SKB

TECHNICAL REPORT 91-27

Verification of HYDRASTAR – A code for stochastic continuum simulation of groundwater flow

Sven Norman
Starprog AB

July 1991

SVENSK KÄRNBRÄNSLEHANTERING AB
SWEDISH NUCLEAR FUEL AND WASTE MANAGEMENT CO
BOX 5864 S-102 48 STOCKHOLM
TEL 08-665 28 00 TELEX 13108 SKB S
TELEFAX 08-661 57 19

VERIFICATION OF HYDRASTAR - A CODE FOR STOCHASTIC CONTINUUM SIMULATION OF GROUNDWATER FLOW

Sven Norman

Starprog AB

July 1991

VERIFICATION OF HYDRASTAR - A CODE FOR STOCHASTIC CONTINUUM SIMULATION OF GROUNDWATER FLOW.

by Sven Norman, Starprog AB
July, 1991

TABLE OF CONTENTS

0. EXECUTIVE SUMMARY

When modelling groundwater flow in fractured rock one must first make a basic choice concerning the conceptual description of the problem. The first alternative is to try to understand the characteristics of the complex geometric shapes of the void space of the rock such as the distributions of channels or fractures. The other alternative is to disregard this and go directly for the effect of the geometry on hydrology which is some kind of permeability or conductivity field. However these two approaches should not be perceived as two competing and entirely different approaches. On the contrary the difference should be viewed as a matter of scale and the former thus being a small scale model and the latter a large scale model. Ideally the connection between the models should be fully understood. Currently this is however not the case. Inspired by [Neuman, 1988] our choice is to describe the rock using a continuous conductivity field.

Since the hydraulic properties of the rock are extremely heterogeneous and sample locations are very sparsely positioned a great deal of uncertainty will always be present in subsurface hydrology. Thus it is natural to regard all parameters or fields as stochastic. In particular if one introduces a conductivity field, it should be a stochastic field. In this connection we also refer to a preliminary study [Lovius at al, 1990] which used a two dimensional stochastical continuum approach for analyzing groundwater flow around a repository. It showed that the spatial variability and the associated uncertainty have a great impact on repository safety predicitions. Another conclusion of this study was that a three dimensional approach is desirable in order to describe the full importance of spatial variability.

Thus, HYDRASTAR is a code developed at Starprog AB for use in the SKB 91 performance assessment project with the following principal function:

- Reads the actual conductivity measurements from a file created from the data base GEOTAB.
- Regularizes the measurements to a user chosen calculation scale.
- Generates three dimensional unconditional realizations of the conductivity field by using a supplied model of the conductivity field as a stochastic function.
- Conditions the simulated conductivity field on the actual regularized measurements.
- Reads the boundary conditions from a regional deterministic NAMMU computation.
- Calculates the hydraulic head field, Darcy velocity field, stream lines and water travel times by solving the stationary hydrology equation and the streamline equation obtained with the velocities calculated from Darcy's law.
- Generates visualizations of the realizations, if desired.
- Calculates statistics such as semivariograms and expectation values of the output fields by repeating the above procedure by iterations of the Monte Carlo type.

When using computer codes for safety assessment purposes validation and verification of the codes are important. Thus this report describes a work performed with the goal of verifying parts of HYDRASTAR. The verification described in this report uses comparisons with two other solutions of related examples:

A. Comparison with a so called perturbation solution of the stochastical stationary hydrology equation. This is an analytical approximation of the

stochastical stationary hydrology equation valid in the case of small variability of the unconditional random conductivity field. The compared output is the semivariogram function of the hydraulic head. Added is also the comparison of the input conductivity semivariograms with the estimated counterparts.

B. Comparison with the [Hydrocoin, 1988], case 2. This is a classical example of a hydrology problem with a deterministic conductivity field. The principal feature of the problem is the presence of narrow fracture zones with high conductivity. The compared output are the hydraulic head field and a number of stream lines originating from a set of given positions.

Example B tests the solver of the hydrology equation and the stream line equation whereas example A tests the hydrology equation solver, the estimation of output statistics, and the generator of the unconditional realizations. Some further verification of the conditioning of the simulations is contained in [Norman, 1991]

In this executive summary we show the dominating results. For the case A these are contained in the Figs. 0.1 to 0.3 which show the perturbation solution head semivariograms in the directions parallel and orthogonal to the applied head gradient and the input conductivity semivariograms all compared to the confidence levels of the estimated counterparts.

As the reader easily can see the results for the head semivariograms embodied in Figs. 0.1 and 0.2 are not quite satisfactory. Instead we note the following flaws:

1. Too high semivariogram values for the short lags in the direction of the applied head gradient.
2. Too high semivariogram values for lags orthogonal to the direction of the applied head gradient.
3. Too low semivariogram values for large lags in the direction of the applied head gradient.
4. The convergence is bad for large lags in the direction of the applied gradient.

Discussing these points one at a time:

Point 1 is probably due to a too coarse mesh. The mesh used (31x71x71=138069 head nodes) contains three nodes per effective correlation length for the conductivity. Experiments have been performed with six nodes per effective correlation length for the conductivity in the direction of the head gradient and this produces a marked improvement for the short lags in the direction of the applied gradient.

Point 2 is not affected by the refinement mentioned above. The obvious test is to refine the mesh in the directions orthogonal to the applied gradient directions as well. This has not been done primarily because the total number of nodes increases as the square of the number of nodes in the directions orthogonal to the head gradient.

Point 3 is probably due to the finite approximation of the whole space. Tests that have been performed increasing the computational domain seem to validate this.

There is not much to do about point 4.

Thus we see that in the search of better results we are led to put larger and larger demands on our computer resources by increasing the mesh density and the size of the computational domain. The above example is, however, already on the computational limit.

The dominating results for case B is contained in the comparison between the diagrams shown in Fig. 0.4 and comparison between the diagrams in Fig. 0.5 on one hand and the diagrams in Fig. 0.6 on the other.

Covariance model (see section 2.2.2.):	Exponential
Range parameter, a (see section 2.2.2.):	1
Bandwidth factor, f (see section 6.2.):	0. 2
Bandwidth, T (see section 2.2.)::	0. 2
R (see section 2.2.):	20
Number of icosahedron sets (see section 2.3.) :	4
Number of random lines (see section 2.3.) :	0
Spatial sample size(for head), I P I (see section 4.):	121
Spatial sample size(for head), I L I (see section 4.):	1
Number of realizations:	400

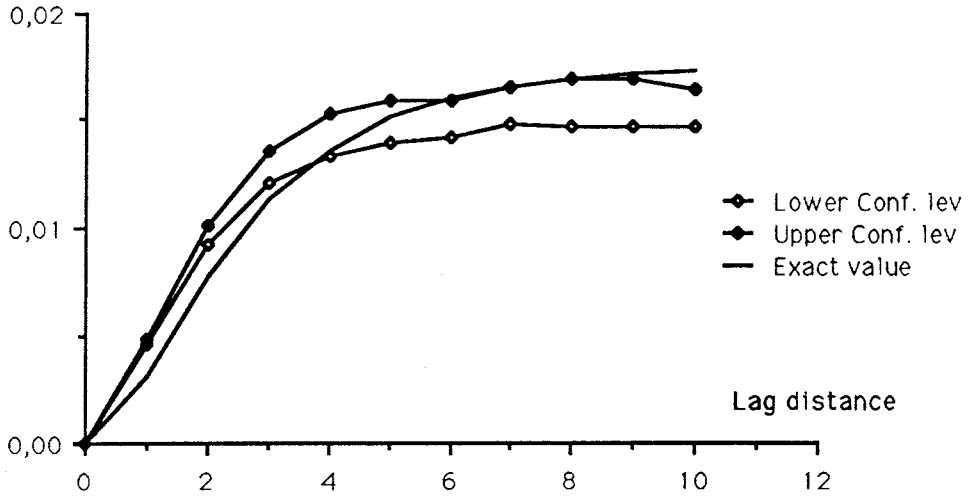


Fig. 0.1. Exact semivariogram and the Chebyshev confidence intervals on the 0. 95 -level for the heads in the direction parallel to the applied head gradient.

