Nordic Hydrology, 18, 1987, 121-150

An Evaluation of the Importance of Transmissivity, Head and Concentration Data for Contaminant Transport Modelling

Dirk Van Rooy

Institute of Hydrodrynamics and Hydraulic Engineering,
Technical University of Denmark, Copenhagen

The 3 basic data types of contaminant hydrology are examined by stochastic modelling of a groundwater contamination case. The stochastic transport model, which is of the Monte Carlo type, uses a numerical flow and transport model, and views transmissivity as a random autocorrelated field. A large set of transmissivity realisations is generated using the turning bands technique. Conditioning is done with regard to transmissivity, head and concentration observations. The unconditional approach assumes, explicitly, a stationary stochastic process of logtransmissivity. This is implicitly turned into a non-stationary process by the conditioning procedures. These use simple and universal kriging, and utilize the kriging uncertainties to determine subsets of realisations that are in agreement with the observations at a predefined confidence level. The approach followed allows quantification of the uncertainties of predicted head and concentration through space and time. Conditioning on head observations leaves large transport uncertainties. Conditioning on the transmissivity data has a more prominent effect. The single, most effective data type is the concentration data. Smallest transport uncertainties occur when all the data are simultaneously taken into account. The conditioning effect depends on the number and spatial configuration of the data. A trade-off between the stochastic and deterministic transport approach is suggested. In modelling terms this corresponds to a trade-off between advection and dispersion.

Introduction

During recent years the interest in subsurface hydrology has shifted from topics related to water resource investigations to problems of water quality. In particular,

quality aspects that are related to human activities interfering with the natural state of the groundwater attract the attention. Man-made groundwater contamination can generally be classified as stemming from plane- or point-sources. The former are usually characterized by low concentrations of the contaminants but they extend over horizontal areas comparable with the size of the affected aquifer; the latter sources are of limited areal extent as compared to the groundwater system on which they operate, but they usually contain high levels of the contaminants.

This study is confined to *point-sources* of contamination, examples of which are: spills of chemical substances at manufacturing plants, leakage from storage tanks or pipe-lines and, last but not least, leakage from *waste disposal sites* located at the land surface or in the subsurface. Waste disposal sites are one of the major known threats to groundwater systems: firstly, because the sources already exist all over the world and, secondly, because they were laid out during a time when the potential danger was not recognized. Therefore, in general no safe-guarding measures have been taken and removal of the sources is only possible at great expense since many sites actually occupy relatively large areas and/or volumes. Even in the cases where the damage done so far is limited, the seed of future problems has been sown, and this calls for an increased attention that will last for many years to come.

Subsurface hydrological problems are invariably of a complex nature. The main issue, however, is the *inaccessibility* caused by the complexity of the subsurface features. Any data set, no matter how large, will inevitably only represent a sample of the real-life system, and an evaluation of the total system behaviour is to be based on the available data. An important question then arises: How much useful information does an actual data set contain with regard to a particular system? And an important extension of this question is: How is this amount of information related to the type, the number and the spatial and temporal configuration of the data?

In terms of practical applications, these and similar questions become: Given a data set, what uncertainties can we expect in predicting the entire groundwater flow and contaminant transport through time and space? Or, given a field budget, what data should be collected to obtain the optimal description of the contaminated groundwater system? At what locations should observations preferably be made?

Such questions can be investigated by stochastic and geostatistical methods. These view the system variables and parameters as being partly governed by random components, and they aim to quantify the prediction uncertainties arising from the random behaviour. In the present study, the spatial distributions of the *hydraulic conductivity,* the *hydraulic head* and the *contaminant concentration* are viewed as uncertain quantities. These are the three basic data types encountered in groundwater pollution problems.

Another reason for casting a contaminant transport problem in stochastic terms

is the complex nature of the dispersive mechanism of solute transport. The hydrodynamic dispersion is mainly related to the heterogeneity of the flow velocity at *all scales* of the aquifer. An important factor in explaining this heterogeneity is the spatial distribution of the hydraulic conductivity. By viewing the conductivity as a stochastic parameter it is expressed in terms of probabilities, acknowledging the fact that its exact spatial distribution remains uncertain at the locations where it has not been measured. It should be noted that other factors contribute to the erratic spreading of a contaminant in an aquifer. One could mention the spatial distribution of the effective porosity, of the surface recharge a.o. In the present study these aquifer properties are assumed to be known deterministically.

Review

Numerous studies dealing with stochastic aspects of groundwater flow and solute transport have been published since the key paper by Freeze (1975) revived a dormant interest in the approach. Freeze's work dealt with one-dimensional flow systems only, but it is of general importance, since the author thoroughly formulates the stochastic framework. A survey of field data included in this study strongly supports a lognormal probability distribution function of hydraulic conductivity. The lognormal distribution of hydraulic conductivity is widely accepted nowadays. Freeze employed a Monte Carlo technique to investigate the predictive ability of one-dimensional flow modelling. This approach was extended to combined flow and solute transport modelling by Smith and Schwartz (1980). These investigators couple the Monte Carlo scheme to a numerical simulation model. They study the validity of the Fickian formulation of dispersion by computational experiments, and conclude that the classic theory, which assumes constant dispersivities multiplied by a known flow velocity, is not generally valid. In a subsequent paper Smith and Schwartz (1981a) continue along the same lines, but the focus is now on the uncertainty in transport predictions. The predictive capability is limited by partial sampling of a heterogeneous porous medium. They conclude that large uncertainties may occur when predicting the distribution of a tracer, and that those parameters capable of changing both the magnitude and the direction of flow are of primary importance.

Whereas the previous studies were concerned with flow and transport uncertainty from a modelling point of view, another school emphasizes the analysis of the hydraulic data using geostatistical tools. Delhomme's (1978) paper was a breakthrough in subsurface hydrology of the geostatistical approach, originally designed to solve problems of mining. Especially the technique of structural analysis by semivariogram recognition and the optimal interpolation method of kriging have become important research tools, nowadays. Delhomme (1979) combined the Monte Carlo method with a geostatistical analysis to investigate the flow regime of

an aquifer. The author concludes that due to the spatial variability of transmissivity large uncertainty with respect to the flow regime exist. Furthermore, only a small reduction of the uncertainty was obtained by conditioning the simulations on the measured transmissivity values. The author recognizes the need for conditioning the flow calculations on measurements of hydraulic head. Clifton and Neuman (1982) apply both transmissivity and head conditioning, using an inverse technique, to analyse the flow regime of an aquifer in Arizona. They conclude that conditioning on measured head values drastically reduces the head uncertainty and that well-calibrated deterministic simulations provide an adequate description of the *flow* regime of an aquifer. Transport was not treated in this study.

Of special interest to the present work are studies that investigate the *solute transport* forecasting capability as a function of a limited data set. The technique of incorporating field observations into the simulations is known as conditional modelling. By contrast, *unconditional* simulations only take the statistical properties of the data into account. Then, a location-invariant distribution of the stochastic variable exists everywhere throughout the domain, including those points where measurements were actually made. *Conditional* simulations aim to preserve both the stochastic properties, and the data at their measurement locations. Smith and Schwartz (1981b) combine the Monte Carlo technique with numerical solutions of the flow and transport equations. They describe the spatial structure of hydraulic conductivity by a nearest-neighbour relationship. The authors present a theoretical study and evaluate the effect of conditioning the transport simulations on conductivity values. The focus of this work is on tracer arrival times, and it is concluded that the conditioning effect is small, and thus large transport uncertainties remain. The modelling technique used by Smith and Schwartz (1981b) is very general in the sense that few restricting assumptions have to be made. As a working hypothesis a stationary stochastic process and a lognormal distribution of the hydraulic conductivity was assumed. Apart from that, the approach is applicable to nonuniform flow fields with irregular boundaries and with a arbitrary degree of variability of the conductivity distribution. Another approach was taken by Dagan (1982, 1984). This author adopts analytical solution methods that involve numerous simplifying assumptions with regard to the flow and transport configuration. The approach is based on first order perturbations and conditional probability functions, and the solution applies to infinite domains, uniform flow fields and small variances of logtransmissivities. Although of limited applicability in practical transport modelling, this study is instructive with regard to the basic behaviour of the transport processes. Dagan evaluates the variance reducing effect of conditioning the transport on point values of transmissivity and head. The main conclusions drawn from this work are that large concentration uncertainties exist when unconditional probability distributions of transmissivity are used. Incorporating head data has little impact on the concentration variances, while preserving transmissivity values may reduce the transport uncertainty considerably.

The Stochastic Simulation Approach

In the present study a numerical flow and transport simulation model is combined with geostatistical techniques. The link is obtained by a Monte Carlo approach. In this sense, the methodology adopoted here is a combination of the stochastic approach taken by Smith and Schwartz (1981b) and of the geostatistical techniques employed by Delhomme (1979).

The major goal of this work is to investigate the transport prediction uncertainty as a function of the available field observations of transmissivity, hydraulic head and contaminant concentration. In this respect, the approach followed here is an extension of the previous work. Another goal is to include hydrodynamic dispersion as a function of the variability of the transmissivity field, rather than using it as a lump parameter. All other aquifer characteristics such as porosity, boundaries, contaminant source strength and release etc. are assumed to be known in a deterministic sense.

Groundwater Flow and Solute Transport

The hydrodynamic laws governing groundwater flow in fully saturated porous aquifers are well-described in the literature, consult e.g. Freeze and Cherry (1979). The governing partial differential equation can be written as

$$\mathrm{div}(\bar{\bar{K}}\ \mathrm{grad}\ h) - W^* \equiv S_s \frac{\partial h}{\partial t} \tag{1}$$

where

$\bar{\bar{K}}$ – hydraulic conductivity tensor [m/s],
h – hydraulic head [m],
S_s – specific storage [m^{-1}],
W^* – source/sink term [s^{-1}],
t – time [s].

