0.93	14	78		9							
11	p27K	0.43		78		16	1		7				
12	p29K	0.82	19	66		3	2	4	2	3			
13	p31K	0.45	3	94		6							
14	p35K	0.56		91		3	5	1	8				
15	p36K	0.56	70		11				3	11		1	2
16	p37K	0.64	6	24	2	52	28	3		2			
17	p38	0.31	60	19		19	1						
18	p4–76	0.96	14	80		6							
19	p4a	–0.03		15		33		5	19	4	15	5	5
20	p4b	0.39		71		11		8	6	2			

Only the significant values (>1%) are given

- The prediction statistics are estimated on the residuals of the whole prediction database (300 residual values).

Table 3 gives the determination coefficients obtained on prediction residuals. For several outputs, boosting algorithm improves the quality of approximation as compared to monotonic model, though for other outputs monotonic model gives better approximation. At the same time, certain outputs display non-uniform approximation quality of response surface that is caused by a large amount of output values that are close to zero.

For certain output variables, boosting algorithm has not brought any improvement compared to the regression on the ranks. However, constructed response surfaces allow the calculation of Sobol sensitivity indices that give more rich information than simple correlation and regression coefficients. We will perform calculations for all outputs, though it should be remembered that the “degree of trust” to these results will be expressed by the  $R^2$  value.

#### 4.5 Global sensitivity analysis

Taking into account the quality of response surface adjustment expressed by the  $R^2$  value we can now replace the MARTHE code calculation by the response surface and calculate the Sobol sensitivity indices using the Monte Carlo method. For each output, calculation of Sobol indices is repeated 10 times in order to obtain its mean and variation interval. Total sensitivity indices (in percent)  $S_{T_i}$  for the most influent parameters for each model output are presented in Table 3.

Comparing the total Sobol indices and sensitivity indices calculated in the monotonic context, we obtain coherent sets of the most influent parameters. Note that the three most influent inputs are the same for all types of analysis. These are distribution coefficient of the second model layer (kd2), distribution coefficient of the first model layer (kd1) and infiltration intensity in the zones of leakage (i3). Moreover, these inputs are the only influent parameters for the outputs that display the best quality of response surface adjustment. We should also note that the values of Sobol indices of the first order and the total sensitivity indices are quite close. It means that sensitivities to input parameters are essentially due to the first order, i.e. interactions between inputs are insignificant.

For every model output parameter, one of the three inputs mentioned above is the most influent. Based on this criterion, all outputs (and all observation wells, accordingly) can be divided into three groups: those

that are mostly influenced by kd1, those that are mostly influenced by kd2 and those that are mostly influenced by i3. Such a division correlates with spatial distribution of corresponding observation wells.

It can be also seen that model values of longitudinal and transversal dispersivities as well as model porosity value have practically no influence as compared to the other inputs.

## 5 Discussion and conclusion

The methods of uncertainty propagation and global sensitivity analysis applied to a combined numerical model of groundwater flow and radionuclide transport developed in the MARTHE hydrogeological program package allow estimating the uncertainty of prediction results and quantifying the influence of the model input parameters on Sr-90 concentrations predicted in 20 observation wells. Sensitivity analysis conducted for the radionuclide spreading model shows that Sr-90 concentration values predicted by the model for the end of the year 2010 are mostly influenced by uncertainty in the values of distribution coefficient of the first and second model layers and infiltration intensity in the zones of pipe leakage on the site. At the same time model values of longitudinal and transversal dispersivities as well as model porosity value have practically no influence as compared to the other inputs. Thus, more precise knowledge of the values of the most influent parameters will lead to considerable reduction of the model prediction uncertainty.

This knowledge on the input–output relations gives additional information for site characterization and can help to define, for example, effective sampling or measurements planning, and also to determine more precisely the values of sensible parameters describing the site.

The statistical analysis shows the greatest influence of the Sr-90 sorption effect in soils on the model results as compared with other hydrogeological parameters (permeability or porosity). These distribution coefficients, assigned for Sr-90 at each geological layer, are frequently estimated for transport modeling basing on coefficient data base, and this study shows the importance to have real measurements on the site. We also have time variability of this parameter in relation with chemical conditions of the field (meteoric water infiltrations, etc.) and usually it is not taken into account in distribution coefficient estimations.

In the case of leakage zones, the water infiltration rates are not measured in the field, because it is prac-

tically impossible. Estimation of these values also leads to errors in model responses.

Several remarks can be made that concern further possibilities and extensions of this analysis. First, response surfaces constructed with the boosting method did not show good adjustment for all of the output parameters. In the context of further model analysis, it would be interesting to try other methods based on local polynomials, kriging, etc. Marrel et al. (2006) have done a specific work on the kriging model for complex computer models, and promising results have been obtained on our data: eight output variables are better explained by kriging, while six remain badly simulated. Second, several output variables cause problems for regression since there are many values that are close to zero. This type of data is difficult to model, the results giving elevated values of relative residual statistics. This problem can be treated by using specific weighting of the data in the regression algorithms, by ignoring all small values or by transforming the output variable by taking its logarithm. Finally, the form of zones 1–4 in the second model layer (see Fig. 4), presenting alternation in the layer hydrogeological properties, influences the groundwater flow direction modeled by MARTHE. At the same time, the form of these zones used in this study has been found by interpolation of spatial data that are quite few in number. It seems interesting in a future work to make geostatistical simulations of the form of these zones and to analyze corresponding variation in the model output variables. Recent studies (Tarantola et al. 2002; Iooss and Ribatet 2006) have tackled the global sensitivity analysis problem when one of the input parameters is not a simple scalar random variable, but a complex spatial random field.