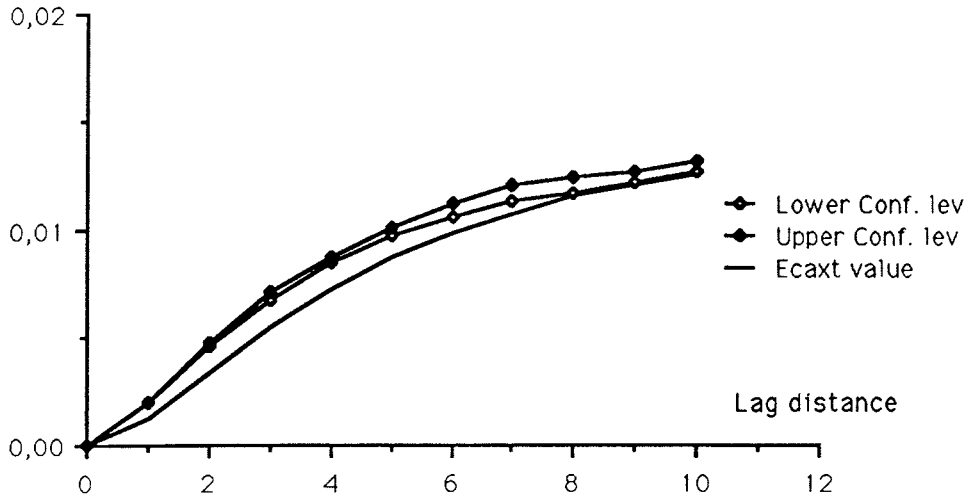


Fig. 0.2. Exact semivariogram and the Chebyshev confidence intervals on the 0. 95 -level for the heads in the direction orthogonal to the applied head gradient.

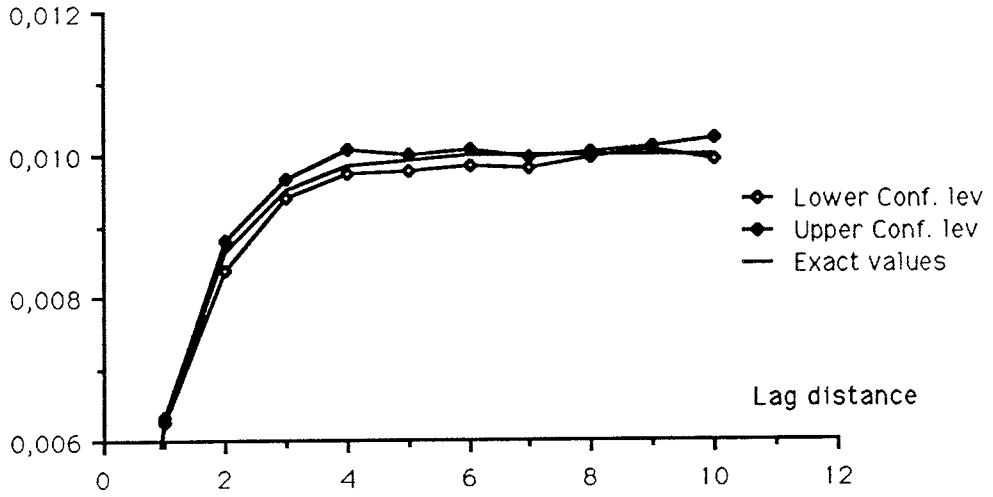


Fig. 0.3. Exact semivariogram and the Chebyshev confidence intervals on the 0. 95 -level for the conductivities in the direction of the gradient. The diagram shows only the upper part since all curves coincides on the lower.

The diagram in Fig. 0.4 shows the head values at a certain depth obtained from HYDRASTAR calculations and from [Hydrocoin, 1988]. The results are in good agreement, except for a small offset, round +3 m, in the head values. The offset is due to our approximation of the top boundary values. This could be overcome by either choosing a more sophisticated way of representing the top boundary or simply using other teams calculated head values as our boundary values. See chapter 7.

Finally Fig. 0.5 shows the stream lines and travel distance versus travel time for one of the trajectories obtained from a HYDRASTAR calculation. These favorably compare with the corresponding diagrams in Fig. 0.6 which are taken directly from [Hydrocoin, 1988].

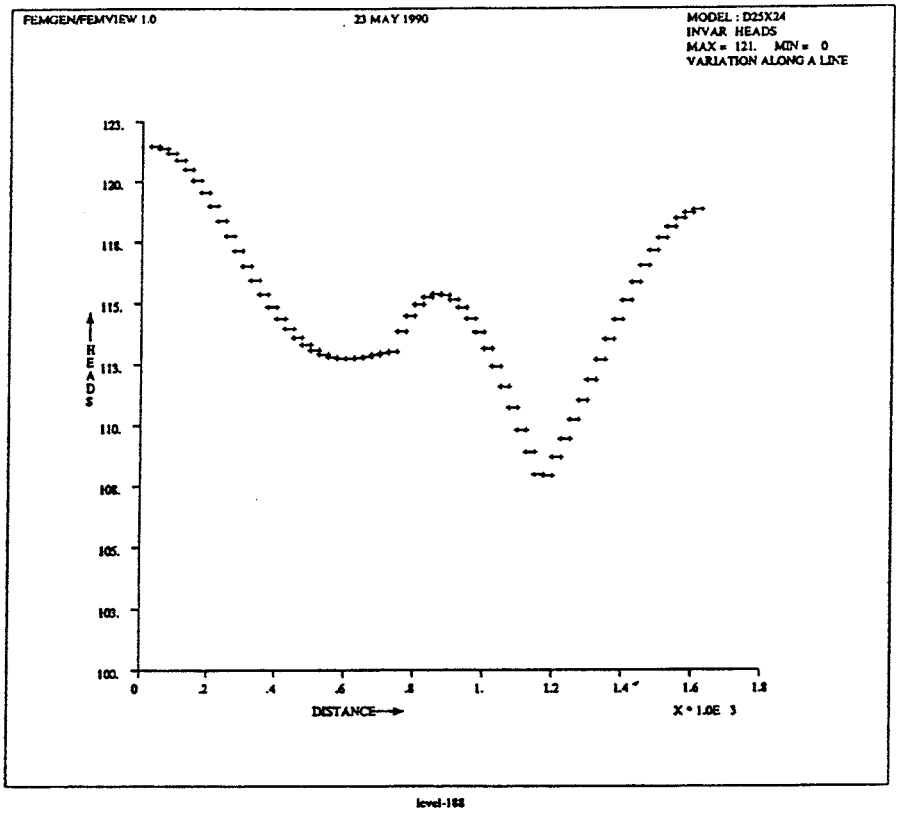
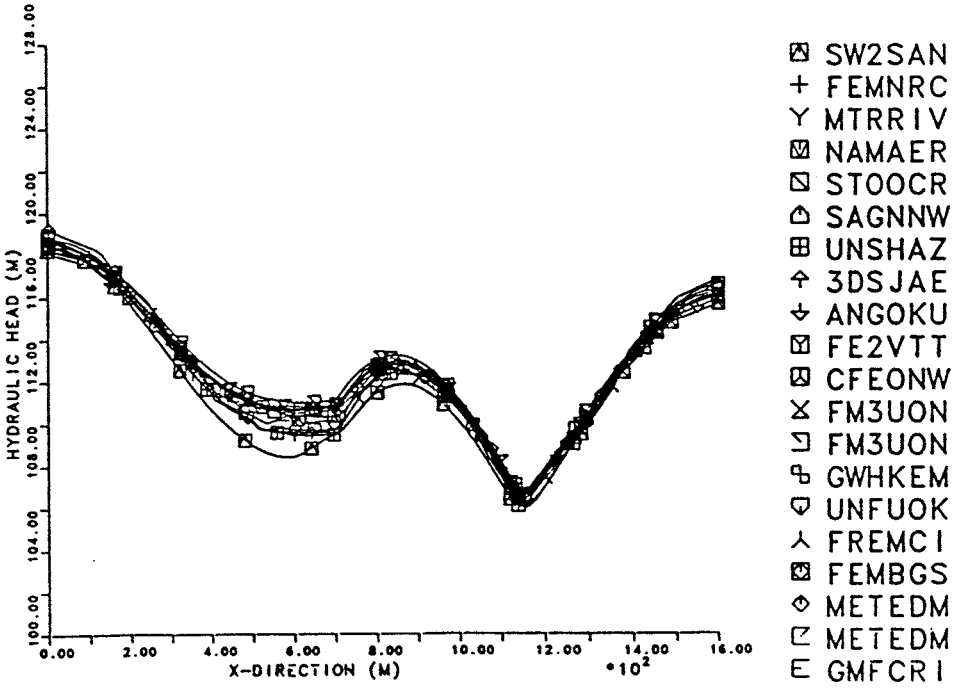


Fig. 0.4. Head values at the level -188 m from HYDRASTAR using mesh 2 (above) and from [Hydrocoin, 1988] at the level -200 m using medium meshes (below).