When dissolved chemical species are present in flowing groundwater, these will be transported through the porous matrix. The mechanisms responsible for the spreading of the solutes are of physical and chemical nature. The physical processes that control transport are advection and hydrodynamic dispersion. When only non-reactive species such as chloride, Cl^-, are considered, the physical mechanisms fully account for the migration of the contaminant. The transport of reactive solutes is beyond the scope of this study.

The governing partial differential equation of conservative contaminant transport in groundwater, known as the advection-dispersion equation, can be written as

$$\mathrm{div}(\bar{\bar{D}}\ \mathrm{grad}\ C) - \bar{v}(\mathrm{grad}\ C) = \frac{C'W^*}{n} = \frac{\partial C}{\partial t} \tag{2}$$

where

C = solute concentration in groundwater [mg/l],
$\overline{\overline{D}}$ = hydrodynamic dispersion tensor [m^2/s],
C' = solute concentration of source [mg/l],
$\bar{v}$ = seepage velocity [m/s],
n = effective porosity [–].

and

$$\bar{v} = \frac{\bar{q}}{n} \tag{3}$$

The hydrodynamic dispersion coefficient, formulated in the classical sense, is a function of the groundwater flow velocity and of the aquifer material. In two areal dimensions and assuming isotropy it can be characterized by a longitudinal and a transverse component

$$\begin{aligned} D_L &= \alpha_L \, |\bar{v}| \\ D_T &= \alpha_T \, |\bar{v}| \end{aligned} \tag{4}$$

where

D_L, D_T = longitudinal and transverse dispersion coefficients [m^2/s],
α_L, α_T = longitudinal and transverse dispersivity [m].

Konikow and Bredehoeft (1978) have developed presumably the most widespread numerical model. This model, which has been used in the present work, is based on a finite difference approximation to solve the 2-D flow equation and it solves the 2-D transport equation by a method-of-characteristics formulation.

The Stochastic Solute Transport Model

From a stochastical viewpoint the real transmissivity field, as it is in nature, is just one realisation out of an infinite number of possible realisations, denoted the ensemble. The transmissivity is the *independent* parameter. Based on its stochastic properties a large number of synthetic fields is generated. The set of these synthetic realisations or members represents a finite sample of the ensemble that obeys the prescribed stochastic properties. Hydraulic head, flow velocities and contaminant concentrations are the *dependent* variables in the sense that they are derived by applying the basic hydrodynamic laws. Thus, corresponding to each transmissivity realisation, realisations of head and concentration are obtained.

With regard to the stochastic process that governs the transmissivity, it is assumed that the lognormal probability density function is valid. Strong evidence in support of this assumption is available in the literature (see previously). Further, the validity of the ergodicitity principle is assumed, i.e. the statistical properties derived from the one member available are an acceptable approximation of the theoretical ensemble properties. It is worth recognising that the single realisation available, the aquifer, is only sampled at a limited number of locations. Thus

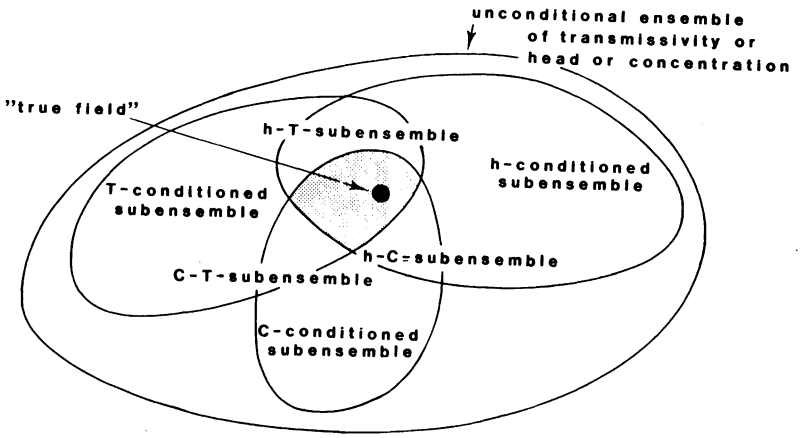


Fig. 1. Visualisation of unconditional ensembles and conditioned subensembles of transmissivity, head and concentration.

ergodicity is really assumed with respect to the available data set rather than to the entire realisation. In practice, one has no choice but to adopt the ergodicity principle as a working hypothesis for the stochastic approach. The Monte Carlo technique combined with a numerical transport model has been chosen, since this approach is not restricted by limiting assumptions that prohibit application to field situations, and it allows one to quantify the uncertainties of head and concentration through time and space. This is of primary importance in the present work.

In the unconditional simulation the stochastic process of logtransmissivity is *explicitly* defined as a stationary process with regard to the first two moments, i.e. the medium is statistically homogeneous. The conditioning proceeds conceptually by determining subsets of the transmissivity, head and concentration ensembles, see Fig. 1. The conditioning is applied to the ensemble that corresponds to the data type under consideration. Thus transmissivity conditioning is applied to the transmissivity ensemble, conditioning on head data is applied to the head ensemble and concentration conditioning to the concentration ensemble. In the conditional simulations it is *implied* that the governing stochastic process of logtransmissivity is non-stationary.

Unconditional Model

In the unconditional stochastic simulations only the statistical information of logtransmissivity is taken into account. This includes the mean and the autocovariance structure. The turning bands method is used to synthesize a large number of transmissivity fields, see appendix , and the same uncertainty exists everywhere throughout the domain. Each individual transmissivity member is then successively used as deterministic input to a solute transport model. By using elementary statis-

tics the distributions, expected values and variances of head and concentrations are calculated in every grid point, and in time to construct probabilistic breakthrough curves.

Transmissivity-conditional Model

In the T-conditional simulations both the structural information and the actual transmissivity values at their measurement locations are taken into account. Thus the T-conditional ensemble and realisations preserve the observed transmissivity values at their measurement locations in addition to obeying the prescribed stochastic properties. While disregarding the measurement uncertainties, no uncertainty exists at these locations. Also the transmissivity variablility is suppressed in the area surrounding the data location due to the imposed correlation structure. The influence of an observation gradually decreases with increasing distance, and its functional behaviour depends on the shape of the autocovariance function.

As transmissivity is the independent Monte Calo parameter, conditioning can be done using a direct technique. This technique was devised by Matheron (1973) and makes use of simple kriging, see appendix. Consider the stationary process $Z(x)$ and let the real value at a location x_r be $z(x_r)$, then

$$z(x_r) = z^*(x_r) + [z(x_r) - z^*(x_r)] \tag{5}$$

where z^* is the kriging estimate based on a data set consisting of m measured values $z_m(s_m)$. The term in square brackets, i.e. the difference between the true value and the kriging estimate, is the kriging error and is not accessible. However, it can be simulated by picking a substitute for z, called s, from an unconditional realisation with the same stochastic properties as those used in the kriging procedure. Then

$$z_s(x_r) = z^*(x_r) + [s(x_r) - s^*(x_r)] \tag{6}$$

where s^* is the kriging estimate based on the *unconditional* values at the m measurement locations, but using the kriging weights that were used for calculating z^*. It can be proven that Eq. (6) is a conditionally simulated field in the sense that its spatial characteristics are identical with the prescribed statistical properties of the unconditional field. Furthermore, since kriging is an exact interpolator, then when $x_r = x_m$ the bracketed term in Eq. (6) vanishes and $z^* = z_m$. Thus the measured values will effectively be retained at their measurement locations.

Head-conditional Model

Conditioning on the hydraulic head observations is done by determining the subset of *head members* that is in agreement with the head data. When a sufficient number of head data is available, the *universal* kriging technique (see appendix) offers the means to do so. Fortunately, in practice it is often possible to collect a substantial amount of hydraulic head data of good quality and at a reasonable cost. Since the

head field as a rule is governed by a non-stationary stochastic process, a geostatistical analysis involves the determination of a generalized covariance function. Kriging then provides a head map and a corresponding map of estimation variances. Delhomme (1978) suggested using the kriging standard deviations as a tool for hightlighting the parts of a groundwater flow domain on which to concentrate in the course of a traditional calibration procedure. Here, this idea is used to screen, successively, the individual members of the head ensemble.

Generally speaking, it makes sense to require that at a grid point the calculated head falls in a confidence interval defined by the kriging expectation value and the standard deviation at that location. Thus an acceptance criterion is defined as follows

$$\frac{\sum_{i=1}^{NG} |h_i - h_i^*|}{\sum_{i=1}^{NG} \sigma_i^*} \leq p \qquad (7)$$

where i refers to grid points and NG is the number of grid points considered. h_i and h^*_i are the calculated and kriged head values, respectively, and σ^*_i is the kriging standard deviation. Assuming a normal distribution, then, when $p = 1$ the 68% confidence level is applied and when $p = 2$ and $p = 3$ the 95% and 99% levels are used, respectively. The criterion can be applied selectively to individual grid points, to subregions of the domain, or to all grid points at once. E.g. only the model cells containing a head observation could be considered, or several clusters of grid cells could be included. Furthermore, different criteria can be applied to different regions of the domain demanding simultaneous fulfillment of all the criteria. When dealing with transient flow problems, the head values at different instants can be included straightforwardly. In a synthetical study of groundwater flow by Hefez *et al.* (1975), several criteria involving observed and calculated head values were examined and it was concluded, on empirical grounds, that best results were obtained with a criterion based on absolute deviations. Criterion Eq. (7) is based on the *absolute deviation* between the kriged and calculated head values. The criterion ensures a *plausible* solution, i.e. one that is in agreement with the data material at a predefined confidence level, but not an optimal solution. Thus each individual member of the *head* ensemble is screened, one at a time, using Eq. (7), effectively filtering out the subset of head members that obeys the observations. Once the acceptable head members have been identified, the corresponding concentration and transmissivity members are also known.

