Environ Monit Assess (2008) 137:261–273
DOI 10.1007/s10661-007-9758-3

Spatial evaluation of the risk of groundwater quality degradation. A comparison between disjunctive kriging and geostatistical simulation

E. Barca · G. Passarella

Received: 19 July 2006 / Revised: 24 March 2007 / Accepted: 6 April 2007 / Published online: 13 June 2007

Abstract In some previous papers a probabilistic methodology was introduced to estimate a spatial index of risk of groundwater quality degradation, defined as the conditional probability of exceeding assigned thresholds of concentration of a generic chemical sampled in the studied water system. A crucial stage of this methodology was the use of geostatistical techniques to provide an estimation of the above-mentioned probability in a number of selected points by crossing spatial and temporal information. In this work, spatial risk values were obtained using alternatively stochastic conditional simulation and disjunctive kriging. A comparison between the resulting two sets of spatial risks, based on global and local statistical tests, showed that they do not come from the same statistical population and, consequently, they cannot be viewed as equivalent in a statistical sense. At a first glance, geostatistical conditional simulation may appear to represent the spatial variability of the phenomenon more effectively, as the latter tends to be smoothed by DK. However, a close examination of real case study results suggests that disjunctive kriging is more effective than simulation in estimating the spatial risk of groundwater quality degradation. In the study case, the potentially 'harmful event' considered, threatening a natural 'vulnerable groundwater system,' is fertilizer and manure application.

Keywords Groundwater · Disjunctive kriging · Geostatistical simulation · Risk assessment

E. Barca (✉) · G. Passarella
Water Research Institute, IRSA – CNR,
National Research Council,
via F. De Blasio, 5,
70123 Bari, Italy
e-mail: emanuele.barca@ba.irsa.cnr.it

Introduction

Usually, the risk of degradation of a natural system is defined as functionally related to the vulnerability of the system itself and to the hazard associated to a specific dangerous event.

The vulnerability is defined as the capability of a specific system to defend itself against a dangerous event; the hazard is the probability that a specific event could occur in some part of the natural system considered (Varnes 1984).

The elements involved in the definition of vulnerability are sensitiveness, resilience, renewability and weak points. A system is sensitive, to a specific dangerous event, if the latter can easily change the status of the system. Resilience is a property which allows the system to return to its initial status after being subjected to damage. Renewability measures the capability of a damaged system to return to its initial status using artificial means (therefore the cost of remediation is a factor closely linked to this parameter). Finally, the presence of weak points in the structure of a natural system could produce dangerous effects in very wide

areas. When a hazard is studied in relation to an event, it is necessary to evaluate its intensity, spatial extension, duration, and, in addition, to find out the magnitude and the persistence of the effects that such an event could produce. This approach has been widely used in the past years by civil protection agencies in order to manage various risks, e.g. managing nuclear plants and wastes risks (Claiborne and Gera 1974), environmental risks (Lin et al. 2001, 2002) etc.

In some previous studies (Passarella et al. 2002) an a posteriori methodology, evaluating the probability of exceeding a threshold concentration value of a specific pollutant, was proposed as the quantification of the risk of water quality degradation associated with a generic spatial point of the considered aquifer. This definition of risk presents a couple of advantages with respect to the traditional definition. Firstly, the proposed approach sidesteps the need to evaluate the two components of the risk (namely, the vulnerability and the hazard) and allows the risk to be estimated directly by means of probabilistic tools. In addition, it allows the resulting estimated probability to be represented by means of risk classification maps that are indeed a very useful tool from a qualitative standpoint for the ordinary and extraordinary management of the water resource considered.

The probabilistic definition of the risk requires that, in order to evaluate the probability that certain critical events will occur in a given spatial point, the sampled variable has to be considered as a random variable (RV). Hence the solution of the problem depends on the recognition of its theoretical distribution model (probability density function, or pdf) and the estimation of its parameters (Ott 1995). Applying the above definition to a groundwater system involves selecting an arbitrary set of no-data points and defining a suitable pdf in every selected point. After that, to complete the description of the whole groundwater system, a probabilistic n-dimensional model, called random function (RF) is needed.

Matheron (1970) popularized the theory of RF applied to natural resource control. The RF theory applied to earth sciences has evolved into geostatistics. In particular, disjunctive kriging (Rivoirad 1994) and geostatistical simulation (Armstrong and Dowd 1993) are two methodologies which have been developed within this discipline.

In the present paper, a study based on values of nitrate concentration sampled in the groundwater system of the Modena plain (central Italy) is illustrated. The dangerous event considered is the spreading of fertilizers on the soil for agricultural purposes while the vulnerable system is, of course, the groundwater system.

This work will show a comparison between the two methodologies, simulation and disjunctive kriging, applied to the assessment of the 'spatial' risk over the whole study area.

Ranking the spatial risk values in a given number of classes, it is possible to draw maps representing, graphically, the considered groundwater system, distinguishing areas more or less compromised by nitrate pollution at the current monitoring season.

When a sufficient number of maps are available in time, spatial risk can be crossed with the risk 'trends' producing what we call the 'effective' risk classification (Passarella et al. 2002). The general objective of this paper is to compare disjunctive kriging and conditional simulations as basic methodologies for assessing the 'spatial risk.' Consequently, risk trends and effective risk classification have been neglected here, since they are assessed after the spatial risk and do not add anything to the paper results.

Working methodology

Theoretical background

The main purpose of applied earth sciences is the study of the spatial (and/or time-spatial) behavior of one or more distributed (or territorial) properties.

Geostatistics provides, for this purpose, a number of techniques based on the *Random Functions* (*RF*) theory. *RFs*, from a physical standpoint, represent the spatial laws of distribution of the considered properties; they are described mathematically by a set of *Random Variables* (*RV*), each of which is associated with a specific spatial position. The *RF* model is very complex because it depends on a knowledge of all the conjoint distributions taken among all the *RVs*. In general, unless some simplifying hypotheses (first and second order stationarity or intrinsic hypothesis) on those distributions are accepted, the information obtained from the sampling survey in the observed domain is not sufficient to rebuild the *RF* analytically.

Nevertheless, geostatistics offers a great number of reliable techniques (e.g.: *disjunctive kriging* and

geostatistical simulation) to reproduce *numerically* the spatial law of the considered property on the observed domain.

In the following sections the theoretical characteristics of the proposed methodology are reported giving particular emphasis to stochastic simulations; in fact, in a previous paper (Passarella et al. 2002) disjunctive kriging has already been applied for assessing groundwater risk of degradation and described in detail.

An overview of stochastic simulation

Let us consider a random variable X; the values of X are called *realizations* of the *RV*. Simulating an *RV* consists in generating randomly realizations whose mean and variance reflect those of the original *RV*.