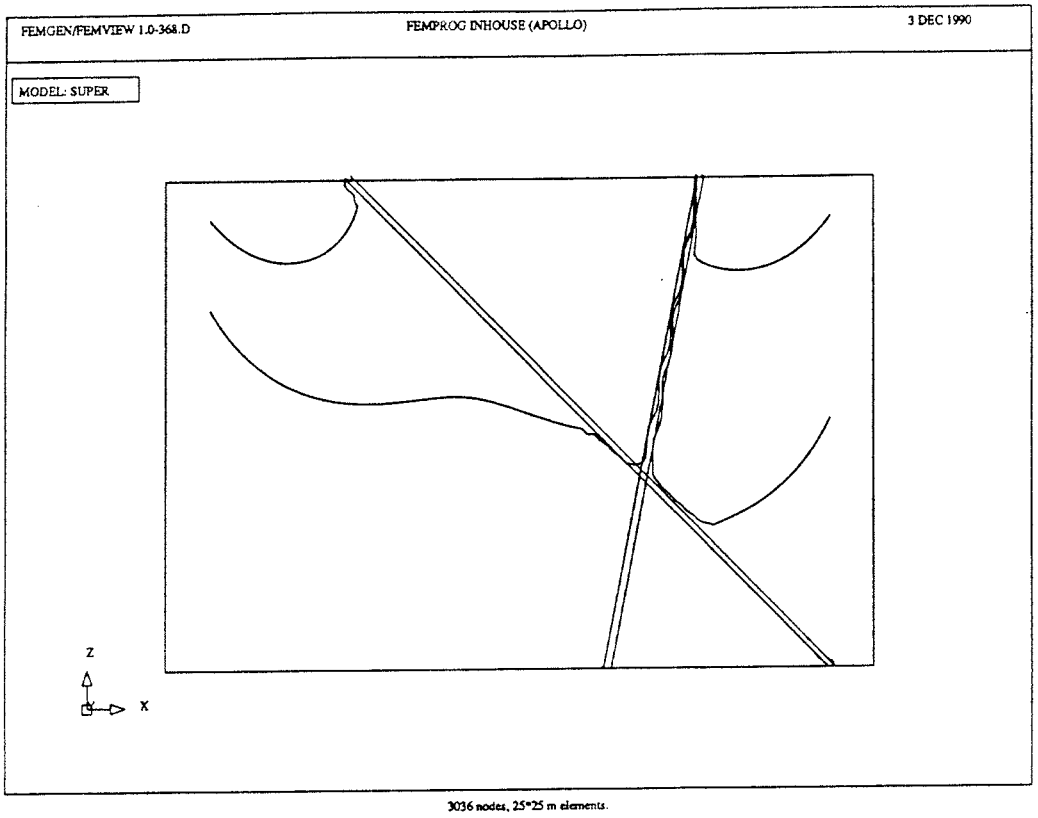
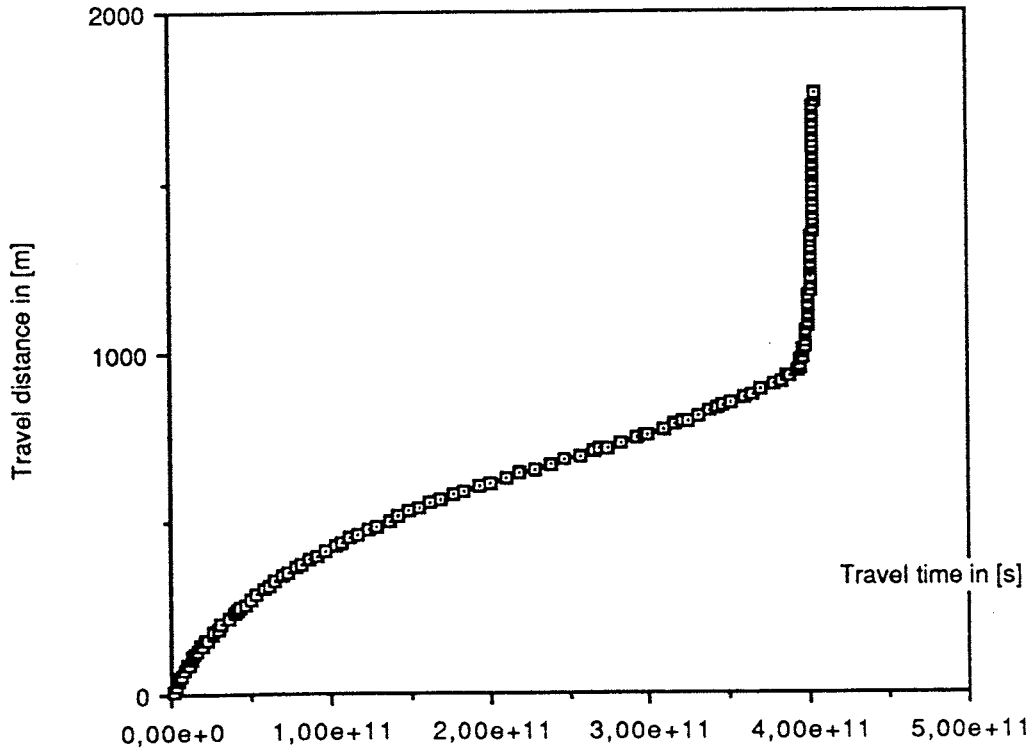


Fig. 0.5. Stream lines from HYDRASTAR using mesh2 (above). Travel distance versus time from HYDRASTAR and for trajectory number 2 (lowest leftmost tracer) using mesh2 (below).



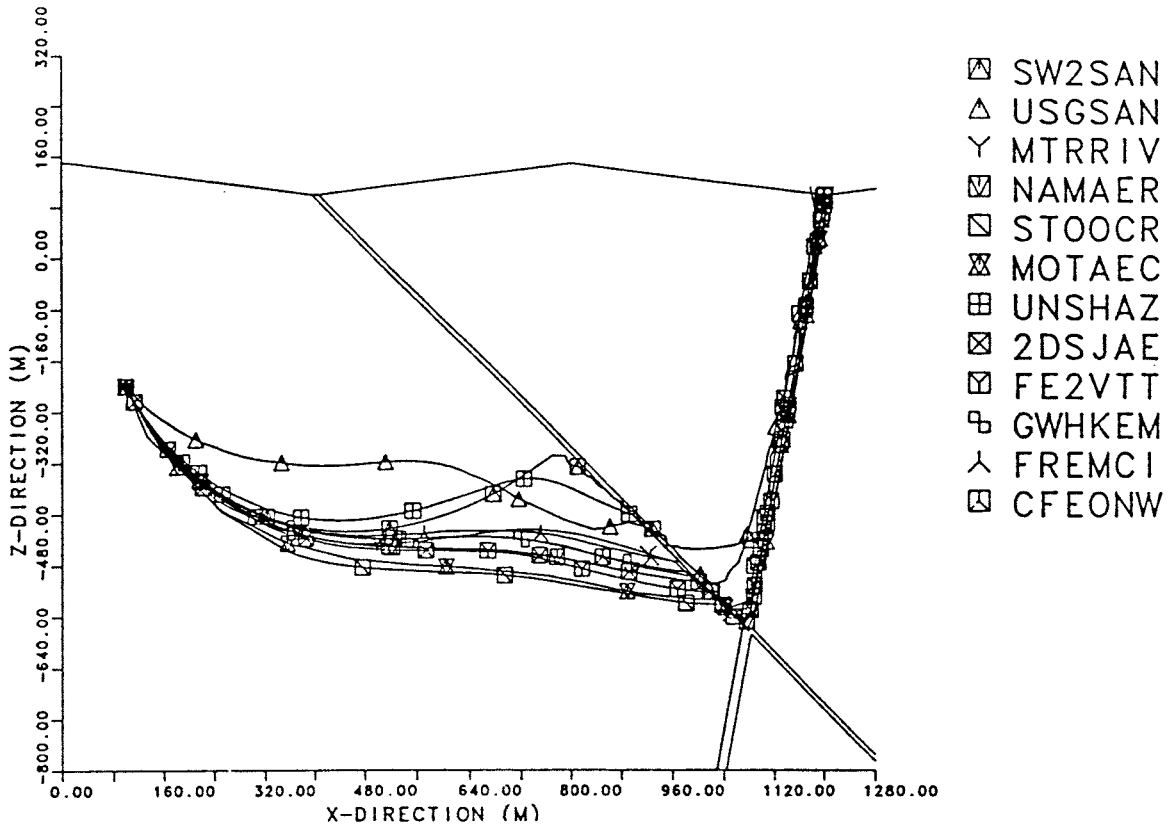
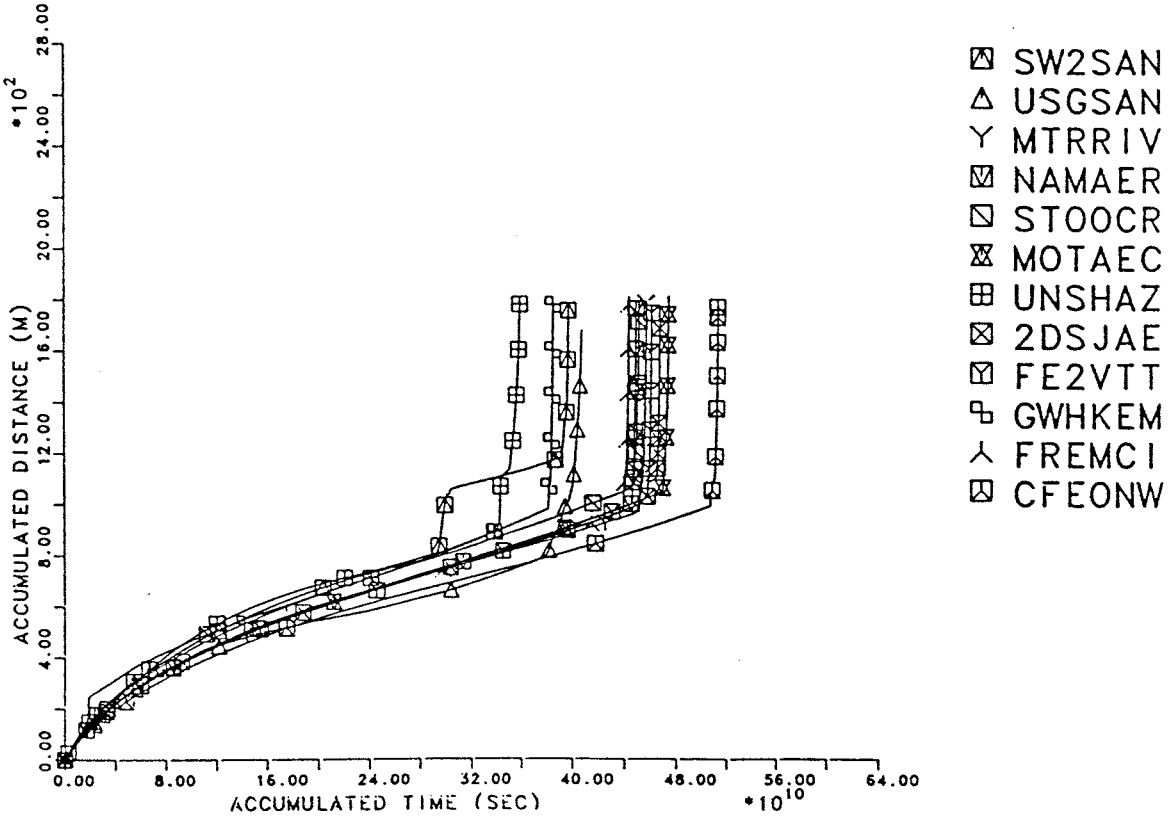