Concentration-conditional Model

In principle, the same conditioning procedure as applied to the hydraulic head could be used for concentration conditioning. This would involve determining the

subset of concentration members that fulfill the concentration acceptance criterion, formulated analogously to Eq. (7). In practical problems of groundwater contamination, however, often only a *small* number of concentration data is available. Some of them will represent background levels, while others sample the contaminated area. The data set of background levels is generally governed by a stationary stochastic process, whereas the plume definitely obeys a non-stationary process. Thus separate geostatistical analyses of the two data sets would be required. In most cases an insufficient number of high quality concentration data will be available to rigorously perform the exercise of structural analysis and to apply kriging. Still it is often possible to put forward a concentration range for a particular observation well. Let us assume that several Cl^- observations at different instants were made in a well. These values will usually show significant variation, which allows upper and lower limits of a concentration interval for that well to be defined. Further, the background concentration field is usually characterized by a range of values, adding uncertainty to the observed plume concentrations. Thus, by specifying upper and lower limits of concentration for the observation wells, the set of concentration members can be screened for consistency with respect to prescribed concentration intervals. The subset of accepted concentration members is labelled, and consequently the corresponding head and transmissivity members are hereby also known.

Case Study

The stochastic solute transport model was applied to a case of groundwater contamination. The site, known as the Løgtved landfill, is located on the island of Zealand, Denmark, about 100 km west of Copenhagen.

Aquifer Configuration and Waste Disposal Site

Figs. 2 and 3 display an areal and a cross-sectional view of the study area, respectively. The Quaternary sediments are about 60 m in thickness and are underlain by pre-Quaternary clay. Two aquifers, an upper unconfined and a lower confined, have been identified. They are separated by an aquitard consisting of glacial till. The aquifers are formed by glaciofluvial outwash material and consist of highly-permeable sand and gravel. The thickness of the upper formation is up to 20 m of which 7-15 m is fully saturated. The unsaturated zone ranges from 1.5 to 6 m. The thickness of the aquitard ranges from 10 to 20 m, but locally thicknesses up to 50 m have been observed. Since the clayey aquitard is of substantial thickness and no contamination has been observed in the confined aquifer, the present study is limited to the upper aquifer. Two streams that are directly embedded in the upper formation delineate the study area, see Fig. 2. This constitutes an area of roughly 2.75×4 km^2. Also shown are the wells that cover the area. For ease of further reference some of the wells have been named.

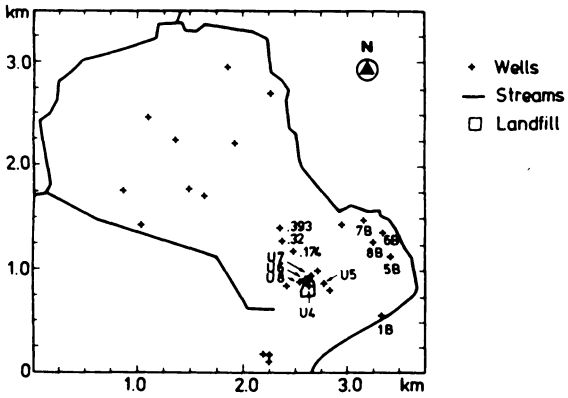


Fig. 2.
Areal view of the study area.

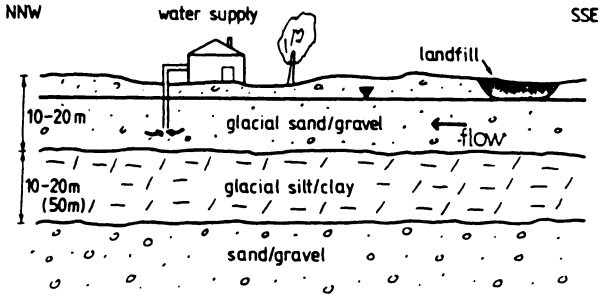


Fig. 3.
Schematic cross-sectional view of study area.

The landfill is situated in a partly abandoned gravel pit, excavated in the upper sand/gravel formation. The waste thickness ranges from 5 to 10 m and the areal extent is 30,000 – 50,000 m^2 (uncertainty with regard to the area exists, as the fill is covered with soil and vegetation; here the value of 50,000 m^2 has been retained). Operation of the site started in 1959, and the last recorded disposals are from 1982. The fill is *unprotected* i.e. there is no liner placed at the bottom of the waste, which is near the water table. Both municipal and toxic industrial waste products have been disposed.

Transmissivity Data

Transmissivity values were estimated from slug tests and specific capacity data. For more information about the interpretation of slug tests and about the data see Van Rooy (1986b, 1987). The geometric mean of the transmissivity is about 5.4×10^{-3} m^2/s. The geometric mean is valid when a lognormal probability distribution of the permeability and 2-D uniform flow can be assumed (Marsily 1986).

A semivariogram analysis of the transmissivity data was done. The transmissivity values were logarithmically transformed prior to the analysis. Since too few data were available, it was not possible to test the validity of the lognormal distribution. Fig. 4 shows the sample semivariogram of logtransmissivity together with the expo-

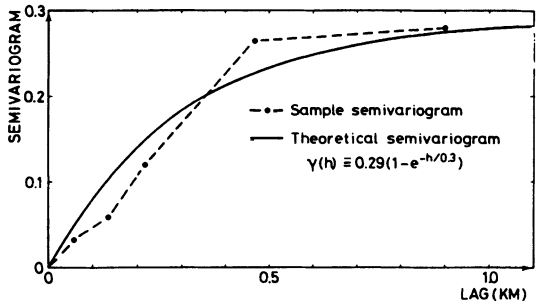


Fig. 4.
Experimental and theoretical semivariogram of logtransmissivity.

nential model fitted to it. An exponential rather than a simple linear model has been chosen, as the *experimental* variogram values are based on 11 observations only and are subject to large *uncertainty*. The exponential type is the one most frequently reported in the literature dealing with subsurface hydrology. Hoeksema and Kitanidis (1985) compiled and analysed data from N-American aquifers and found that the covariance of logtransmissivity can be represented by an exponential type function. In the present study other models such as the spherical type or even composite models would have been acceptable too, however, the limited data set does not justify the use of elaborated variogram functions. The coefficients of the exponential function were optimised using cross-validation. It was not possible to analyse rigorously for statistical anisotropy. The correlation length of the theoretical variogram is 300 m, which corresponds to a *range of influence* of about 900 m. The sill, i.e. variance, of the variogram of logtransmissivity is 0.29, which compares well with the field variance of 0.275. Delhomme (1979) reports large-scale horizontal ranges of influence from less than 1 km for alluvial aquifers and up to 20 km for limestone and chalk reservoirs. This author also reports logtransmissivity variances going from 0.7 for sandy/alluvial aquifers and up to 5.0 for limestone aquifers. Clifton and Neuman (1982), in studying an irregularly shaped area of about 45 × 15 km^2, operate with a range of 9 km. They report sill values of 0.37 and 0.74. Thus the range and sill values found here belong to the lower end of the scale of reported values. It should be kept in in mind that the data only cover the lower half of the study area and thus sample a fairly small area of about 1.5 × 1.5 km^2. Fortunately, the landfill and the plume are situated in this region. It is noted that the transmissivity values correspond to relatively small aquifer volumes. It is assumed, however, that they adequately represent *larger aquifer volumes*, but this is not easily verifiable. At present the proposed variogram model is assumed to be a true representation of the actual structure of the transmissivity field at the macroscopic scale.

Hydraulic Head Data

The hydraulic head of the upper aquifer was sampled during a two-day field campaign, April 18-19, 1985 (Terraqua 1985). These measurements represent a snap-

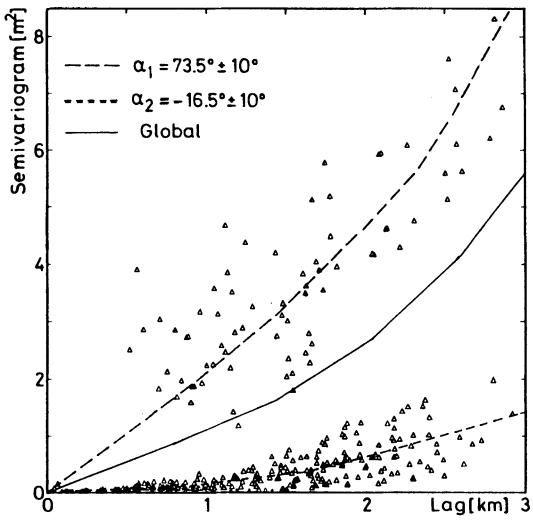


Fig. 5. Grouped and ungrouped variograms of head.

shot sample of the potential head surface. After scrutinizing the available data 28 head values covering the whole study area were selected for this study.