A random function $Z(X_i)$ consists of a set of *RVs* for i=1,..n, thus a realization of a *RF* is a set of realizations of the n *RVs* components. Therefore, simulating a *RF* means generating a large set of realizations which reflect the characteristic parameters of the considered *RF*. In particular, a simulation process generating realizations that honor the measured values is called a 'conditional simulation.' One of the advantages in using simulation is that it can be used also as a spatial interpolator, that is, it is possible to assess estimates of an unknown variable for an arbitrary number of no-data points of the spatial domain. The estimation assessment is carried out in the following way: the simulation process is run for sufficient time to reproduce the frequency histogram of the sampled data, at the end of the process a set of realizations will be associated with each unsampled point. Averaging these realizations at each point of the domain, the estimation will be obtained for each no-data point. Many procedures are capable of simulating realizations of a one-dimensional *RF* with a known mean and variance. However, when these procedures are generalized up to 2 or 3-dimensions, they become extremely demanding from a computational standpoint.

In the present paper, the *Turning Bands* simulation method (*TB*) is considered. The *TB* method provides independent realizations of a gaussian *RF*. If it is to be unbiased, this method requires the normality hypothesis, and if the hypothesis is not verified, a transformation of data has to be carried out by means of the gaussian *anamorphosis*, as described in a previous paper (Passarella et al. 2002). The *TB* method created by Matheron (1970), consists, firstly, in breaking down the 3-D (or 2-D) simulations into an arbitrary

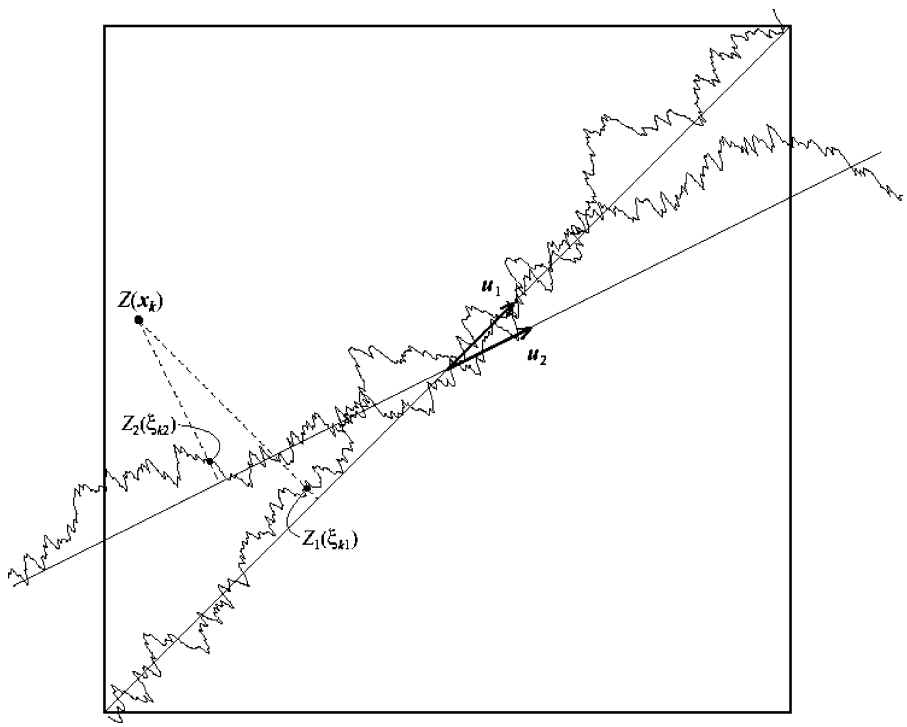

Fig. 1 The Turning Bands method: the contributions supplied by the n unidimensional stochastic processes $Z_i(\xi_i)$ are summed in a bidimensional process $Z(x)$ evaluated at point x_k (Thompson et al. 1989)

number of 1-D independent simulations and, then, in rebuilding the 3-D (or 2-D) domain starting from these simpler simulations (Fig. 1). Hence, from a computational complexity standpoint, the TB method is, by definition, equivalent to a 1-D simulation method.

The following formula defines the *TB* algorithm (Thompson et al. 1989):

$$Z(x_k) = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} Z_i(x_k * u_i)$$

Of course, in the *TB* method it is necessary to make some choices regarding the number of the turning bands and the number of simulations to be carried out. Obviously, the larger the turning bands are, the better is the estimation of $Z(\chi_k)$, thus minimizing the loss of information inherent in the transformation of a 3-D (or 2-D) domain as a sum of many 1-D domains.

Similarly, the estimation obtained using the simulations improves with the square of the number of simulations carried out, nevertheless, it is impossible to increase to infinity the number of simulations and/or of the turning bands (Journel and Huijbregts 1978).

Doing so, the computational load (time and resource) would grow to undesirable values. It is necessary, therefore, to find a compromise between the number of simulations to be carried out, the computational resources available, and, finally, the acceptable degree of uncertainty in the prediction.

Structural analysis

Neither method greatly influences the structure of the risk estimation procedure, because *DK* and *TB*, share similar working hypotheses and steps included in the stages that immediately precede (e.g. gaussian anamorphosis) and follow their application. Figure 2 compares the risk estimation methodologies based on the two geostatistical techniques, from the gathering of concentration data related to the quality parameter selected, to the mapping of the spatial risk of the groundwater qualitative degradation, defined as the probability of exceeding the concentration threshold assigned to the quality parameter.

An essential stage common to both the geostatistical methods consists in the *structural analysis* or *variography*. At this stage, the calculation of the

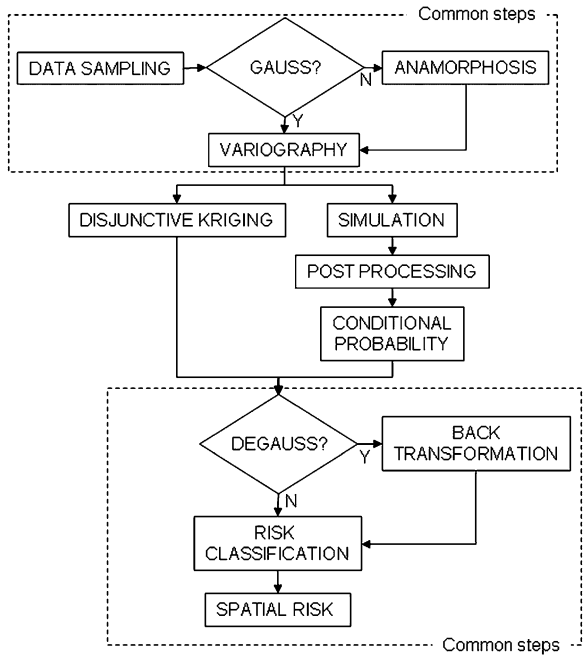

Fig. 2 Main structure of the methodology used for spatial risk assessment

experimental variogram (*EV*) and the determination of a variogram model (*VM*), which defines the spatial auto-correlation of the variable within the domain, are carried out. Once the *EV* has been calculated, determining the *VM* consists in selecting a particular function, chosen among some 'given' mathematical models, and estimating three parameters (range, sill, nugget) characterizing the function (Isaaks and Srivastava 1989; Krajewski and Gibbs 1993) so that the model best fits the experimental variogram.