Fig. 0.6. Stream lines (above) and travel distance versus time (below) for trajectory number 2 (lowest leftmost tracer) from [Hydrocoin, 1988] using the fine meshes.



1. INTRODUCTION

Generally speaking a person with a desire to model the hydraulic performance of a subsurface repository will face a very complex problem. To start with a certain amount of knowledge about the rock is given. This knowledge has a lot of different forms, for instance

- location of inferred fracture zones,
- fracture statistics from surface outcrops, tunnels and core mappings,
- single hole packer tests, transient and stationary,
- interference tests, cross-hole tests,
- tracer tests,
- general geological information.

From this information it is possible to form different models of the rock as a medium for water and radionuclide transport. Such a model will always be uncertain for several reasons

a. the translation of knowledge into a model and model parameters such as conductivity values or distributions for fracture densities does always involve some assumptions difficult to assess,

b. the hydrological properties of the rock is very heterogeneous whereas the measurement locations are sparsely located.

HYDRASTAR is a code developed at Starprog AB at the request of SKB AB designed to deal with a restricted form of the formidable problem above. The restriction is mainly that the knowledge used is almost only the stationary single-hole packer test measurements. Referring to the point a. above the basic assumptions employed are

- The rock is assumed to behave like a stochastic continuum in the following sense: for some range of scales s, $L_1 \leq s \leq L_2$ the rock obeys the isotropic form of Darcy's law and the introduced conductivity field is regarded as a stochastic function.
- The stationary single-hole packer tests can be used to calculate the value of the conductivity at a position given as the midpoint of the packer interval.

In this connection it should be stated that this approach has been inspired by [Neuman, 1988].

With these assumptions the problem to take the uncertainty in unsampled regions into account, point b. above can be posed in a stringent fashion for the conductivities. This requires that one first infers a model for the stochastical function equal the conductivity field and then uses this to simulate the conductivity field in the unsampled regions conditioned on the performed measurements. This conditional simulation is then repeated and the related fields such as heads, velocities and so forth are calculated. From these repeated simulations it is possible to collect statistics and thus obtain a grip on the resulting spatial variability and the uncertainty in the hydraulic situation around a repository.

The comprehensive detailed description of HYDRASTAR is deferred to a later report. The scope of this report is to compare results generated using HYDRASTAR with:

A. Input semivariogram functions and analytical perturbation solution to the hydrology equation.

B. Finite element solutions to the hydrology equation with a deterministic conductivity field. The computational example and the finite element solutions are taken from the international verification project for groundwater computer codes [Hydrocoin, 1988].

This leads to the division of the report in four main parts:

- In order to discuss the sometimes rather technical aspects of the comparisons, in particular with regard to the random function generator we start with a thorough description of those parts of HYDRASTAR being tested by A and B above. These are the generator of the unconditional conductivity simulations in chapter 2, the hydrology equation solver in chapter 3 and the estimation of output statistics in chapter 4.
- Derivation of the analytical perturbation solution in chapter 5.
- Description of the results generated in part A in chapter 6.
- Description of the results generated in part B in section 7.

2. STOCHASTICAL FUNCTION GENERATOR

In the following we shall describe how to generate unconditional Gaussian regionalization with a given covariance or semivariogram function. The method used by HYDRASTAR is the turning bands algorithm as described by [Journel and Huijbregts, 1978] and [Tompson, Ababou and Gelhar, 1989]. The original reference to this method is [Matheron, 1973].

2.1. The turning bands method

Regard the problem of simulating a three dimensional random function Y. We assume it to be weak second order stationary, that is the following two requirements should hold. First, the expected value of Y should be constant i.e.

$$E[Y(x)] = m = \text{constant} ,$$

where we introduced x for an arbitrary point in the three dimensional real space, $\mathbf{R}^3$, and E[·] for the expectation value operator. Secondly, introducing the residual process Y' by

$$Y'(x) = Y(x) - m ,$$

the covariance function

$$C_Y(\xi, x) = E[Y'(x + \xi)Y'(x)] .$$

should depend only on the lag vector ξ connecting two points in $\mathbf{R}^3$. Let us assume further that the random function Y is Gaussian and that the moments above, i.e. the expectation value and the covariance function, are known. This specifies the distribution of the regionalization completely.

The basic idea with the turning bands algorithm is to write

$$Y(x) = \int_{S_{1/2}} Y_1((x, l), l)\sqrt{f(l)}\, dS(l) \qquad 2.1.1$$

where f(l) is a probability density function on the half unit sphere $S_{1/2}$, l is a unit vector on $S_{1/2}$, (· , ·) signifies the ordinary scalar product and $Y_1(\cdot, l)$ is a family of stochastic processes indexed by l with the following properties

- $Y_1(\cdot, l_1)$ and $Y_1(\cdot, l_2)$ are uncorrelated if $l_1 \neq l_2$ and
- $Y_1(\cdot, l_1)$ is stationary with zero expectation and a covariance function $C_1(\cdot, l_1)$.

Above we also introduced a general notation exemplified with $Y_1(\cdot, l_1)$. This simply means that this object is the function or process of one variable obtained by fixing the variable given as the second argument to the value l_1.

As an approximation for the case of a uniform distribution on the half unit sphere, i.e. $f = 1/2\pi$ identically, we write

$$Y(x) = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} Y_1((x, l_i), l_i) \qquad 2.1.2$$

where $\{l_i\}_{i=1}^{N}$ is a set of unit vectors on $S_{1/2}$. The first thing to note is that the marginal distributions of this process will tend to a normal distributions as the number of lines tends to infinity as follows from the central limit theorem [Fisz, 1963, p 196.]. The second thing to notice is that trivially the stationarity of Y follows from the stationarity of $Y_1(\cdot, l)$. In particular

$$E[Y_1(\cdot, l)] = 0, \forall\ l \Rightarrow E[Y] = 0.$$

However, the main thing about the method is to find the covariances $C_1(\xi, l)$ which produces the given covariance $C(\xi)$. Hence using the properties of the family $Y_1(\cdot, l)$ we express the covariance of Y in the covariances of $Y_1(\cdot, l)$ as

$$C_Y(\xi) = \int_{S_{1/2}} C_1((\xi, l), l) f(l) dS(l) = E_l[C_1((\xi, l), l)] \qquad 2.1.3$$

and in the approximate case

$$C_Y(\xi) = \frac{1}{N}\sum_{i=1}^{N} C_i((\xi, l_i), l_i).$$

To produce a stringent derivation of 2.1.3 from the exact formula 2.1.1 necessitates the introduction of a whole machinery and is therefore excluded. [1]

Using the spectral theorem [Yaglom, 1962] we write utilizing the first equality of equation 2.1.3

$$\frac{1}{4\pi^2}\int_{\Re^3} S_Y(\lambda) e^{i(\lambda,\xi)} d\lambda = \int_{S_{1/2}} \int_{-\infty}^{\infty} S_1(\lambda_1, l) e^{i(\xi, \lambda_1 l)} f(l) d\lambda_1 dS(l) \qquad 2.1.4$$

where $S_Y(\cdot,)$ and $S_1(\cdot, l)$ are the spectrums of the processes $Y(\cdot, l)$ and $Y_1(\cdot, l)$ respectively.