The geostatistical analyses of the head data clearly show the presence of a drift or trend (see Fig. 5). The underlying drift is indicated by the increasingly growing variogram in the direction of the drift, and this kind of behaviour is inevitably present in head data by the nature of groundwater flow. Two distinct clusters of points can clearly be distinguished in the *ungrouped* variogram values. They are an indication of statistical anisotropy, which happens to be another typical characteristic of head variograms. Consequently, a directional analysis where the data were grouped in two angle classes was applied. These two directions were chosen approximately orthogonal and parallel to the head isolines. The *grouped* global and directional variograms clearly reveal the statistical anisotropy and the drift. The non-stationarity of head implies that simple kriging can only be applied after the trend has been removed, as done by Virdee and Kottega (1984). Alternatively, one can use the techniques of universal kriging (see appendix) which is done here. The head trend was found to be of the order 1, and the following polynomial generalized covariance function was determined

$$K(h) = -0.1733\,|h| \tag{8}$$

using a nearest neighbourhood of 12 data points. Cross-validation was used to compare the performance of the different models and different sizes of neighbourhood were evaluated during the analyses. Kriging of the hydraulic head was done in a 23 × 33 discrete grid with a distance of 125 m between the neighbouring nodes. Fig. 6 displays the kriged head and standard deviations. The latter increase with the distance from the data locations, reflecting an increasing interpolation uncertainty.

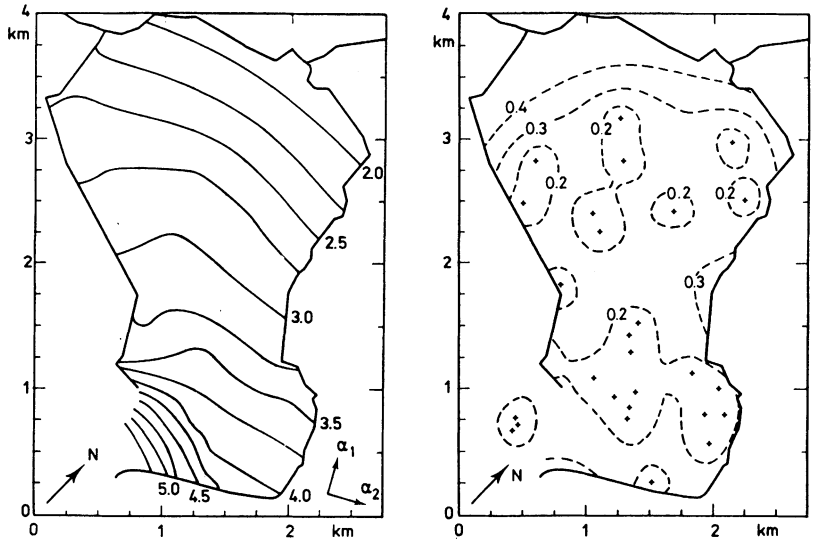


Fig. 6. Universal kriging of head.

Pollution Indicators

Water samples taken from a number of wells spread over the study area were analysed for various pollution indicators. These analyses show low values of organic compounds and irregular patterns of enhanced levels of inorganic constituents. Interpretation of the concentration data was very much obscured by the presence of other contaminant sources than the landfill. In large parts of the study area *pharmaceutical* sludge has been applied as a substitute for fertilizer. Many of the existing wells are located at *farming* settlements and anomalous concentrations could be due to local sources. Furthermore, a highway on which *de-icing salts* are applied during the winter crosses the study area. This constitutes a periodically activated line source of salts. Many of the *observation wells* were drilled for the purpose of small-scale water supply and are not adequate for contaminant sampling. E.g. well .32 (see Fig. 2) is apparently located in the contaminated area as the elevated concentrations of well .174 (91-204 mg/l) and .393 (184-212 mg/l) indicate. However, analyses of water samples from that well do not show any contamination, and this is presumably due to the shallowness of the well, namely 1.2 m of saturated penetration compared with 11.3 m and 11.4 m for the other two wells, respectively. Further well .174 and .393 are only screened over 2 and 3.5 m, and consequently the sampling is subject to *partial penetration* of the aquifer. In summary, no consistent picture of the extent and spatial pattern of the plume could be obtained. The background chloride concentration is estimated on basis of samples from 8 apparently unaffected wells, and these values span from 16 to 65 mg/l. Chloride values in well U4 range from 800 to 1,400 mg/l, with a centered value

around 1,000 mg/l. It is believed that this value reflects the leachate concentration. Cl^- concentrations in wells U6, U7 and U8 range from 98 to 247 mg/l. These wells fully penetrate the upper aquifer. The saturated thickness is from 9 to 11 m, and the well screens are implemented over the entire thickness. Temperature and electrical conductivity profiles were measured a few months after completion of the wells, allowing the water to achieve equilibrium with its environment (see Van Rooy 1986b). The water temperature is significantly enhanced to 12-14°C. The background temperature is 8-9° C. The electrical conductivity at reference level 25° C ranges from about 100 to 140 mS/m, with a background level from 60 to 100 mS/m. A few meters below the water table an increase of the order of 10 % was noted in all three wells, but the conductivity and temperature are fairly constant over the rest of the profiles. Thus the plume seems to occupy the lower three quarters of the aquifer at these locations.

Stochastic Model Study

In this section the results of the unconditional and conditional stochastic simulations of the Løgtved case of groundwater contamination are presented. The prefixes T-, h-, C- and combinations of them refer to conditioning on transmissivity, hydraulic head and concentration data, respectively.

Unconditional and T-conditional Transmissivity Ensembles

Two sets of 300 transmissivity realisations were generated. The unconditional set was generated using the *t*urning *b*ands *m*ethod, and the T-conditional set was derived from it using simple kriging. An isotropic and exponentially decaying autocovariance function, $C(h) = \exp(-h/0.3)$, corresponding to the semivariogram of Fig. 4, is imposed. The logtransmissivity fields are required to obey the $N(m,\sigma)$ distribution, with $m = -5.2$ and $\sigma = 0.54$, and were subsequently exponentially transformed to obtain the desired transmissivity distributions. The transmissivity values were generated in an equidistant net of 23×33 points. Delhomme (1979) states that the smoothing effect due to gridding can be neglected, provided that the size of the model cells is small in comparison to the range of the variogram. In the present case the ratio of the range to the cell distance of 125 m is 7-8. The quantity and quality of the present data set does not justify a denser net of points. The generated transmissivities can be viewed as block values, based on the assumption of large-scale representativeness of the T-observations previously proposed, or as punctual values that are based on the semivariogram of point observations. The latter view, adopted by Delhomme (1979), gives valid model outcome when the discretization is considerably smaller than the range. In the present study no further investigation of the effect of scale is done. A thorough discussion of different scales can be found in Dagan (1986).

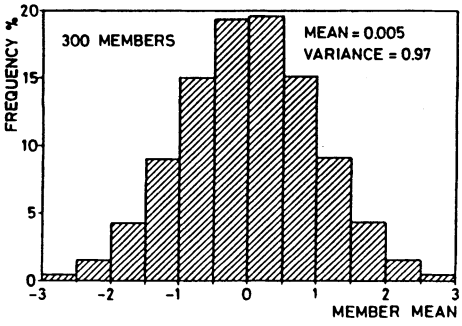


Fig. 7. Standardized ensemble histogram of logtransmissivity.

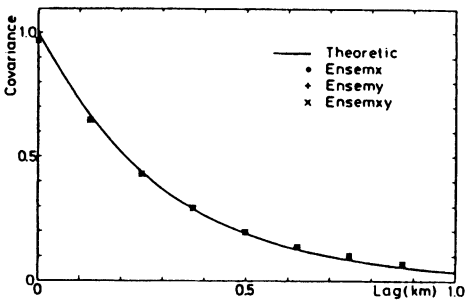


Fig. 8. Normalized ensemble autocovariance of logtransmissivity in x-, y- and xy-directions.

In Fig. 7 the standardized histogram of the TBM generated normal fields is shown. The mean is to 0.005 and the variance is 0.97, which is close to the prescribed values of 0.0 and 1.0, respectively. In Fig. 8 the prescibed and generated autocovariance structures are shown. The discrete values follow the exponentially decaying model closely. The unconditional transmissivity fields were conditioned with respect to 11 transmissivity values estimated from specific capacity and slug data. A more detailed analysis of the synthetic fields is provided in Van Rooy (1986b).

Stochastic Flow and Transport Simulations

All transmissivity realisations, the unconditional as well as the conditional, were successively used as single deterministic input to the numerical flow and transport model. All other model parameters, see Table 1, were kept unchanged during the simulations. These model parameters are partly based on data and partly on an extensive model study, see Van Rooy (1986b). Contaminant migration at the scale of the individual grid cells and larger is fully advection-based for each member, and dispersion is viewed as an ensemble phenomenon. Small-scale dispersion, that is,

Table 1 – Summary of main model parameters

o waste deposit area	47000 m^2
o source concentration of Cl^-	1100 mg/l
o source release fuction	step
o uniform background concentration	50 mg/l
o uniform effective porosity	0.20
o uniform and stationary surface recharge	4.44×10^{-9}m/s
o uniform saturated aquifer thickness	10 m
o longitudinal dispersivity	10 m
o α_L/α_T	3
o transport time	22 yrs
o constant head boundaries	

at scales smaller than the grid cells, is applied as a Fickian process. The applied value of 10 m for the small-scale longitudinal dispersivity corresponding to a transport distance of one cell, i.e. 125 m, was estimated from a compilation of field data by Lallemand-Barres and Peaudecerf (1978). A prestudy using slightly different parameters and using only 100 unconditional transmissivity realisations was presented by Van Rooy (1986a). In that study a grid-scale longitudinal dispersivity of 2 m was employed. A sensitivity analysis made clear that the transport results are rather insensitive to the dispersivity as long as it is smaller than 10 m, or so.