Risk assessment and classification

The choice of the concentration threshold is the last critical step necessary to complete the risk estimation procedure using one of the geostatistical methods described above. This study considers nitrate concentration limits fixed by the Italian law for drinkable water (DLgs N. 152/1999); these limits are, respectively, 10 mg/l, as Guide Value (*GV*) and 50 mg/l as Maximum Allowable Concentration (*MAC*).

The final purpose of the proposed methodology is the definition of an 'effective' water quality degradation risk index resulting from the combination of time related ('trend') and spatial risk indices. Figure 3 shows the general structure of the methodology used

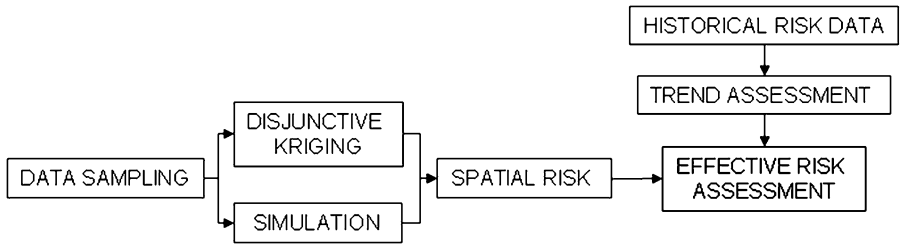

Fig. 3 General structure of the methodology used for assessing and classifying the effective risk

for assessing and classifying the effective risk; disjunctive kriging and geostatistical simulations are used for assessing the spatial risk, while the risk trend is assessed by using a non parametric method applied to a historical risk dataset. Crossing classes of spatial and trend risk, it is possible to define classes of effective risk.

Table 1 reports the effective risk classification scheme, which is based on five classes, ranging from 'very low' to 'very high' risk with regard to nitrate concentration. The five spatial risk classes, in the first column, are characterized by constant amplitude and represent the risk at the sampling time. The three trend classes represent the behavior of spatial risk throughout the previous sampling surveys, and can show an increasing, decreasing or steady risk trend index over time. The number and width of the spatial risk classes was chosen taking into account that synthesis and detail are inversely proportional. The resulting classification has to be useful for management purposes; consequently it has to be 'readable' for technicians and not only for researchers or scientists. In this frame, a loss of detail can be accepted if it is counterbalanced by a simplification of the classification, which can make it more understandable. When we attribute 'high risk' or 'very high risk' to a zone of the considered area, water authorities will assign a weight to these classes on the base of technical, political and economical factors.

Table 1 Effective groundwater quality degradation risk classes

Spatial groundwater quality degradation risk index	Trend index		
	Positive trend	No trend	Negative trend
0.0–0.2	Low	Very low	Very low
0.2–0.4	Medium	Very low	Very low
0.4–0.6	High	Medium	Low
0.6–0.8	Very high	Very high	Medium
0.8–1.0	Very high	Very high	High

Study area

In the present paper, the application of the described methodology to a domain of nearly 1,200 km^2, located in the high and middle Modena plain (central Italy), is considered (Fig. 4). This area, delimited on the southern side by the Tosco-Emiliano Apennines and crossed by the Secchia and Panaro rivers, has been subjected to many studies and research projects aimed at finding solutions to the freshwater supply problem (Visentini 1935; Colombetti et al. 1980; Paltrinieri and Pellegrini 1990; Barelli et al. 1990; Vicari and Zavatti 1990).

A detailed research related to the hydrogeology and intrinsic groundwater vulnerability of the studied aquifer (AA.VV. 1996) provided a large amount of digitalized thematic cartography concerning the hydrogeology, superficial depot textures, permeability, vulnerability and potential and real pollution of its groundwater. From a geological standpoint, the high and middle Modena plain, which lies between the Apennines to the south and the sedimentary Po river basin to the north, is characterized essentially by a mountain conoid system constituted by the main conoids of the Secchia and Panaro rivers intersected by the secondary conoids of minor streams. In the study area there is considerable agricultural activity, associated with the use of chemical fertilizers and the spreading of wastewater on the soil, producing serious pollution in the groundwater. For this reason, groundwater is accurately monitored. In particular, for the present study we considered the available values of nitrate concentration, sampled every spring and autumn between 1990 and 1996, relative to 90 sampling surveys carried out over 14 sampling

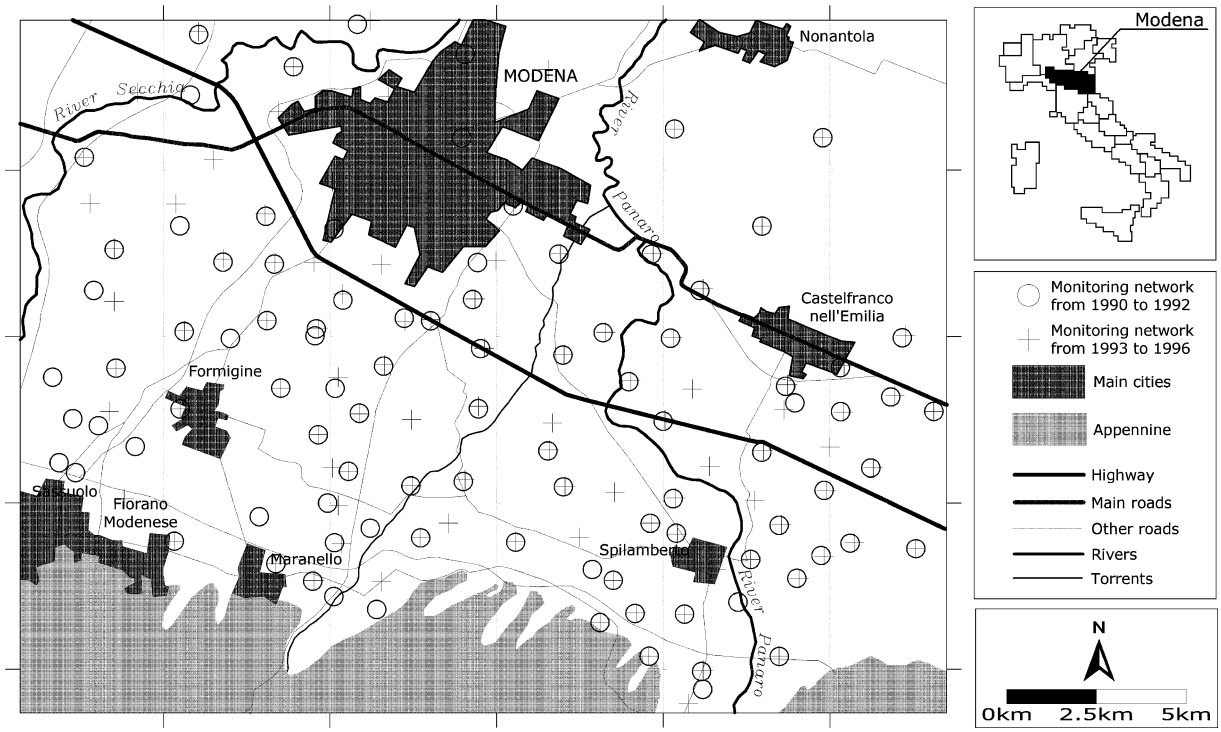

Fig. 4 Study area and monitoring networks

seasons (Giuliano et al. 1995). Nitrate concentration was chosen as the reference variable because it is a critical parameter affecting the groundwater in that zone. Figure 5 shows the behavior of the mean seasonal nitrate concentration during the time period considered. This figure demonstrates that nitrate concentration is persistent, in average, over the years. The other statistics (i.e. the first and third quartile, also reported in Fig. 5) confirm this behavior.