The idea is now to rewrite the right hand side into the same form as the left hand side. The first step in doing this is to divide the right hand side into two parts

$$\int_{S_{1/2}} \int_{0}^{\infty} S_1(\lambda_1, l) e^{i(\xi, \lambda_1 l)} f(l) d\lambda_1 dS(l) + \int_{S_{1/2}} \int_{-\infty}^{0} S_1(\lambda_1, l) e^{i(\xi, \lambda_1 l)} f(l) d\lambda_1 dS(l)$$

and then perform a change of variables as

---

[1] However what is needed is

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \psi_1 \psi_2 dF(x, l_1, \psi_1) dF(y, l_2, \psi_2) = C_1(x - y)\delta(l_1 - l_2)$$

in the sense of distributions on $S_{1/2} \otimes S_{1/2}$ where $F(x, l, \cdot)$ is the probability distribution function for the stochastic variable $Y_1(x, l)$.

$$\begin{cases} \lambda_1 l = \lambda \\ \lambda_1 = |\lambda| \\ l = \lambda/|\lambda| \\ |\lambda|^2 d\lambda_1 dS(l) = d\lambda \end{cases}$$

for the first integral and

$$\begin{cases} \lambda_1 l = \lambda \\ \lambda_1 = -|\lambda| \\ l = -\lambda/|\lambda| \\ |\lambda|^2 d\lambda_1 dS(l) = d\lambda \end{cases}$$

for the second. Introducing the notation $\Re^3_{1/2}$ for the part of $\Re^3$ intersected by rays through $S_{1/2}$ and $\Re^3_{-1/2}$ for $-\Re^3_{1/2} = \{-x : x \in \Re^3_{1/2}\}$ we then rewrite the previous expression as

$$\int_{\Re^3_{1/2}} S_1(|\lambda|, \lambda/|\lambda|)\, e^{i(\xi,\lambda)} f(\lambda/|\lambda|) \frac{d\lambda}{|\lambda|^2} + \int_{\Re^3_{-1/2}} S_1(-|\lambda|, -\lambda/|\lambda|)\, e^{i(\xi,\lambda)} f(-\lambda/|\lambda|) \frac{d\lambda}{|\lambda|^2}.$$

Finally, introducing the very natural definition $f(-l) = f(l)$ and $S_1(\lambda, -l) = S_1(-\lambda, l)$ for $l \in S_{1/2}$ we have that the covariance function of the turning bands representation of Y can be written as

$$\int_{\Re} S_1(|\lambda|, \lambda/|\lambda|) e^{i(\xi,\lambda)} f(\lambda/|\lambda|) \frac{d\lambda}{|\lambda|^2}$$

and thus by the uniqueness of Fourier transforms and 2.1.4 we have

$$S_Y(\lambda) = 4\pi^2 S_1(|\lambda|, \lambda/|\lambda|) \frac{f(\lambda/|\lambda|)}{|\lambda|^2}. \qquad 2.1.5$$

Now in the particular case that the probability density function, f, is uniform, i.e.. equal to $1/2\pi$, and the family of line processes $Y_1(\cdot, l)$ satisfies the additional condition that $C_1(\xi_1, l)$ is independent of l we have

$$S_Y(\lambda) = S_1(|\lambda|) \frac{2\pi}{|\lambda|^2}.$$

This result together with 2.1.5 is also obtained by [Tompson, Ababou and Gelhar, 1989] although they use another definition of the spectral density with regard to the factor 2π. These expressions are then used to determine the spectral density along each line from the given spectral density S_Y.

In the case of isotropic processes and uniform probability density function f one may directly relate the covariances of Y and $Y_1(\cdot, l)$ by the following computation given in [Journel and Huijbregts, 1978]. Starting from 2.1.3 and suppressing the direct dependence of C_1 on l, i.e. the second argument, we have

$$C_Y(\xi) = \frac{1}{2\pi} \int_{S_{1/2}} C_1((\xi, l))dS(l) = \frac{1}{4\pi} \int_S C_1((\xi, l))dS(l)$$

where we used the symmetry of C_1 and the notation S for the unit sphere.

Switching to polar coordinates with the direction of the polar axis parallel to ξ and θ as the angle between the radius vector and the polar axis we rewrite this as

$$\frac{1}{4\pi} \int_0^{2\pi} d\varphi \int_0^{\pi} C_1(|\xi|\cos\theta)\sin\theta d\theta = \frac{1}{2} \int_0^{\pi} C_1(|\xi|\cos\theta)\sin\theta d\theta .$$

Employing the change of variables

$$\begin{cases} |\xi|\cos\theta = s \\ -|\xi|\sin\theta = ds \end{cases}$$

and the symmetry of C_1 once again we have

$$\int_0^{|\xi|} C_1(s)\frac{ds}{|\xi|}$$

or

$$C_1(|\xi|) = \frac{d}{d|\xi|}(|\xi| C_Y(|\xi|)). \qquad 2.1.6$$

In practice one generates a finite number of lines, N, each line is divided in bands of width T. If the value of the line process $Y_1(\cdot, l)$ in the k:th band is taken to be constant and equal $y_{i,k}$ 2.1.2 is replaced by

$$Y(x) = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} y_{i,k} \qquad 2.1.7$$

where k is chosen so that

$$(k-1)T \le (x, l_i) < kT .$$

In order to make this a useful algorithm one has to find out how to generate the line processes $Y_1(., l)$ and the lines themselves so that the result is "sufficiently good".

We will treat this in the following two sections.

2.2. Line processes

This section will solely occupy itself with describing methods for simulating one dimensional stochastical processes with a given covariance function $C_1(\xi)$. This will be based on convolution techniques and follows essentially [Journel&Huijbregts]. We

stress from the beginning that no importance whatsoever will be given to the form of the distribution of the generated process but only to its two first moments.

We start with the observation that a covariance function $C_1(\xi)$ always allows a convolution representation as

$$C_1(\xi) = f * \check{f}(\xi) = \int_{-\infty}^{\infty} f(\xi - u) f(-u)\, du =$$

$$\int_{-\infty}^{\infty} f(\xi + u) f(u)\, du, \qquad 2.2.1$$

where f is a realvalued function and `f is a fairly nonstandard notation defined by

$$\check{f}(x) = f(-x).$$

This result follows easily from Bochners theorem [Yaglom, 1962] if we assume that C_1 corresponds to a spectral density function [2] since we may then write

$$C_1(\xi) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\xi\lambda} dF(\lambda) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\xi\lambda} S(\lambda) d\lambda$$

with S as the spectral density function and F as the spectral distribution function. The first integral is evaluated in the Stieltjes sense.

Referring again to Bochners theorem we know that $F(\omega)$ is a nondecreasing function. Thus $dF(\omega) \geq 0$ and $S(\lambda) \geq 0$.

Introducing the tilde sign ~ for Fourier transform and superindex * for complex conjugate we see that 2.2.1 gives

$$\tilde{C}_1(\lambda) = \tilde{f}(\lambda)\tilde{f}(\lambda)^* = |\tilde{f}(\lambda)|^2$$

and thus the choice of the function f only needs to satisfy

$$S(\lambda) = |\tilde{f}(\lambda)|^2$$

which has an infinite number of real valued solutions since $S(\lambda) \geq 0$. however essentially only one is continuous.

Taking the definition of a function against a stochastic measure for granted we write

$$Y(u) = \int_{-\infty}^{\infty} f(u + r) dT(r)$$

where dT(r) is a stochastic measure on the real line such that for two intervals I_1 and I_2 with lengths $|I_1|$ and $|I_2|$ we have

---

[2] That is we are assuming for simplicity that the covariance function has an absolutely continuous Fourier transform.

$$\begin{cases} E\,[T\,(I_1)] = E\left[\int_{I_1} dT\,(r)\right] = 0 \\ E\,[T\,(I_1)T\,(I_2)] = E\left[\int\int_{I_1 I_2} dT\,(r)dT\,(r')\right] = |I_1 \cap I_2|\sigma_T{}^2 . \end{cases} \qquad 2.2.2$$

It is now possible to show, at least formally [3] that the process so defined has the desired covariance function. In fact

$$C_1(s) = E[Y\,(u+s)Y\,(u)] = E\left[\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f\,(u+s+r)f\,(u+r')dT\,(r)dT\,(r')\right] =$$

$$\int_{-\infty}^{\infty} f\,(s+r)f\,(r)dr\sigma_T{}^2$$

which is what we wanted apart from the constant factor $\sigma_T{}^2$.