The outcome from the simulations are the sets of head, of steady-state plume concentrations and of breakthrough concentrations at a few observation points. Conditioning with regard to the 28 head observations involved the model cells containing a head observation, and the neighbouring cells. 4 individual subregions were chosen and the 68% confidence criterion was applied to each of them, demanding simultaneous fulfillment of all the criteria. These subregions comprise nodes with standard deviations smaller than 0.2 m, see Fig. 6. Nodes close to the boundaries were excluded. Concentration acceptance intervals of Cl^- were applied to 3 nodal locations, approximately coinciding with wells U6-U7-U8, .174 and .393. As the concentration observations are of low quality and are perhaps not representative for the plume, somewhat arbitrary criteria demanding Cl^- concentrations going from 125 to 225, 205 and 175 mg/l, corresponding to increasing distance from the fill, were applied. These criteria should be revised when more and better controlled data material becomes available. The set statistics, i.e. mean and standard deviation, were calculated at each nodal point. The standard deviation and its use in constructing confidence intervals is only meaningful when it can be assumed that the data are normal or near-normal distributed. Delhomme (1979) analysed head histograms which were found to be slightly skewed, but Marsily (1986) pointed out that the standard deviations still can be used as guiding measures for constructing confidence intervals. Here the primary interest is in *transport* rather than in flow, and the concentration distributions, in the ensemble sense, of the various simulations were examined by the X^2-test. The normal distribution was found acceptable at a 5% significance level in most cases. See Fig. 9 for some typical concentration histograms. Further, a sensitivity test was done in order to determine the minimum number of members needed to represent the ensemble statistics. This was done by calculating the statistical moments of successively smaller subsets, and comparing them with each other. It was found that when the number of members dropped below 40 no reliable unconditional concentration estimates could be obtained. With regard to the conditional simulations, when 30 members or more were taken into account, the calculated statistical moments changed little, fluctuating slightly up and down when more members were added. The smaller number of members needed in the conditioned cases expresses the smaller degree of variability caused by the conditioning process. Table 2 lists the various simulations with the number of members belonging to the corresponding

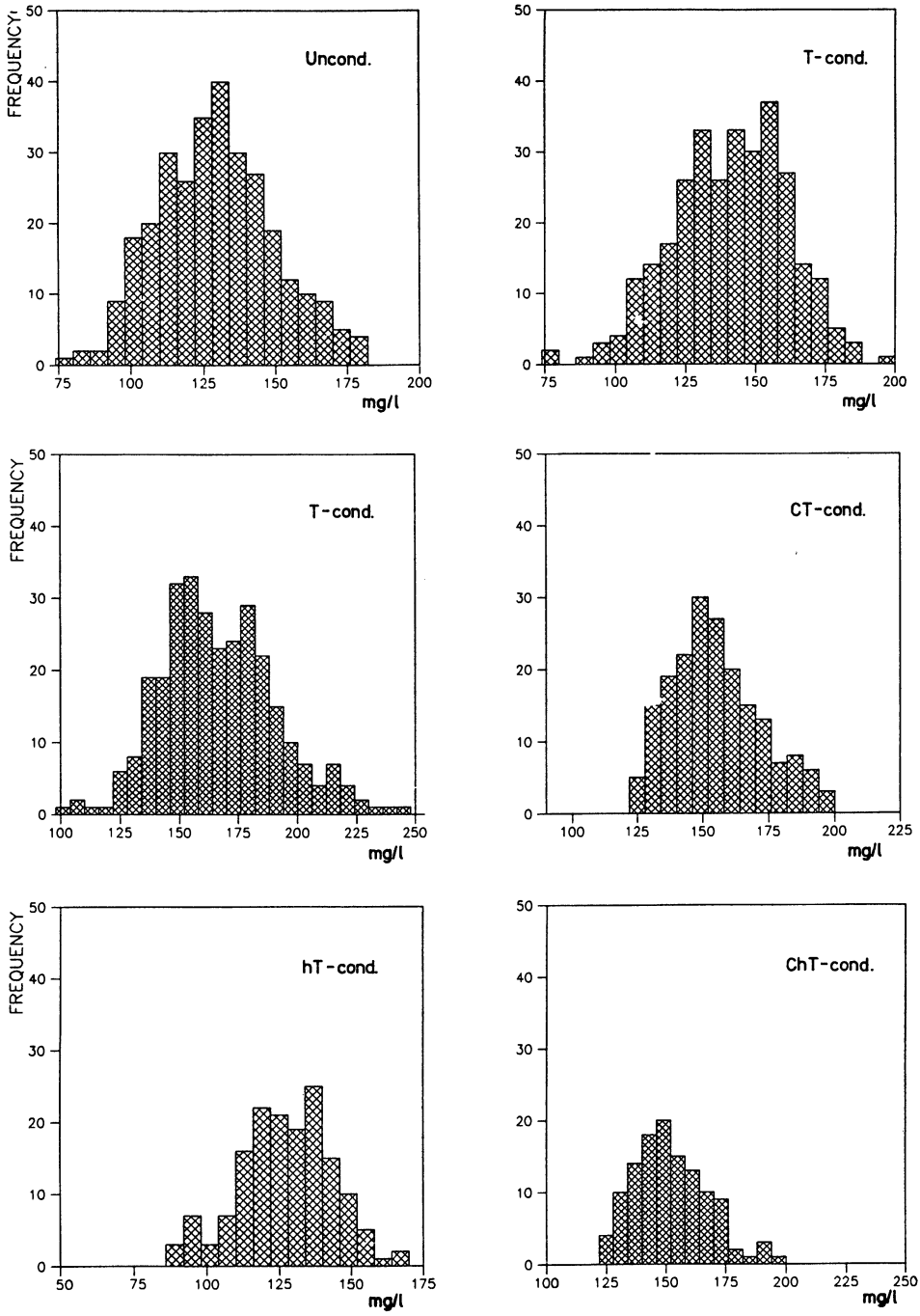


Fig. 9. Typical Cl^- concentration histograms.

Table 2 – Number of members

simulation	*members*
unconditional	300
T-condit	300
h-condit	80
C-condit	85
hT-condit	156
CT-condit	190
ChT-condit1	103

subensembles. All subensembles contain at least twice the minimum number of members required for an adequate representation of the statistical properties.

Conditioning Effect on Head Uncertainty

The conditioning effect consists of reducing the head variance throughout the flow domain. The conditioning effect is shown in Table 3 in terms of reduction of the head variance at a few observation points in the lower half of the study area. It is seen that the head data are the single, most effective data type with regard to reducing the head variance. Also the transmissivity data are of importance, and ultimately it is the combination of all the available data that gives the maximum variance reducing effect. The largest reductions, obtained with the ChT-conditional simulation, range from 3.7 to 4.9. The use of constant head boundaries makes the flow regime somewhat insensitive to the transmissivity distribution, and larger reductions are obtainable with more loosely defined flow boundaries. The present case study confirms that the head data are the single, most important data type with regard to macroscopic flow conditioning.

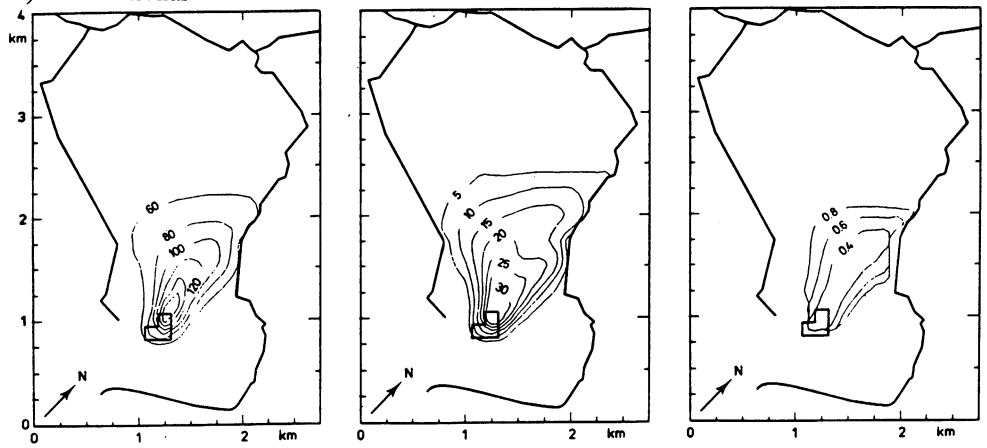
Conditioning Effect on Transport Uncertainty

The transport conditioning consists of reducing the concentration variance throughout the space and time domain. In order to get a full appreciation of the

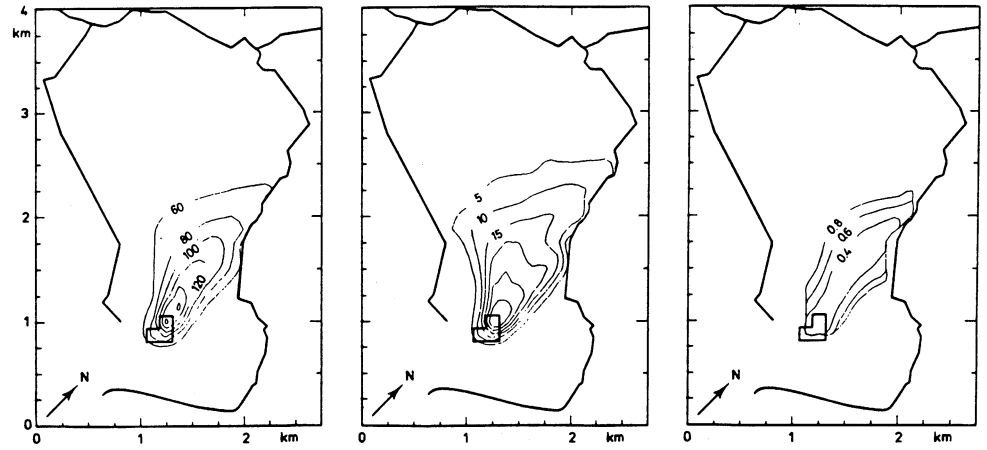
Table 3 – Reduction of the head variance relative to unconditional simulation

simulation	variance reduction at node			
	(11,10)	(13,10)	(13,11)	(13,13)
Uncondit	1.0	1.0	1.0	1.0
T-condit	2.0	2.7	2.4	1.1
h-condit	2.8	2.9	2.8	1.9
C-condit	1.6	1.7	1.5	1.2
hT-condit	3.3	4.2	3.4	1.6
CT-condit	2.5	3.3	2.6	1.3
ChT-condit1	3.7	4.9	4.1	1.7