It appears that the autumn means increased during the observation period, reaching concentration values of just over 39 mg/l in the fall of 1996. The spring means, on the other hand, appear to have fluctuated around 34 mg/l. Both for autumn and spring values,

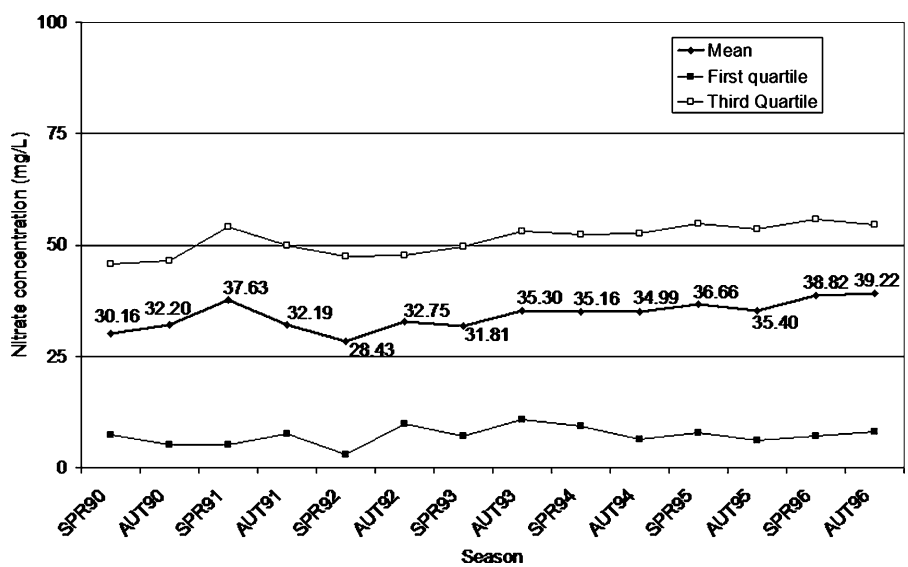

Fig. 5 Nitrate concentration behavior during the considered period

the maximum mean value was reached in 1996. The median, with the exception of spring 1990, was always far from the mean: this is a clear sign of the asymmetry of the distribution of the sampled concentration values. In fact, the histograms and graphic representations of cumulated sampled values all seem to be characterized by a marked asymmetry. The Kolmogorov-Smirnov test (KS), a tool allowing the normality of a given distribution to be evaluated statistically, confirmed this through rejecting the normality hypothesis (Passarella and Caputo 2006).

Results and discussion

Preliminary elaborations

A fundamental requirement for this type of study is a knowledge of the spatial behavior of the considered variable. (Isaaks and Srivastava 1989; Krajewski and Gibbs 1993). The software tool used in this study to carry out preliminary elaboration, kriging and simulations was ISATIS (Geovariances 2001).

The spatial behavior of nitrate concentration was determined calculating the related experimental variograms (*EVs*) for each of the considered seasons and fitting them using variogram models (*VMs*) (Passarella and Caputo 2006). The *VMs* used to interpolate the *EVs* are all authorized mathematical models, with the property that the resulting spatial covariance matrix is positive semidefinite (Journel and Huijbregts 1978).

The availability of a large amount of (time-space) data, though it causes data elaboration and management problems, allows considerations to be made on the time fluctuations of the spatial characteristics of the considered variable. Putting in sequence the obtained variogram representations for all the sampling seasons and for all the parameters, a storyboard is obtained, describing perfectly the time evolution of the parameters. As previously said, in the present study, nitrates show a steady behavior over time (Fig. 5), where the term steady is used in the sense of 'less disposed to modify in time their mean behavior in space.' This characteristic can be found again in the series of the variograms.

The experimental variograms of the measured concentrations were easily interpolated by means of an exponential model with sufficiently regular parameters. In general, the nugget did not exceed 25% of the total variability and the range was also sufficiently regular, assuming values around 6–8 km. Cross-validation, a procedure of pseudo-validation of the chosen model variogram, confirmed, generally, the correctness of the choice (Journel and Huijbregts 1978).

Table 2 shows the variogram models for each of the seasons considered in this paper and the results of the cross-validation.

Spatial risk index maps

After the preliminary elaboration, which was common to both the geostatistical techniques (Fig. 2), it was possible to apply each technique independently in order to assess the *spatial* risk of groundwater quality degradation with respect to the thresholds of 10 and 50 mg/l of the nitrate concentration. As described previously, the choice of the nitrate concentration threshold was made taking into account the values fixed by the Italian law for water for drinking purposes (DLgs N. 152/1999), 10 mg/l is defined as guide value (*GV*) and 50 mg/l as the maximum allowable value (*MAC*). Obviously, since the risk is the probability of exceeding the given threshold, it ranges between 0 (no risk) and 1 (maximum risk). Ranking such values, assessed in each of the node of the discretization mesh, in four classes, two maps were obtained for each season and threshold. In

Table 2 Variogram models and parameters per season and results of the cross validation

Season	Model	Nugget $(mg/l)^2$	Sill $(mg/l)^2$	Range (m)	Cross validation	
					Mean mg/l	Std.dev. mg/l
Spr-90	EXP	211	535	18,630	0.059	1.02
Aut-90	EXP	233	857	8,830	0.049	1.153
Spr-91	EXP	300	1246	6,820	0.029	0.923
Aut-91	EXP	58	974	13,970	0.046	1.012
Spr-92	EXP	181	719	5,961	0.053	1.074
Aut-92	EXP	261	569	5,292	0.051	1.031
Spr-93	EXP	193	646	6,945	0.044	1.065
Aut-93	EXP	253	803	4,830	0.055	1.035
Spr-94	EXP	264	880	7,140	0.021	0.769
Aut-94	EXP	260	1131	5,460	0.061	1.053
Spr-95	EXP	310	1058	7,059	0.057	0.988
Aut-95	EXP	221	1183	6,300	0.038	1.046
Spr-96	EXP	282	1453	6,834	0.041	1.044
Aut-96	EXP	616	924	8,288	0.045	0.886

general, the spatial risk maps resulting from the *DK* application show very high probabilities of exceeding the *GV* and a low risk of exceeding the *MAC*.

Carrying out the simulations, all the DK settings were, obviously, maintained to allow a comparison between the results of the two different approaches to be made. With regards to the specific settings for the simulations, 500 *trials* with 100 *turning bands* were used. The spatial risk maps resulting from simulations, as could be expected, are extremely similar to those provided by *DK*.