In practise we will make a discrete approximation of these formulas on a grid with spacing b by

$$Y\,(ib) = \int_{-\infty}^{\infty} f\,(r)dT\,(r-ib) = \sum_{k=-\infty}^{\infty} \int_{\left(k-\frac{1}{2}\right)b}^{\left(k+\frac{1}{2}\right)b} f\,(r)dT\,(r-ib) \approx$$

$$\sum_{k=-\infty}^{\infty} f\,(kb) \int_{\left(k-\frac{1}{2}\right)b}^{\left(k+\frac{1}{2}\right)b} dT\,(r-ib) = \sum_{k=-\infty}^{\infty} f\,(kb) \int_{\left(k-i-\frac{1}{2}\right)b}^{\left(k-i+\frac{1}{2}\right)b} dT\,(r) = \sum_{k=-\infty}^{\infty} f\,(kb)t_{k-i} . \qquad 2.2.3$$

Here t_k by the implicit definition in the equation above are independent stochastic variables with moments following directly from 2.2.2 i.e.

$$\begin{cases} E\,[t_k] = 0 \\ E\left[t_k{}^2\right] = b\sigma_T{}^2 = \sigma_t{}^2 . \end{cases}$$

This approximation is simple and in principle the one used by [Journel and Huijbregts, 1978]. The disadvantage is of course that in order for it to be a good approximation the grid spacing b has to be small. One could contemplate better integration schemes. This is however not done in this report.

For use in the turning bands method i.e. in formula 2.1.7 we want to have the values of the stochastic process generated in the points with spacing equal to the bandwidth T whereas the method described above generates values at points with spacing b, the magnitude of which is determined by the need of a sufficiently good approximation. Thus we require that

$$T = N_q b \qquad 2.2.4$$

where N_q is some integer.

---

[3] It does not pose any serious problems to prove this if we define all involved integrals as Riemann integrals. We only approximate with Riemann sums.

Another method for generating realizations of one dimensional processes is the FFT method which is advocated in [Tompson , Ababou and Gelhar, 1989].

2.2.1. Random number generator

The random number generator currently used in HYDRASTAR to generate random reals uniformly distributed on the interval [0, 1] is based on a combined linear congruential generator taken from [Brately, Fox and Schrage, 1987], claimed to originate from L'Ecuyer:

$$x_{i+1} = 40014 * x_i \bmod 2147483563$$

$$y_{i+1} = 40692 * y_i \bmod 2147483399$$

$$z_{i+1} = (x_{i+1} + y_{i+1}) \bmod 2147483563$$

yielding a random real

$$r_{i+1} = z_{i+1}/2147483563 \,. \qquad 2.2.1.1$$

This has proven to be fast and is claimed to have been extensively checked. In particular it has good spectral properties, a very large period and do not suffer from the well known weakness of linear congruential generators of having successive overlapping sequences of numbers falling on parallel hyperplanes. It should be portable across all machines having a word length of at least 32 bits.

In order to validate this generator further we have also used the following generator employed by the PROPER Monitor, see [McGrath and Irving, 1975],

$$x_{i+1} = 5^{15} * x_i \bmod 2^{47} \,. \qquad 2.2.1.2$$

that have been thoroughly tested [Pörn, 1986], [Coveyou and MacPherson, 1967]. Comparisons has been performed between semivariogram functions estimated from series of one dimensional realizations using either one of these random number generators and the results are almost identical. See section 6.1.

Due to the difficulties in handling the large integers involved in the generator 2.2.1.2 it is slower than the generator 2.2.1.1 and this is the reason for choosing the latter.

2.2.2. Application to spherical and exponential models

Two isotropic covariance models C(r) are often used [Journel and Huijbregts, 1978], the spherical model

$$C(r) = \begin{cases} V\left(1 - \frac{3}{2}\frac{r}{a} + \frac{1}{2}\frac{r^3}{a^3}\right) & 0 \le r \le a \\ 0 & r > a \end{cases}$$

and the exponential

$$C(r) = V \exp(-\lambda r) \qquad r \ge 0$$

where V signifies the variance, r the norm of the vector separating two measurement points, i.e. the lag vector and a is a range parameter.

According to 2.1.6 we derive the corresponding covariances required for the line processes as

$$C_1(s) = \begin{cases} V\left(1 - 3\frac{s}{a} + 2\frac{s^3}{a^3}\right) & 0 \le s \le a \\ 0 & s > a \end{cases}$$

and

$$C_1(s) = V(1 - \lambda s)\exp(-\lambda s) \quad s \ge 0$$

respectively.

For these two models we propose to use convolution separations of the type 2.2.1 with the convolution function f given by [Journel and Huijbregts, 1978, p 507 - 508]

$$f(u) = \begin{cases} \left(\frac{12V}{a^3}\right)^{1/2} u & |u| \le \frac{a}{2} \\ 0 & |u| > \frac{a}{2} \end{cases} \qquad 2.2.2.1$$

and

$$f(u) = \begin{cases} 2(V\lambda)^{1/2}(1 - \lambda u)e^{-\lambda u} & u > 0 \\ 0 & u < 0 \end{cases} \qquad 2.2.2.2$$

respectively. It is easily verified by direct, although somewhat tedious, computation that these functions satisfy the equation 2.2.1. However it is not clear whether these particular choices are the best.

Now we want to specialize the approximate formula 2.2.3 to these convolution divisions. In the spherical case we write

$$Y(ib) = \int_{-a/2}^{a/2} f(r)\,dT(r - ib) = \sum_{k=-R}^{R} \int_{\left(k-\frac{1}{2}\right)b}^{\left(k+\frac{1}{2}\right)b} f(r)\,dT(r - ib)$$

$$\approx \sum_{k=-R}^{R} f(kb)\, t_{k-i} \qquad 2.2.2.3$$

with

$$t_k = \int_{\left(k-\frac{1}{2}\right)b}^{\left(k+\frac{1}{2}\right)b} dT(r)$$

and thus we require that

$$a = (2R + 1)b. \qquad 2.2.2.4$$

For the choice of the important parameter R [Journel and Huijbregts, 1978] recommends the value R = 20.

In the exponential case we first approximate

$$Y(ib) = \int_0^{\infty} f(r)dT(r - ib) = \int_0^{4/\lambda} f(r)dT(r - ib)$$

the error associated with this approximation can be estimated by

$$\frac{E\left[\left(\int_{4/\lambda}^{\infty} f(r)dT(r - ib)\right)^2\right]}{E\left[\left(\int_0^{4/\lambda} f(r)dT(r - ib)\right)^2\right]} = \frac{\int_{4/\lambda}^{\infty} f(r)^2 dr}{\int_0^{4/\lambda} f(r)^2 dr} = \frac{25e^{-8}}{1-25e^{-8}} \approx 3.3E-4.$$

Approximating further

$$Y(ib) = \int_0^{4a} f(r)dT(r - ib) = \sum_{k=0}^{8R} \int_{kb}^{(k+1)b} f(r)dT(r - ib) \approx$$

$$\sum_{k=0}^{8R} f\left(\left(k + \frac{1}{2}\right)b\right) t_{k-i} \qquad 2.2.2.5$$

and we require that

$$(8R + 1)b = 4a. \qquad 2.2.2.6$$

where the value of the parameter R again is recommended to be chosen equal to 20 by [Journel and Huijbregts, 1978].

These formulas differ in details only from those of [Journel and Huijbregts, 1978]. However the formula 2.2.2.3 together with 2.2.2.4 generates five to ten times more accurate values than the corresponding formulas in [Journel and Huijbregts, 1978] due to better integral approximation. Unfortunately no such cheap improvements resulted by using the formula 2.2.2.5 together with 2.2.2.6 instead of the counterparts in [Journel and Huijbregts, 1978]. As pointed out in the end of 2.2 one should try to improve the scheme for numerical integration in the exponential case.