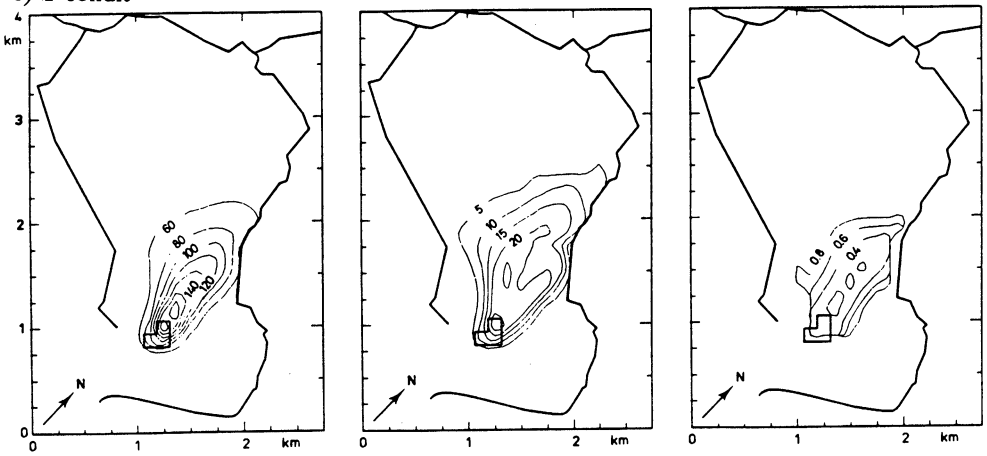
a) unconditional



b) h-condit



c) T-condit



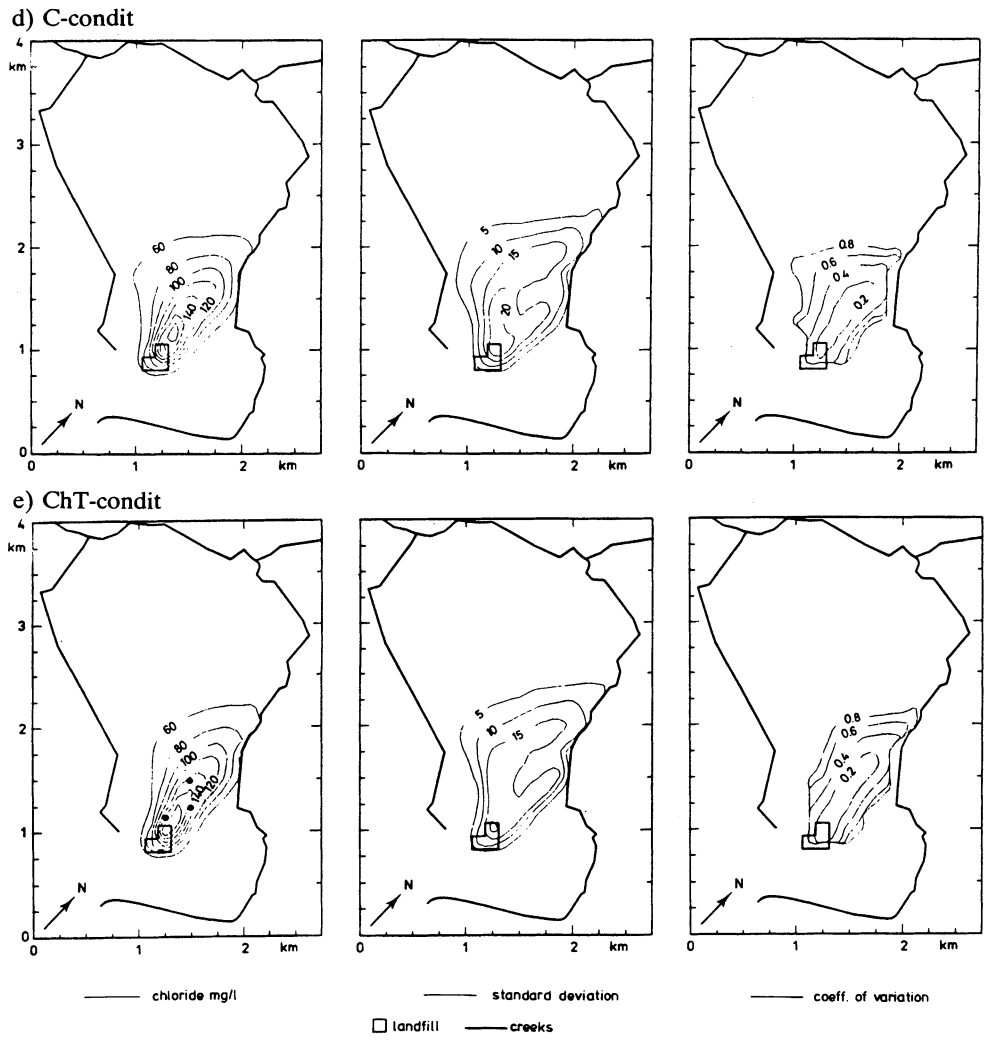


Fig. 10. Stochastic transport results: expected Cl^- plume, standard deviation, coefficient of variation. a) unconditional b) h-condit c) T-condit d) C-condit e) ChT-condit.

stochastic simulations and of the conditioning effect in particular, the results are expressed by three measures, namely expected concentrations, standard deviations and coefficients of variation. These are presented as maps of isolines in Fig. 10.

The *unconditional* plume of Cl^-, Fig. 10a, exhibits substantial transverse spreading. The standard deviations are largest along the center of the plume and decrease towards its outer boundaries. The coefficient of variation behaves in a reversed manner. Looking at the *h-conditional* model, Fig. 10b, it is noticed that the plume has a distinctly different shape. The plume is narrowed down and stretches over a

larger distance in the longitudinal transport direction. The map of the concentration standard deviations shows somewhat reduced levels. When *T-conditioning,* Fig. 10c, is applied, the plume appearance is similar to the previous one with regard to its transverse spreading. Contaminant migration along the longitudinal plume direction is reduced, and the spatial distribution of the standard deviations is considerably affected. The conditioning effect, however, is most clear from the concentration coefficient of variation, where three subareas of levels below 0.20 appear on the map. Continuing to the *C-conditional* simulation, Fig. 10d, the expected plume resembles the unconditional one in its general shape, though the concentration gradients are considerably larger. The standard deviations and especially the coefficients of variation are substantially reduced, particularly along the longitudinal plume axis. The smallest uncertainties are obtained when all the data are taken into account, i.e. the *ChT-conditional* model, Fig. 10e.

Some of the phenomena pointed out above are clarified in Table 4 that shows the coefficient of variation averaged over different plume sections defined by plume concentration intervals. The reduction factor, i.e. the average coefficient of variation relative to the unconditional simulation, is shown in bold-face. It is the *tendencies*, rather than the absolute figures, that are of importance. The average coefficient of variation for the unconditional simulation ranges from 0.78 for the plume as a whole to 0.33 for the concentration levels larger than 120 mg/l. The reduction obtained by the head data only is small at all concentration levels. The effect of conditioning on transmissivity data is much more prominent and differentiates with respect to the various sections of the plume. Its effect is largest at the central part, more particularly at the levels larger than 120 mg/l. The same notion applies to the C-conditional model for which the levels of 100 mg/l and more are most affected by the conditioning effect. It is seen that the concentration data are the single, most

Table 4 – Concentration coefficient of variation and its reduction for various simultions and plume sections; areal extent of the plume.

	Cl^- plume or part of it (mg/l)								cells occupied	
Simulation	global		⩾100		⩾120		100-120		global	⩾100
Unconditional	.78	**1.0**	.35	**1.0**	.33	**1.0**	.37	**1.0**	95	22
h-condit	.75	**1.0**	.33	**1.1**	.31	**1.1**	.35	**1.0**	89	20
T-condit	.72	**1.1**	.26	**1.3**	.22	**1.5**	.35	**1.0**	80	21
C-condit	.66	**1.2**	.22	**1.6**	.19	**1.7**	.30	**1.2**	85	23
hT-condit	.68	**1.1**	.25	**1.4**	.20	**1.7**	.34	**1.1**	79	21
CT-condit	.66	**1.2**	.20	**1.8**	.18	**1.8**	.30	**1.2**	80	22
ChT-condit1	.61	**1.3**	.19	**1.8**	.17	**1.9**	.25	**1.4**	80	21
Cht-condit2	.56	**1.4**	.18	**1.9**	.16	**2.1**	.25	**1.5**	79	22
Cht-condit3	.50	**1.6**	.16	**2.2**	.14	**2.4**	.21	**1.7**	79	22

bold-faced figures are reduction factors; see also text

effective conditioning variable. It is the ChT-conditional simulations (three ChT-conditional models are listed, see later) that give the best results. It must be remembered that the reduction is expressed *relative to the unconditional case.* This explains partly the small reduction obtained by the head data, as they are an important ingredient of the deterministic model set-up. Concludingly it is remarked that for the ChT-conditional cases the average reduction of the coefficient of variation at the central part of the plume is of the order 2. In terms of concentration variances, the numerical output revealed variance reductions of the order 4-10.