Figure 6 shows the maps related to the two concentration thresholds for the autumn of 1996, the last sampling season available. As mentioned above, just to improve the readability of the maps, four probability value classes were defined and a different color was associated to each class.

On large scale, the maps, resulting from the two different methods, show a similar spatial behavior. However, in general, a first, evident difference between the two spatial risk maps is that risk areas in the simulated maps are irregularly shaped while in the DK maps they are much more smoothed. This difference is surely due to the well-known property of variance minimization of kriging estimation (Isaaks and Srivastava 1989).

A more detailed analysis of the differences between *DK* and simulations risk values, cell by cell, revealed values generally nearly equal to zero but also outlined that *DK* tends to overestimate small values and underestimate large values as expected from literature (Goovaerts 1997). This important difference was, then, investigated using two different, but complementary checks. First, an analysis of the similarity between the risk maps resulting from the two methods was carried out using the cross-correlogram index (Stein et al. 1997) to compare their spatial patterns. Later, a normality test was carried out on the distribution of the differences between *kriged* and simulated risk values to establish if they could be attributed only to random fluctuations.

The first test was used to compare the two map patterns, where a pattern is defined as the spatial arrangement of representations for high and low values. Two maps display a similar pattern if the arrangements occur at the same positions on both maps.

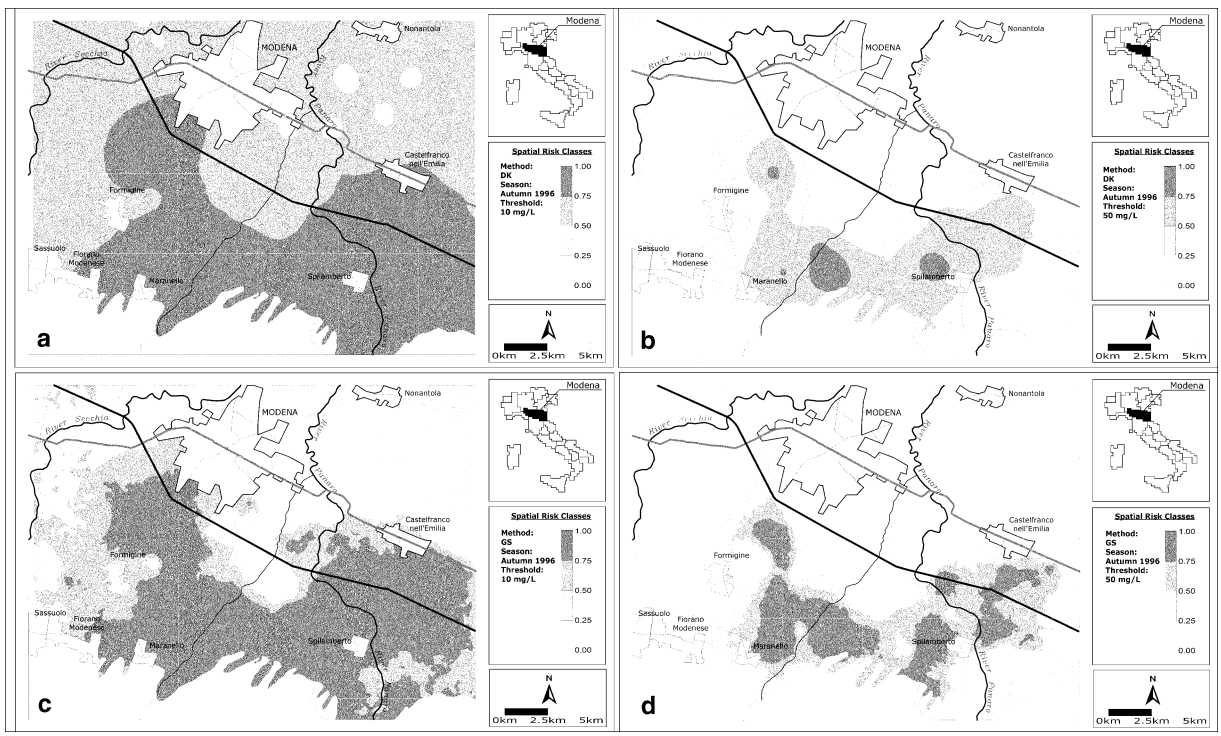

Fig. 6 Spatial risk maps for the 1996 autumn season: **a** DK method, 10 mg/l threshold; **b** DK method, 50 mg/l threshold; **c** TB method, 10 mg/l threshold; **d** TB method, 50 mg/l threshold

Stein et al. (1997) propose several measures to compare two different maps. In this paper, the cross-correlogram $\rho_{\text{Sim-DK}}(h)$ has been chosen:

$$\rho_{\text{Sim}-DK}(h) = \frac{E\left[z_{ij,\text{Sim}}, z_{i'j',DK}\right] - m_{\text{Sim}} m_{DK}}{s_{\text{Sim}} s_{DK}} \quad (1)$$

where $z_{ij,\text{Sim}}$ and $z_{i'j'DK}$ represent the values at locations i, j, and respectively i' and j' of the two maps, $h = \sqrt{(i-i')^2 + (j-j')^2}$ represents the distance between the two locations, E denotes the mathematical expectation, m_{Sim} and m_{DK} represent the population means and s_{Sim} and s_{DK} represent the population standard deviations. If patterns are completely similar, apart from a constant, $\rho_{\text{Sim-DK}}(0)$ will be equal to 1. To estimate $\rho_{\text{Sim-}DK}(h)$ from the available data the following equation can be used:

$$r_{\text{Sim}-DK}(h) = \frac{\sum_{i,j=1}^{N(h)} z_{ij,\text{Sim}} z_{i'j',\,DK} - m_{\text{Sim}} m_{DK}}{s_{\text{Sim}} s_{DK}} \quad (2)$$

where m_{Sim} and m_{DK} represent the sample means, s_{Sim} and s_{DK} represent the sample standard deviations and $N(h)$ is the total number of pairs separated by h. Results from the reported method showed a very good similarity between the two map patterns; in fact, $r_{\text{Sim-}DK}(0)$ was 0.93 and 0.94 for the thresholds of 50 and 10 mg/l respectively.

As explained above, this test, even though it is necessary in order to establish whether the two maps are similar, is not alone sufficient; in fact, high, positive values of $r_{\text{Sim-}DK}(0)$ can be obtained even if the pairs differ for a constant value. This means that if we want to prove a statistical similarity between the two maps we need also to test the distribution of the differences between simulated and *kriged* values, cell by cell. If the distribution is normal, then we can state the two maps are completely similar. The Kolmogorov-Smirnov test (*KS*), was used for assessing the normality of the difference distribution, and it rejected the normality hypothesis for all the considered seasons and both the thresholds.

As a result of the two above described tests on the risk values, it can be affirmed that the two datasets produced by simulations and *DK* do not belong to the same statistical population and, consequently, the two methods cannot be considered statistically equivalent: from a practical standpoint this means they are not interchangeable.