2.2.3. Corrections of the approximative formulas

In order to check and correct the approximations above resulting in the formulas 2.2.2.3 and 2.2.2.5 we calculate the covariances from the approximative expressions directly. Starting with the spherical case we note the formula

$$E[Y(ib)Y(ib + sb)] = \sum_{k'=-R}^{R} \sum_{k=-R}^{R} f(kb)f(k'b)E\left[t_{k-i}\, t_{k'-i-s}\right].$$

Since

$$E\left[t_{k-i}\, t_{k'-i-s}\right] = \begin{cases} \sigma_t^2 & \text{if } k + s = k' \\ 0 & \text{if } k + s \neq k' \end{cases}$$

we have in the case s > 0 (note that s is an integer)

$$E[Y(ib)Y(ib+sb)] = \sigma_t^2 \sum_{k=-R}^{R} f(kb)f((k+s)b) =$$

$$\sigma_t^2 b^2 \frac{12V}{a^3} \sum_{k=-R}^{R-s} k(k+s) = \quad {}^4$$

$$\sigma_t^2 b^2 \frac{12V}{a^3}\left(\frac{s^3}{6} - \left(R^2+R+\frac{1}{6}\right)s + \frac{1}{3}(2R^3+3R^2+R)\right) \qquad 2.2.3.1$$

and in particular by putting s = 0 we obtain the variance as

$$E[Y(ib)^2] = \sigma_t^2 b^2 \frac{4V}{a^3}(2R^3+3R^2+R).$$

This should be compared with the required value V.

The approximative formula 2.2.2.3 is then corrected with a multiplicative constant

$$\left[\sigma_t^2 b^2 \frac{4}{a^3}(2R^3+3R^2+R)\right]^{-1/2}$$

in order to give the correct value of the variance. Thus

$$Y(ib) = \sqrt{\frac{3V}{\sigma_t^2(2R^3+3R^2+R)}} \sum_{k=-R}^{R} kt_{k-i}$$

which is the formula used in HYDRASTAR for simulating processes with spherical covariances. The value of R currently used is 20, the stochastic variables t_k are uniformly distributed in the interval [-0. 5, 0. 5] and thus $\sigma_t = 1/12$.

Turning to the exponential model we perform a similar computation

$$E[Y(ib)Y(ib+sb)] = \sigma_t^2 \sum_{k=0}^{8R-s} f\left(\left(k+\tfrac{1}{2}\right)b\right)f\left(\left(k+\tfrac{1}{2}+s\right)b\right) =$$

$$4V\lambda\sigma_t^2 e^{\lambda b(s+1)} \sum_{k=0}^{8R-s} \left(1-\lambda b\left(k+\tfrac{1}{2}\right)\right)\left(1-\lambda b\left(k+s+\tfrac{1}{2}\right)\right) e^{-2\lambda bk}.$$

This series is possible to calculate by for instance neglecting the influence of the tail of this power series and using the formula for a geometrical series. Since this involves tedious computations which are prone to error we put s = 0 and write instead

$$E[Y(ib)^2] = 4V\lambda\sigma_t^2 e^{-\lambda b} \sum_{k=0}^{8R} \left(1-\lambda b\left(k+\tfrac{1}{2}\right)\right)^2 e^{-2\lambda bk} = 4V\lambda\sigma_t^2 e^{-\lambda b}\Sigma$$

with

$$\Sigma = \sum_{k=0}^{8R} \left(1-\lambda b\left(k+\tfrac{1}{2}\right)\right)^2 e^{-2\lambda bk}.$$

---

[4]This equality requires some calculations.

Recognizing that this should be equal to V we get the correction factor

$$\left(4\lambda\sigma_t^{\,2}e^{-\lambda b}\Sigma\right)^{-1/2}$$

and thus the formula used in HYDRASTAR for simulating one dimensional processes with a covariance function of exponential type is

$$Y(ib) = \sqrt{\frac{V}{\sigma_t^{\,2}\Sigma}}\sum_{k=0}^{8R}\left(1-\lambda\left(k+\frac{1}{2}\right)b\right)e^{-\lambda kb}t_{k-i}\,.$$

where the value of R and the stochastic variables t_k are the same as the ones used in the generation of the spherical model.

2.3. Line generation

Given the possibility to generate the one dimensional processes Y one would then proceed to generate three dimensional realizations from the approximative formula

$$Y(x) = \frac{1}{\sqrt{N}}\sum_{i=1}^{N} Y_i((x,l_i),l_i)\,.$$

Recalling formula 2.1.1 it would be natural to choose the lines l_i in an evenly spread fashion over the unit sphere as if evaluating an integral numerically. However it is not easy to identify directions that are evenly distributed on the unit sphere for an arbitrary number of lines. Therefore this method is usually restricted to taking N = 15 lines joining the mid-points of the opposite edges of a regular icosahedron which is the regular polyhedron with the maximum number of faces (in fact 20 faces). We will in the following refer to such a set of lines as an icosahedron set. For a discussion on how to do this in practice see [Journel and Huijbregts, 1978, p. 503]. On the other hand formula 2.1.3 suggests drawing the lines randomly from a uniform distribution on the half unit sphere.

A study of the merits of different schemes for line generation has been made by [Tompson, Ababou and Gelhar, 1989] from which the following discussion is inferred.

The use of N = 15 evenly spaced lines reproduces the mean and variance statistics rather well but the realizations show a number of sets of parallel, linelike patterns. The appearance of these is explained by the following reasoning.

In Fig. 2.3.1 the process of simulating a random function with the turning bands method is depicted in two dimensions and using three lines. It is then clear that when averaging the contributions from the three lines the spatial variability along lines of type A becomes roughly σ if the variability of each line process is σ. However the spatial variability on lines of type B only becomes roughly $\frac{2}{3}\sigma$ since the contribution from the line perpendicular to it, i.e. line 2, is constant in each separate realization. Also the spatial average along the line of type A becomes approximately zero whereas the spatial average along lines of type B is determined by the contribution from line 2.

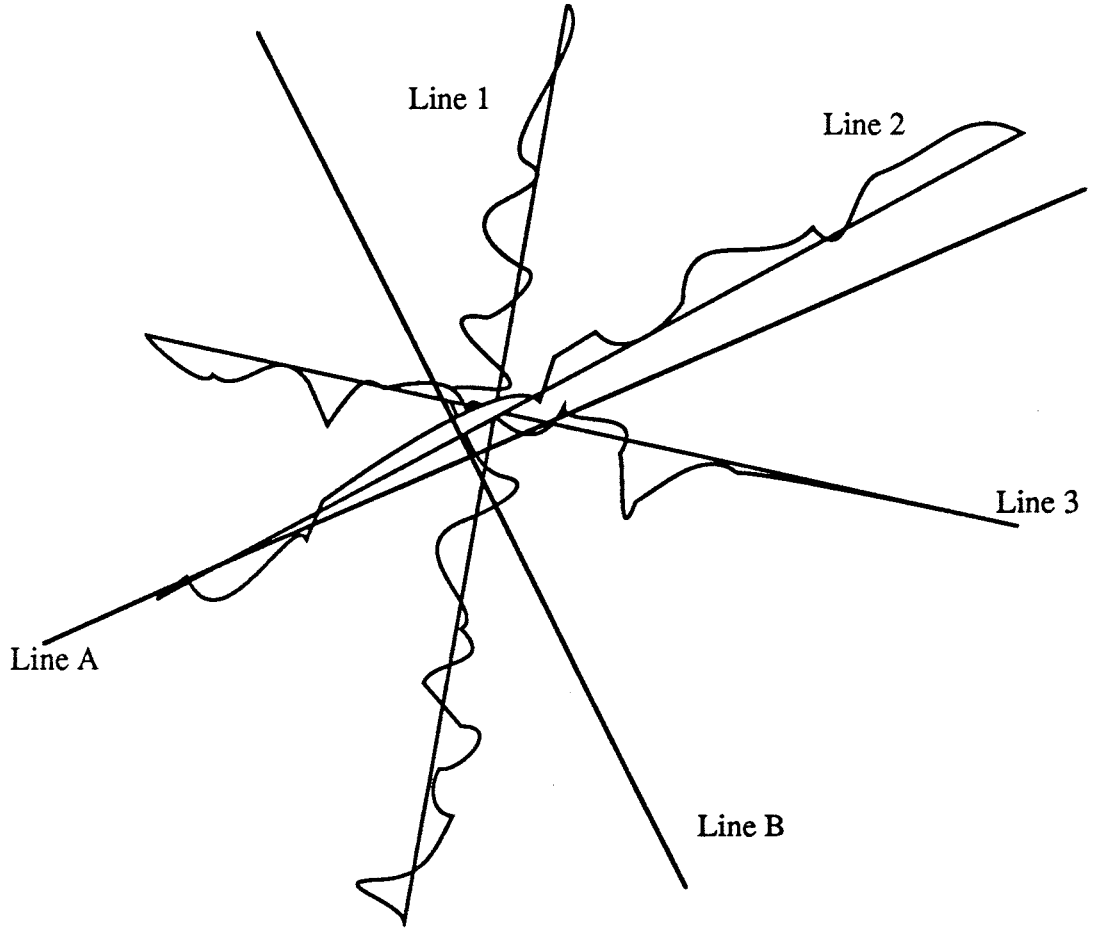


Fig. 2.3.1. Explaining the origin of linelike patterns.