The number of grid cells affected by leachate from the contaminant source drops from 95 for the unconditional simulation to about 80 for the conditional ones. This decrease of the spatial extent of the plume corresponds to the previously observed reduced spatial spreading of the migrating solute. With the size of one cell being $.125 \times .125$ km^2, the aquifer area affected by leachate is about 1.25 km^2. The central area of the plume, characterized by nodal concentrations of > 100 mg/l, occupies about 22 grid cels and differs little from one simulation to the other.

The areal extent over which a concentration observation exerts influence when applied to a conditional transport simulation was examined for three ChT-data configurations. These cases have been denoted ChT-condit1, ChT-condit2 and ChT-condit3 in Table 4. The first and the third simulation are based on 3 concentration data, while the second includes only 2 data points. The difference is the spatial arrangement of the data: in the ChT-condit1 model the distance from the landfill to the most remote point is about 0.6 km, while this is more than 1 km for the other two cases. The expected plumes were very similar in all 3 cases, but the uncertainties are significantly smaller for the ChT-condit2 and -condit3 cases.

Probabilistic Breakthrough Curve

In Fig. 11 the expected breakthrough curves in well .147 for three different simulations are shown. Also shown are the 95% confidence envelopes. The background concentration is equal to 50 mg/l. As can be seen, both the expected values and the uncertainties are affected by the conditioning process. In the unconditional case the 95% confidence interval leaves room for unacceptable large uncertainties. The smallest uncertainties are found for the ChT-conditional simulation, in the transient as well as in the steady-state part of the curve.

Discussion and Conclusion

The effect of conditioning solute transport simulations on transmissivity, head and concentration data has been examined for a practical case of groundwater contamination. The unconditional simulation, which is used as a reference, is based on a randomly behaving but autocorrelated transmissivity field whose logarithmic trans-

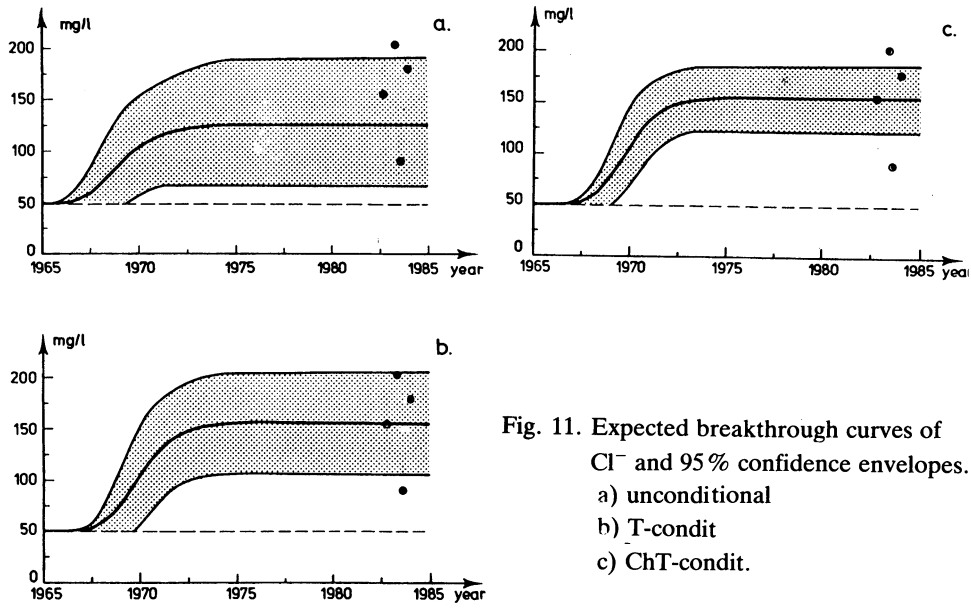


Fig. 11. Expected breakthrough curves of Cl^- and 95% confidence envelopes.
a) unconditional
b) T-condit
c) ChT-condit.

form is governed by a stationary normal stochastic process. This results in the largest transport uncertainties among all the simulations done. When the simulations are conditioned on the data, the transport uncertainties are reduced by different amounts, depending on the number, type and spatial arrangement of the data.

It was found that when the simulations are conditioned on the head data only the *head uncertainty* is reduced substantially, but the *transport uncertainty* is little affected. The relative concentration uncertainty of the main body of the Cl^- plume was of the order 35%. This implies that very different spatial transport patterns may exist. Generalizing this notion means that practical model studies of solute transport that only employ calibration against head data should be viewed with a sound scepticism. It is noted that the variance reducing effect is a relative measure and depends on the model set-up, in particular on the boundary conditions. However, the important notion is that even when the measured head is fairly well reproduced by a simulation model, the transport uncertainty due to the remaining head uncertainty and to the unknown spatial distribution of the hydraulic conductivity may still be very substantial. But there is absolutely no doubt that an adequate knowledge of the head distribution is of fundamental importance to any groundwater study. In summary, this means that the head observations are necessary but insufficient for predicting the migration of a contaminant in a groundwater system. Conditioning the simulations on the transmissivity data resulted in a much more prominent reduction of the transport uncertainty. But the single, most effective data type with regard to transport conditioning is the concentration data. These findings confirm the intimate relationship that exists bet-

ween the migration pattern of a solute and the spatial distribution of the hydraulic conductivity. Not surprisingly the smallest uncertainties are obtained when alle the data are simultaneously taken into account. The reduction of the concentration variance of the main body of the plume (> 100 mg/l) was locally up to a factor 10. The reduction of the average relative uncertainty was of the size 2-3, and it varies throughout the domain.

It was shown, quantitatively, that the spatial arrangement of the data relative to the location of the contaminant source and to the plume is of utmost importance. Concentration observations that are made in the immediate vicinity of the source have only limited value in terms of reducing the transport uncertainty further downgradient. Observations that are made at some larger distance from the source have a larger global conditioning effect and affect particularly the stretch upgradient from the data point.

With respect to the stochastic process that governs the logtransmissivity distribution, the conditioning procedure turns it from a stationary into a non-stationary process. This is immediately clear in the case of conditioning on the transmissivity data. When applying concentration conditioning, only those spatial transmissivity constellations that promote contaminant spreading in agreement with the concentration observations are recognized as being proper realisations of the governing process. This process is in general non-stationary, but in the approach that has been followed here a specification of its nature is not required.

The conditional simulations show that the solute migration is described in an increasingly unambiguous sense when more and different types of data are taken into account. Thus, when the degree of conditioning is increased, the transport gradually becomes more deterministic in its nature. In terms of the two physical transport mechanisms, advection and dispersion (minus diffusion), it is the advective component that is related to the deterministic aspect, and it is the dispersion that provokes the stochastic approach. In its essence, the mechanical mixing of the solute is related to the velocity heterogeneity, and it is the inability to collect sufficient deterministic information about this heterogeneity that creates the need for resorting to a dispersive transport term in a *modelling* situation. In this respect it should be noted that a large spreading of the solute is not equivalent to a large dispersion. It is the extent to which this spreading is known in a deterministic sense or not, i.e. by advection or not, that determines the nature of the transport description. Thus, in a practical modelling situation one should strive for a maximum advective description of the transport, utilizing *all* the available data and consequently the need for resorting to dispersion as a transport mechanism will be reduced. In this view, one could speak of the data-dependency of the dispersive mechanism of solute spreading. Putting this perception of dispersion into a historical perspective, the following scenario can be recognized:

Dispersitivity was originally thought of as a porous medium property that could be characterized by a *scalar* value. Later on its *tensorial* nature was recognized, and

for 2-dimensional isotropic problems it is characterized by a longitudinal and transverse component (Scheidegger, 1961). Discrepancies between laboratory and field measurements of dispersivity, and the observation that the dispersion coefficient increases with the mean travel distance of the solute led to the notion of *scale*-dependency. Nowadays the scale effect has been cast in terms of the travel *time*- or *distance*-dependency, and has been related to the variance and the correlation length of the medium (a.o. Dagan (1982), Gelhar and Axness (1983)). The present study suggests that in a *modelling situation* one should bear in mind that the dispersive mechanism of transport is a *data*-dependent phenomenon. Thus a trade-off between dispersion and advection exists, and the governing factor is the amount, type and spatial configuration of data available to constrain the simulations. Dispersivities derived from theoretical principles assume *uniform flow* fields, that is, when no flow-constraining information is available whatsoever. They represent maximum values, i.e. the amount of dispersion needed when no other data than just the stochastic parameters are known. In a practical situation this is of course quite unrealistic. The stochastic parameters themselves are derived from the field conductivity data, and the latter should, together with *all* the other data material, be incorporated in the simulation model in a deterministic sense. Although the dispersivities based on the stochastic properties of the logconductivity represent maximum values, it is still possible that the needed model dispersion is larger. This is due to the fact that it lumps alle the model deficiencies. And these uncertainties tend to be larger, the larger the scales of the flow and transport domain are. On the one hand, the data-dependency of dispersion makes the mechanism even more difficult to handle, as it seems not possible to establish straightforward relationships of general validity. On the other hand, however, there is the prospect of reducing its magnitude when the *advective* transport is properly calibrated against different data sets of sufficient quantity and quality. And this may be a difficult task indeed.

The flow and transport uncertainties predicted by stochastic simulations in general, and in this study in particular should not be interpreted as absolute quantities: Firstly, the uncertainties are related to one model *parameter* only, the transmissivity, assuming the other to be known exactly.
Secondly, the determination of the stochastic properties of logtransmissivity is based on a very sparse data set that was assumed to be without error and to reflect information at the scale of the model cells.

Deterministic model studies that are properly calibrated against extensive data material, including observations of head, transmissivity and concentration, may form an adequate basis for many practical purposes. Stochastic models offer the advantage of being able to quantify the migration uncertainties, but they should include conditioning on the mentioned data, otherwise the predicted uncertainties may be grossly overestimated. And consequently the expected migration pattern may differ considerably from reality.