Let us consider the main statistics calculated from the datasets of differences between simulations and disjunctive kriging risk maps, for autumn 1996 and both thresholds (Table 3), to find some explanations of the result exposed so far.

This table is reported just as an example but it represents fairly the behavior of all the considered seasons. Firstly, it shows that mean, median and mode are always almost close to zero; that is to say, that on average and for both the thresholds, the behavior of the two methods seems similar, although not identical. However, considering the asymmetry, it is evident that the distributions of the differences are differently skewed; in particular, the negative values of asymmetry and sum suggest that simulations tend to underestimate the risk values for the lower threshold.

On the other hand, for the higher threshold, the asymmetry changes to positive; this indicates simulations tend to overestimate more frequently than *DK*; however the sum of the values, even diminishing significantly in absolute value, remains negative indicating that the overestimations of the *DK*, although less frequent, are far larger than simulations.

In fact, this confirms that the two methods are not identical but it is not sufficient to say which of the two methods is more suitable than the other. Nevertheless, observing the nitrate concentration distributions, we expected frequent high risk values when considering the lower threshold and frequent low risk values when considering the higher one, from both the methods. Figure 7 shows the frequency

Table 3 Main statistics calculated from the datasets of differences between simulations and disjunctive kriging risk maps, for autumn1996 and both the thresholds

Threshold		*Autumn 1996*	
		10 mg/l	50 mg/l
N		21,162	21,162
Missed		438	438
Mean	*mg/l*	−0.16475	−0.04975
Median	*mg/l*	−0.172	−0.07
Mode	*mg/l*	0	−0.065
St. Dev.	*mg/l*	0.152996	0.106549
Kurtosis		−0.89895	2.299918
Asymmetry		−0.08235	1.323454
Sum	*mg/l*	−3486.48	−1052.85
Minimum	*mg/l*	−0.615	−0.514
Maximum	*mg/l*	0.348	0.533

histograms of the simulated and *kriged* risk values for the two given thresholds. In particular, Fig. 7a, related to the lower threshold, shows that *DK* risk estimates correspond to what was expected more than simulations. In fact, the most frequent *DK* risk values are higher than 0.5, while simulated risk values are almost equally distributed among the five risk classes. On the contrary, considering the higher threshold (Fig. 7b), both methods confirm what was expected, i.e., high frequencies for lower risk values.

Even though these considerations seem to point towards the *DK* as the preferable method to be chosen when assessing the *spatial* risk, they prompted us to try a conclusive investigation. Figure 8 describes the results of this analysis. They show the behavior, during the considered seasons, of (1) the sample interval estimation of the probability of exceeding the given thresholds (95% confidence) compared with (2) the mean *kriged* and simulated risk values. Practically, the values (1) were assumed as reference points in order to compare the mean global behavior of the two methods. Obviously, this approach is absolutely independent from the spatial arrangement of the risk estimations.

Firstly, Fig. 8 confirms that the probabilities are rather high when related to the lower threshold (Fig. 8a) and are rather low considering the higher

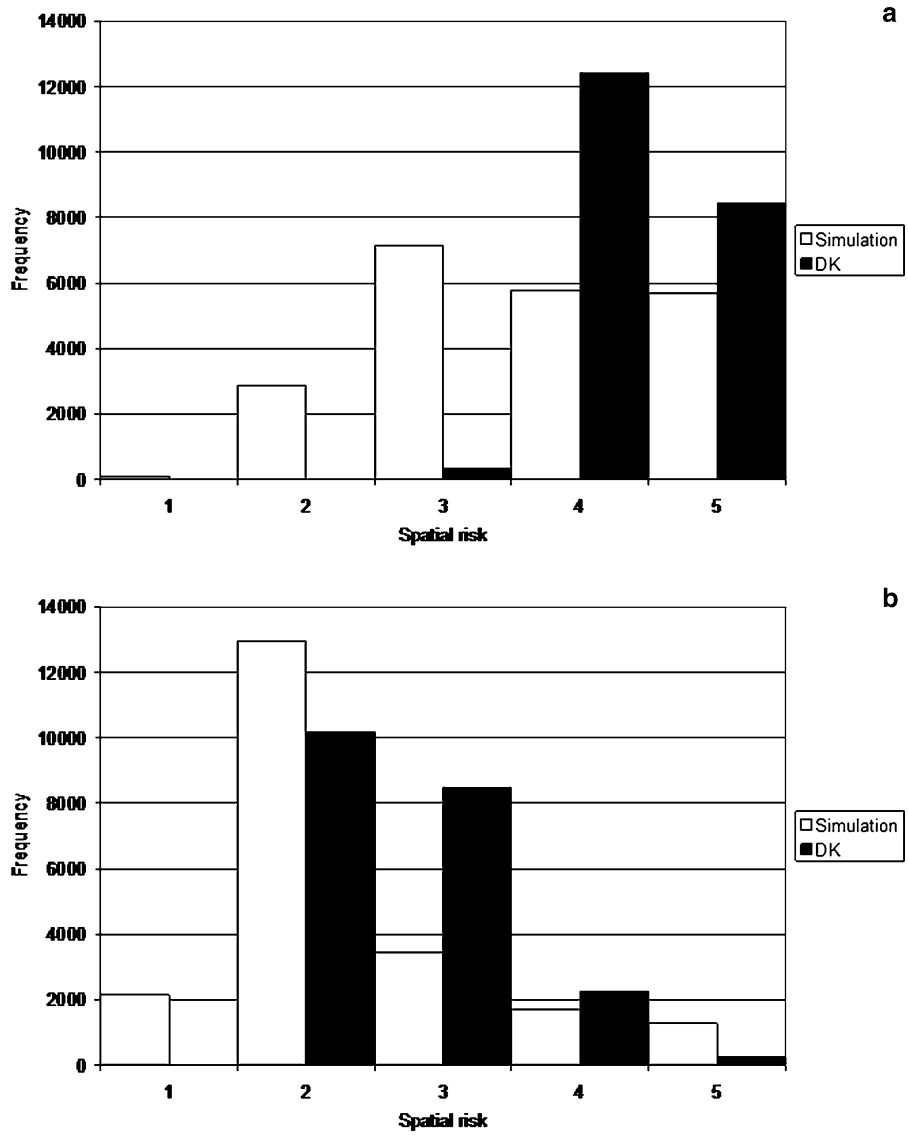

Fig. 7 Frequency histograms of the simulated and *kriged* risk values for the two given thresholds: **a** 10 mg/l threshold; **b** 50 mg/l threshold