**Acknowledgments** This work was supported by the MRIMP project of the “Risk Control Domain” which depends on CEA/Nuclear Energy Division/Nuclear Development and Innovation Division. This work is also a part of the Russian Research Center “Kurchatov Institute” Scientific and Technical Complex “Rehabilitation” project.

## References

- Box GE, Draper NR (1987) Empirical model building and response surfaces. Wiley, New York
- Box GE, Wilson KB (1951) On the experimental attainment of optimum conditions. *J Roy Statist Soc B* 13:1–38
- Campolongo F, Saltelli A (1997) Sensitivity analysis of an environmental model: an application of different analysis methods. *Reliab Eng Syst Saf* 57:49–69
- Devictor N, Bolado Lavin R (2006) Uncertainty and sensitivity methods in support of PSA level 2, SARNET report, SARNET-PSA2-P06 Revision 0, European Commission, 2006
- Downing DJ, Gardner RH, Hoffman FO (1985) An examination of response-surface methodologies for uncertainty analysis in assessment models. *Technometrics* 27:151–163
- Hastie T, Tibshirani R, Friedman J (2002) The elements of statistical learning. Springer, Berlin Heidelberg New York
- Helton JC (1993) Uncertainty and sensitivity analysis techniques for use in performance assessment in radioactive waste disposal. *Reliab Eng Syst Saf* 42:327–367
- Helton JC, Davis FJ (2003) Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. *Reliab Eng Syst Saf* 81:23–69
- Iooss B, Ribatet M (2006) Analyse de sensibilité globale de modèles numériques à paramètres incontrôlables. In: Actes des XXXVIIIèmes Journées de Statistique, Clamart, France
- Iooss B, Van Dorpe F, Devictor N (2006) Response surfaces and sensitivity analyses for an environmental model of dose calculations. *Reliab Eng Syst Saf* 91:1241–1251
- Jourdan A, Zabalza-Mezghani I (2005) Response surface designs for scenario management and uncertainty quantification in reservoir production. *Math Geol* 38:965–985
- Khalil A, Almasri MN, McKee M, Kaluarachchi J (2004) Applicability of statistical learning algorithms in groundwater quality modelling. *Water Resources Res* DOI 10.1029/2004WR003608
- Kleijnen JPC (1997) Sensitivity analysis and related analyses: a review of some statistical techniques. *J Statist Comput Simul* 57:111–142
- Kleijnen JPC (2005) An overview of the design and analysis of simulation experiments for sensitivity analysis. *Eur J Oper Res* 164:287–300
- Kleijnen JPC, Sargent RG (2000) A methodology for fitting and validating metamodels in simulation. *Eur J Oper Res* 120:14–29
- Ma HW (2002) Stochastic multimedia risk assessment for a site with contaminated groundwater. *Stochastic Environ Res Risk Assess* 16:464–478
- Maddalena RL, McKone TE, Hsieh DPH, Geng S (2001) Influential input classification in probabilistic multimedia models. *Stochastic Environ Res Risk Assess* 15:1–17
- Marrel A, Iooss B, Van Dorpe F, Volkova E (2006) An efficient methodology for modelling complex computer codes with Gaussian processes. *Comput Statist Data Anal* (submitted)
- McKay MD, Beckman RJ, Conover WJ (1979) A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 21:239–245
- Sacks J, Welch WJ, Mitchell TJ, Wynn HP (1989) Design and analysis of computer experiments. *Statist Sci* 4:409–435
- Saltelli A, Chan K, Scott M (eds.) (2000) Sensitivity analysis. Probability and statistics series. Wiley, New York
- Schapire R (1990) The strength of weak learnability. *Mach Learn* 5:197–226
- Sobol IM (2001) Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Math Comput Simul* 55:271–280
- Stone M (1974) Cross-validatory choice and assessment of statistical predictions. *J Roy Statist Soc B* 36:111–147
- Tarantola S, Giglioli N, Jesinghaus J, Saltelli A (2002) Can global sensitivity analysis steer the implementation of

- models for environmental assessments and decision making? *Stochastic Environ Res Risk Assess* 16:63–76
- Volkov VG et al (2003) Status of activities on rehabilitation of radioactively contaminated facilities and the site of Russian Research Centre “Kurchatov Institute”. In: *Proceedings of WM’03 Conference*, Tucson, Arizona, USA, February 23–27, 2003, WM Symposia Inc
- Volkov VG et al (2004) The First Stage of Liquidation of Temporary Radwaste Repositories and Rehabilitation of the Radwaste Disposal Site at the Russian Research Centre “Kurchatov Institute”. In: *Proceedings of WM’04 Conference*, Tucson, Arizona, USA, February 29 – March 4, 2004, WM Symposia Inc