The Monte Carlo method is still the most versatile stochastic transport simulation method available at present. It allows for quantifying the uncertainties of head and concentration through space and time, while very few restricting assumptions have to be adopoted. The main disadvantages are the large computation times and the fact that the results are not optimised. The conditioning criteria of head and concentration used in this work are based on absolute deviations and ensure a *plausible* solution, i.e. a solution in agreement with the data material at a predefined confidence level, but not an optimal solution.

The results obtained in this work are related to the specific case study. Although it seems reasonable to generalize the findings to some extent, more studies are required to further examine the effect of conditioning transport simulations on different types of data. This is undoubtedly a subject of great practical importance, and therefore it deserves increased attention.

References

Clifton, P.M., and Neuman, S.P. (1982) Effects of kriging and inverse modelling on conditional simulation of the Avra Valley aquifer in Southern Arizona, *Water Resour. Res., Vol. 18(4),* pp. 1215-1234.

Dagan, G. (1982) Stochastic modelling of groundwater flow by unconditional and conditional probabilities 2, The solute transport, *Water Resour. Res., Vol. 18(4),* pp. 835-848.

Dagan, G. (1984) Solute transport in heterogeneous porous formations, *J. Fluid Mech., Vol. 145,* pp. 151-177.

Dagan, G. (1986) Statistical theory of groundwater flow and transport: pore to laboratory, laboratory to formation, and formation to regional scale, *Water Resour. Res., Vol. 22(9),* pp. 120S-134S.

Delhomme, J.P. (1978) Kriging in hydrosciences, *Adv. in Water Res., Vol. 1(5),* pp. 251-266.

Delhomme, J.P. (1979) Spatial variability and uncertainty in groundwater flow parameters: a geostatistical approach, *Water Resour. Res., Vol. 15(2),* pp. 269-280.

Freeze, R.A. (1975) A stochastic-conceptual analysis of one-dimensional groundwater flow in nonuniform homogeneous media, *Water Resour. Res., Vol. 11(5),* pp. 725-741.

Freeze, R.A., and Cherry, J.A. (1979) *Groundwater,* Prentice-Hall, Englewood Cliffs, New Jersey, 604 p.

Gelhar, L.W., and Axness, C.L. (1983) Three-dimensional stochastic analysis of macrodispersion i aquifers, *Water Resour. Res., Vol. 19(1),* pp. 161-180.

Hefez, E., Shamir, U., and Bear, J. (1975) Identifying the parameters of an aquifer cell model, *Water Resour. Res., Vol. 11(6),* pp. 993-1004.

Hoeksema, R.J., and Kitanidis, P.K. (1985) Analysis of spatial properties of selected aquifers, *Water Resour. Res., Vol. 21(4),* pp. 563-572.

Journel, A.G., and Huijbregts, Ch. J. (1978) *Mining geostatistics,* Academic Press, London, 600 p.

Konikow, L.F., and Bredehoeft, J.D., (1978) Computer model of two-dimensional solute transport and dispersion in ground water, USGS, book 7, chapter C2, 90 p.

Lallemand-Barrès, A., and Peaudecerf, P. (1978) Recherche des relations entre la valeur de la dispersivité macroscopique d'un milieu aquifère, ses autres caractéristiques et les conditions de mesure. Etude bibliographique, Bulletin du BRGM, 2nd. série, section III, No. 4, pp. 227-284.

Mantaglou, A., and Wilson, J.L. (1984) The turning bands method for simulation of random fields using line generation by a spectral method, *Water Resour. Res., Vol 18(5),* pp. 1379-1394.

Marsily, G. de (1986) *Quantitative Hydrogeology – Groundwater hydrology for engineers,* Academic Press, Orlando, 440 pp.

Matheron, G. (1973) The intrinsic random functions and their applications, *Adv. Appl. Prob., Vol. 5,* pp. 439-468.

Reclamation of landfill leachate polluted groundwater (1986) Report prepared for the Commission of the European Communities, Technical University of Denmark, Lyngby, 103 p.

Scheidegger, A.E. (1961) General theory of dispersion in porous media, *J. Geoph. Res., Vol. 66,* pp. 10, 3273-3278.

Smith, L., Schwartz, F.W. (1980) Mass transport 1, A stochastic analysis of macroscopic dispersion, *Water Resour. Res., Vol. 16(2),* pp. 303-313.

Smith, L., and Schwartz, F.W. (1981a) Mass transport 2, Analysis of uncertainty in prediction, *Water Resour. Res., Vol. 17(2),* pp. 351-369.

Smith, L., and Schwartz, F.W. (1981b) Mass transport 3, Role of hydraulic conductivity in prediction, *Water Resour. Res., Vol. 17(5),* pp. 1463-1479.

Terraqua Aps (1985) Undersøgelse af forurening fra Løgtved losseplads. Statusrapport + bilag (in Danish).

Van Rooy, D. (1986a) Stochastic modelling of a contaminated aquifer – the unconditional approach, *Nordic Hydrology, Vol. 17(4/5),* pp. 315-324.

Van Rooy, D. (1986b) Conditional stochastic simulation of groundwater contamination – a case study, Series Paper No. 42, Technical University of Denmark, 154 pp.

Van Rooy, D. (1987) Automatic interpretation of slug test data from unconfined aquifers with completely or partially penetration wells – (submitted).

Virdee, T.S., and Kottegoda, N.T., (1984) A brief review of kriging and its application to optimal interpolation and observation well selection, *J. Hydr. Sc., Vol. 29, (4),* pp. 367-387.

Received: 1 July, 1987

Address:
Institute of hydrodynamics
and hydraulic engineering ISVA,
Technical University of Denmark,
Bldg. 115, DK – 2800 Lyngby,
Denmark.

Appendix

Main concepts of geostatistics and turning bands technique. For detailed information see Journal and Huijbregts (1978) or Marsily (1986).

Semivariogram – The spatial distribution of an earth parameter, property or variable can be characterized by a spatial structure and a randomly fluctuating component. In groundwater hydrology such a variable may be hydraulic head or conductivity, concentration, porosity etc. When stationary conditions can be assumed, the semivariogram γ is directly related to the autocovariance C, thus

$$\gamma(h) = \frac{1}{2}\,\mathrm{var}[Z(x) - Z(x+h)]$$

and $$\gamma(h) \equiv C(0) - C(h)$$

where $C(0)$: finite field variance, $C(h)$: autocovariance, h: lag vector.
The stationary condition is often relaxed to weak stationarity, meaning that only the first two statistical moments are required to be invariant. Then the expectation is constant, and the covariance depends only on the lag

$$E[Z(x)] = m$$

and $$E[Z(x)\,Z(x+h)] - m^2 \equiv C(h)$$

In case of statistical isotropy, the lag vector reduces to a lag distance. For practical purposes a sample or experimental semivariogram is estimated from the data by

$$\gamma^e(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z(x_i) - z(x_i+h)]^2$$

where $N(h)$: number of data pairs separated by lag h. Grouping of the data pairs in discrete lag and angle intervals is usually done. A theoretical model is fitted to the experimental semivariogram.

Simple kriging – kriging is an interpolation technique based on the spatial structure of a random variable. The interpolation is optimal in the sense that it provides the best linear unbiased estimate. The kriged value z^* at an location x_0 is

$$z^* = \sum_{i=1}^{n} \lambda_i z_i$$

The number of data on which the interpolation is based, n, is called the nearest neighbourhood. The essence of kriging is to determine the weight coefficients λ_i, while ensuring a minimum estimation variance and unbiasedness of the estimation, thus

$$\sigma^2 \equiv E[(Z^* - Z)^2] \equiv \mathrm{minimum}$$

and $$E[Z^* - Z] \equiv 0$$

The so-called kriging system is in matrix form

$$[C]\{\lambda\} = \{b\}$$

and contains terms derived from the semivariogram.

Kriging is an exact interpolator, i.e. at the data locations it will reproduce the data values. Kriging can be used to evaluate the semivariogram by the so-called cross-validation procedure. This consists of deleting one data point at a time and estimating it on basis of the remaining data.

Universal kriging – Let the non-stationary phenomenon $Z(x)$ be composed of a random component $R(x)$ and a deterministic trend or drift $m(x)$, then

$$E[Z(x)] \equiv m(x)$$

and $$Z(x) - m(x) = R(x)$$

where the drift is location-dependent. Assume the residual function $R(x)$ to be stationary, then its expected value is location-invariant and equal to zero

$$E[R(x)] = 0$$

Thus, if the drift can be estimated, then the kriging weights can be based on the autocovariance of the residuals, and kriging would proceed just as previously described. However, in practice it is not possible rigorously to estimate simultaneously the drift and the covariance structure. This problem is circumvented by working with differences of differences of the field variable, which filters out the drift without actually determining it. The practical identification of the so-called generalised covariance is done by polynomial fitting and cross-validation is used to compare different models and neighbourhoods, see references.

Turning Bands Method – As kriging is based on the spatial structure underlying a random field, it effectively smooths out the random fluctuations. The field variability is reflected in the kriging variances. Using a random number generator it is possible to simulate the field variability by constructing synthetic versions of it. The turning bands method is one way to do so. The TBM replaces the original multi-dimensional problem by a large number of lines L radiating from a common origin. A second-order stationary process is assumed and

$$Z_N \equiv \frac{1}{\sqrt{L}} \sum_{1}^{L} Z_i(X_N u_i)$$

where Z_N is the field value, X_N and u_i are location vectors in space and along a line, respectively. Z_i is a 1-D value generated using a spectral method, see Mantoglou and Wilson (1982).