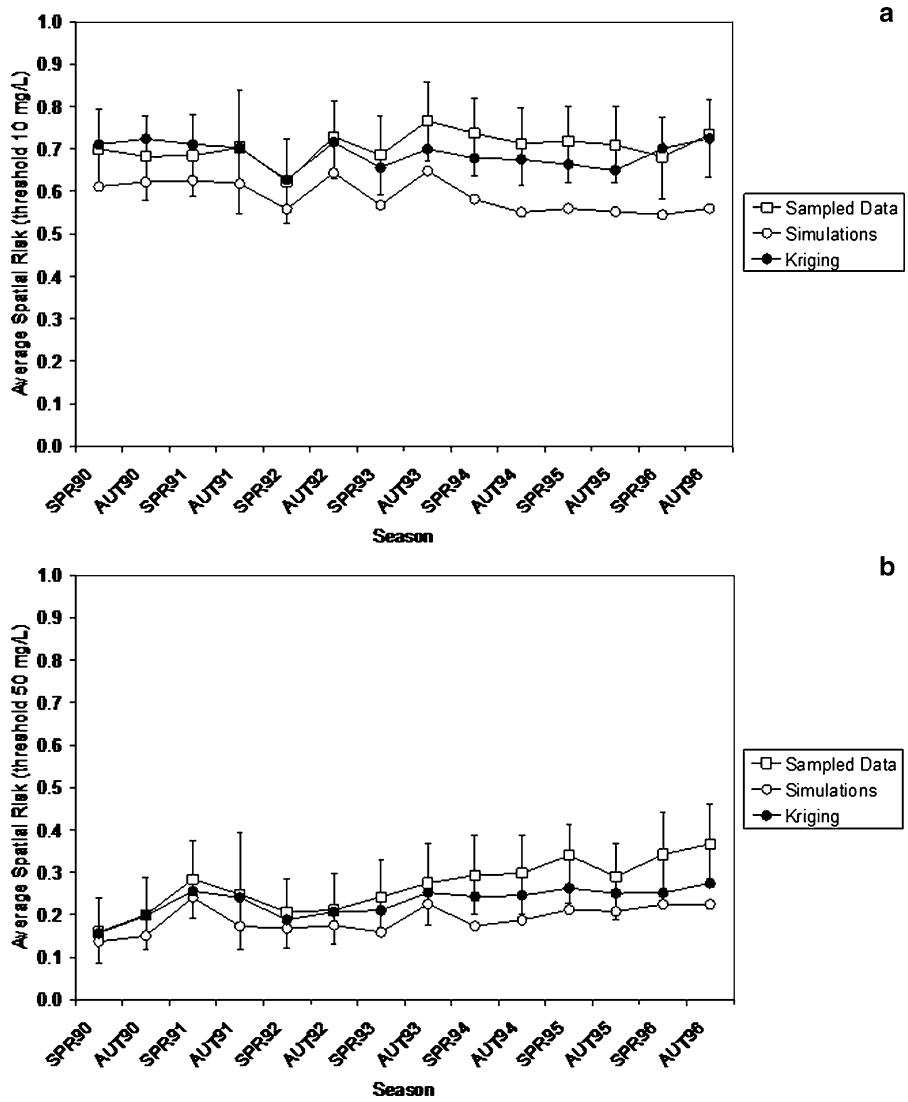

Fig. 8 Behaviour of the sample interval estimation of the probability of exceeding the given thresholds (95% confidence) compared with the mean *kriged* and simulated risk values, during the considered seasons. **a** High probabilities related to the lower threshold. **b** Low probabilities related to the higher threshold

threshold (Fig. 8b). Furthermore, on average, both the methods underestimate with respect to the reference values, for both thresholds. However, it is remarkable that mean *kriged* values always fall within the confidence interval defined for the reference values; on the contrary, simulation values are often outside that interval and always below the mean *kriged* values.

In conclusion, what is reported above confirms the opinion that *disjunctive kriging* is more effective than simulation in order to estimate the *spatial* risk of groundwater quality degradation. Although, in our opinion, the results reached can be considered almost exhaustive and conclusive, differences between the two methods and their significance merit further investigations. For example, increasing the number of simulation trials and of field observations could lead the two methods to converge, at least statistically.

Conclusions

In this paper, two different methods, conditional simulation and disjunctive kriging (DK), have been compared with the aim of using them in the assessment of the 'spatial' risk of groundwater quality degradation defined as the conditional probability of exceeding assigned thresholds of concentration of a generic chemical sampled in the studied water system. Spatial risk values can be ranked in classes, and related risk maps can be drawn outlining areas more or less

compromised by nitrate pollution at the current monitoring season in the considered groundwater system.

In a previous paper, spatial risk had been crossed with risk 'trends' producing the 'effective' risk classification (Passarella et al. 2002). However, given the aim of this paper, risk trends and effective risk classification have been neglected, since they are assessed after the spatial risk and do not add anything to the paper results. Hence, we have focused on spatial risk and, in particular on two geostatistical methods for its assessment: disjunctive kriging and stochastic simulations by turning bands method.

Both methods are based on a knowledge of the concentration of the considered parameter in a number of groundwater sampling points and provide risk maps for given thresholds of concentration. In this study, the Guide Value (10 mg/l) and the Maximum Allowable Concentration (50 mg/l) fixed by the Italian Law for drinking water were used as thresholds. Simulations and *DK* share a large part of their operative stages, nevertheless, they do not have the same computational load, in fact, simulation is much more demanding from a computational standpoint.

Nitrate concentration monitored twice a year for 7 years in a groundwater monitoring network made of around 100 wells were used to assess the spatial risk using both the methods presented. Several statistical and geostatistical tests were carried out in order to compare the results and try to evaluate which is the most suitable method.

A first coarse grain comparison between the results of the two methods showed that, both the methods give 'similar' representations of the qualitative state of the groundwater system, although, a first, evident difference between the two spatial risk maps is that risk areas in the simulated maps are more irregularly shaped than in the DK maps. This difference is surely due to the well-known property of variance minimization of kriging estimation.

Furthermore, a more detailed analysis of the differences between *DK* and simulations risk values, revealed values generally nearly equal to zero but also outlined that *DK* tends to overestimate small values and underestimate large values. This important difference was investigated using the analysis of the similarity between the risk maps resulting from the two methods (Stein et al. 1997). Afterwards, a normality test was carried out on the distribution of the differences between *kriged* and simulated risk values.

Results from the map comparison method showed a very good similarity between the two map patterns, the cross-correlograms being equal to 0.93 and 0.94 for the higher and lower threshold respectively. This test, even though necessary to establish whether the two maps are similar, is not alone sufficient. Hence, the Kolmogorov-Smirnov test (*KS*), was applied to the distribution of the differences between the results of the two methods to prove the statistical equivalence between the two maps. The normality hypothesis was rejected for all the considered seasons and both the thresholds: from a practical standpoint this means that simulations and *DK* are not interchangeable.

In order to establish which of the two methods is more suitable than the other, the behavior, during the considered seasons, of the sample interval estimations of the probability of exceeding the given thresholds were assumed as reference points in order to compare the mean *kriged* and simulated risk values. This analysis confirmed, on average, that both the methods underestimate with respect to the reference values, for both thresholds. However, it is remarkable that mean *kriged* values always fall within the confidence interval defined for the reference values; on the contrary, simulation values are often outside that interval and always below the mean *kriged* values.

In conclusion, the present study confirms the opinion that *disjunctive kriging* is more effective than simulation in estimating the spatial risk of groundwater quality degradation. Although, in our opinion, the results reached can be considered almost exhaustive and conclusive, differences between the two methods and their significance merit further investigations. For example, increasing the number of simulation trials and of field observations could lead the two methods to converge, at least statistically. In fact, geostatistical simulation emphasizes the spatial variability of the phenomenon which is, on the contrary, smoothed by the *DK*.

As reported by Raspa (2000), maps built by means of kriging estimation have a distinctive characteristic: they are smoothed, that is, they show a more regular spatial behavior with respect to the real phenomenon they describe, this smoothing effect influences all kriging methods and, consequently, the results of the applications. Working with simulation allows two specific parameters to be controlled: the number of realizations and turning bands. If this were done, it would probably allow the reliability of the estimation

to be improved, and thus the two methods would converge, at least statistically. Obviously, the computational constraints prevent these two parameters from exceeding reasonable values.

Independently from the chosen approach, the results of the first phase of the methodology represent, exclusively, the qualitative state related to the time when sampling took place (spatial risk). As said above, once the spatial risk map has been assessed for the current monitoring season, risk classification can be improved (effective risk) including the results of the risk trend analysis based on the previous spatial risk maps. This is beyond the aims of this paper and consequently has been neglected here. An example of such further assessment can be found in Passarella et al. 2002.

Acknowledgements The authors wish to acknowledge the courtesy of Dr. V. Boraldi and Dr. A. Zavatti of Emilia Romagna ARPA in providing data used throughout the paper.

References

AA.VV. (1996). Vulnerabilità naturale e rischio di inquinamento delle acque sotterranee nella Pianura Padana. In G. Giuliano (Ed.), *Rapporti dell'Istituto di Ricerca Sulle Acque* (pp. 145). Roma: Irsa – CNR.

Armstrong, M., & Dowd, P. A. (1993). (Eds.) *Geostatistical simulation*. London: Kluwer.

Barelli, G., Marino, L., & Pagotto, A. (1990). (Eds.) Caratterizzazione idraulica degli acquiferi. Studi sulla vulnerabilità degli acquiferi 2. In *Quaderni di tecniche di protezione ambientale – Protezione delle acque sotterranee* (pp. 15–26). Bologna: Pitagora.

Claiborne, H. C., & Gera, F. (1974). Potential containement failure mechanisms and their consequencies at a radioactive waste repository in bedded salt in New Mexico. In *Oak Ridge National Laboratory Report ORNL-TM-4639*.

Colombetti, A., Gelmini, R., & Zavatti, A. (1980). La conoide del fiume Secchia: modalità di alimentazione e rapporti col fiume (Province di Modena e Reggio Emilia). In *Quad. Ist. Ric. Acque* (p. 51). Roma: Irsa – CNR.

DLgs N. 152/99. Disposizioni sulla tutela delle acque dall'inquinamento e recepimento della direttiva 91/271/CEE concernente il trattamento delle acque reflue urbane, e della direttiva 91/676/CEE relativa alla protezione delle acque dall'inquinamento provocato dai nitrati provenienti da fonti agricole. Suppl. ord. G.U. del 29/5/1999.

Geovariances (2001) – ISATIS, Software manual, release 3.5.3. Geovariances Ecole des Mines de Paris, France.

Giuliano, G., Pellegrini, M., & Zavatti, A. (1995). (Eds.) Esempi di carte di vulnerabilità a scala regionale: La Pianura Padana, le province di Parma, Reggio Emilia e Modena. Studi sulla vulnerabilità degli Acquiferi. In *Quaderni di tecniche di protezione ambientale – Protezione delle acque sotterranee* (Vol.2). Bologna: Pitagora.

Goovaerts, P. (1997). *Geostatistics for natural resources evaluation*. New York: Oxford University Press.

Isaaks, E. H., & Srivastava, R. M. (1989). *An introduction to applied geostatistics*. New York: Oxford University Press.

Journel, A. G., & Huijbregts, Ch. J. (1978). *Mining geostatistics*. London: Academic.

Krajewski, S. A., & Gibbs, B. L. (1993). *A variography primer – Special Publication n.8*. Boulder, CO: Gibbs.

Lin, Y. P., Chang, T. K., Shih, C. W., & Tseng C. H. (2002). Factorial and indicator kriging methods using a geographic information system to delineate spatial variation and pollution sources of soil heavy metal. *Environmental Geology, 42*, 900–909.

Lin, Y. P., Chang, T. K., & Teng, T. P. (2001). Characterization of soil lead by comparing sequential Gaussian simulation, simulated annealing and kriging methods. *Environmental Geology, 41*, 189–199.

Matheron, G. (1970). The theory of regionalised variables and its application. In *Les Cahiers du Centre de Morphol. Mathemat. de Fointainebleau* (pp. 211). Paris: Ecole des Mines.

Ott, W. R. (1995). *Environmental statistics and data analysis.* Boca Raton, FL: CRC.

Paltrinieri, N., & Pellegrini, M. (1990). (Eds.) Comportamento idrodinamico dell'acquifero. Studi sulla vulnerabilità degli acquiferi, 2. In *Quaderni di tecniche di protezione ambientale – Protezione delle acque sotterranee* (pp. 9–14) Bologna: Pitagora.

Passarella, G., & Caputo, M. C. (2006). A methodology for space-time classification of groundwater quality. *Environmental Monitoring and Assessment, 115*(1–3), 95–117.

Passarella, G., Vurro, M., D'Agostino, V., Giuliano, G., & Barcelona, M. J. (2002). A probabilistic methodology to assess the risk of groundwater quality degradation. *Environmental Monitoring and Assessment, 79*, 57–74.

Raspa, G. (2000). Il ruolo della geostatistica nella modellizzazione ambientale. In *Aspetti applicativi delle tecniche geostatistiche alle acque sotterranee – Quad. Ist. Ric. Acque 114*. Roma: Irsa – CNR.

Rivoirad, J. (1994). *Introduction to disjunctive kriging and nonlinear geostatistics.* Oxford: Clarendon.

Stein, A., Brouwer, J., & Bouma, J. (1997). Methods for comparing spatial variability patterns of milliet yield and soil data. *Soil Science and Society American Journal, 61*, 861–870.

Thompson, A. F. B., Ababou, R., & Gelhar, L. W. (1989). Implementation of three-dimensional turning bands random field generator. *Water Resources Research, 25*(10), 2227–2243.

Varnes, D. J. (1984). *Commision on landslides and other mass-movements-IAEG. Landslide hazard zonation: A review of principles and practices*. Paris: The UNESCO Press.

Vicari, L., & Zavatti, A. (1990). (Eds.) Idrochimica, Studi sulla Vulnerabilità degli Acquiferi, 2. In *Quaderni di tecniche di protezione ambientale – Protezione delle acque sotterranee* (pp. 27–72). Bologna: Pitagora.

Visentini, M. (1935). *Ricerche sulle acque sotterranee nell'alta Pianura Modenese fra Secchia e Panaro*. Roma: Min. LL. PP. Serv. Idrograf. del Po.