It is important to realize that the magnitude of this effect depends on the number of lines and not on their orientation. We will make use of this fact later on.

The alternative approach of using a larger number of randomly chosen lines, N = 100 seems to be the recommended value, removes this effect but will of course increase the computational burden. Also the variability for the covariance estimates, both between different directions of the lag vectors used in the covariance estimation, and between successive runs are increased in comparison with the case of evenly spaced lines.

Thus it seems to be a good idea to use several icosahedron sets all subjected to a random rotation characterized by a rotation vector chosen from a uniform distribution i.e

$$\omega_j = 2\,\pi U_j \qquad j = 1,2,3$$

where U_j is uniformly distributed over [0, 1] and ω_j, j = 1, 2, 3 are the components of the random rotation vector each giving the rotation angle around the j:th coordinate axis. Results from these different schemes of line generation are shown and discussed in section 6.2.

3. THE HYDROLOGY EQUATION AND ITS NUMERICAL SOLUTION

The hydrology equation solved by HYDRASTAR is

$$\nabla K_s(x)\nabla\langle h\rangle^s(x) = 0 \qquad 3.1$$

where K_S is the isotropic conductivity at the averaging scale s and $\langle h\rangle^s$ is the averaged hydraulic head i.e.

$$\langle h\rangle^s = \frac{1}{|V_s(x) \cap \Omega_f|}\int_{V_s(x)\cap\Omega_f} h(u)\,du$$

where Ω_f is the void space i.e. the fractures, $V_S(x)$ is the averaging volume and $|\cdot|$ is a general notation for volume, or size, of a set. For more details on this see [de Marsily, 1986].

The corresponding integral equation is

$$\int_S K_s(\xi)\nabla\langle h\rangle^s(\xi)\,dS(\xi) = 0 \qquad 3.2$$

where S is the surface of an arbitrary volume V and dS is its directed surface differential.

The integral equation above is solved numerically by a finite difference approximation derived from the specialization of 3.2 to parallelepipeds the boundary surface of which consists of six rectangular faces. These parallelepipeds are referred to as the mass balance elements. This is represented by a staggered prismatic mesh, that is the head nodes are situated at the midpoint of the parallelepipeds and is given by

$$(i_1 - 1)s_1 + (i_2 - 1)s_2 + (i_3 - 1)s_3 \qquad 1 \le i_k \le N_k,\ k = 1,2,3$$

where s_j, j = 1, 2, 3 are the basis vectors of the prismatic mesh and N_k are the number of head nodes in the direction k, k = 1, 2, 3. The conductivities are given on the translated nodes situated on the faces of the parallelepipeds i.e. at

$$\begin{cases} \left(i_1 - \frac{1}{2}\right)s_1 + (i_2 - 1)s_2 + (i_3 - 1)s_3 & 1 \le i_1 \le N_1 - 1,\ 1 \le i_k \le N_k,\ k = 2,3 \\ (i_1 - 1)s_1 + \left(i_2 - \frac{1}{2}\right)s_2 + (i_3 - 1)s_3 & 1 \le i_2 \le N_2 - 1,\ 1 \le i_k \le N_k,\ k = 1,3 \\ (i_1 - 1)s_1 + (i_2 - 1)s_2 + \left(i_3 - \frac{1}{2}\right)s_3 & 1 \le i_3 \le N_3 - 1,\ 1 \le i_k \le N_k,\ k = 1,2 \end{cases}$$

The values of the conductivities at these nodes are denoted

$$\begin{cases} K_1(i_1, i_2, i_3) & 1 \le i_1 \le N_1 - 1,\ 1 \le i_k \le N_k,\ k = 2,3 \\ K_2(i_1, i_2, i_3) & 1 \le i_2 \le N_2 - 1,\ 1 \le i_k \le N_k,\ k = 1,3 \\ K_3(i_1, i_2, i_3) & 1 \le i_3 \le N_3 - 1,\ 1 \le i_k \le N_k,\ k = 1,2 \end{cases}$$

respectively. See Fig. 3.1 where the computational atom is depicted.

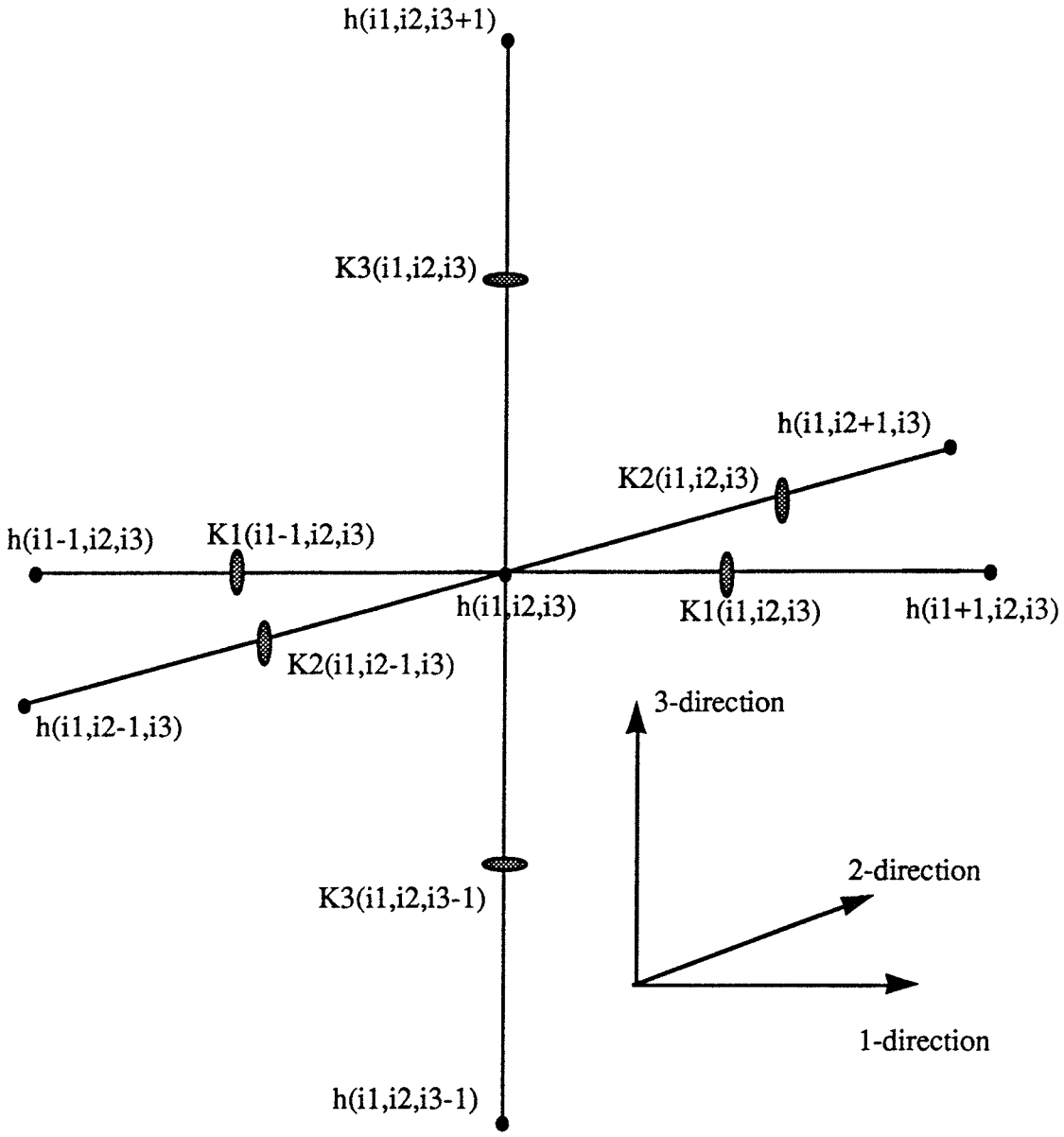


Fig. 3.1. The computational molecule for the hydrology equation 3.1 employed in HYDRASTAR.

The advantage with this is twofold both resulting from the absence of interpolation of conductivity values. First a speed increase results in the solver as a result of the fact that no interpolation takes place. Secondly there is no uncontrolled smoothing of the simulated conductivity field which would result from interpolation. The obvious disadvantage is the increased need for storage.

From these meshes the finite difference equation at the node (i_1, i_2, i_3), $2 \leq i_j \leq N_j - 1$ resulting from 3.1 is written

$$T_1(i_1-1,\, i_2,\, i_3)\, h(i_1-1,\, i_2,\, i_3) + T_1(i_1,\, i_2,\, i_3)\, h(i_1+1,\, i_2,\, i_3) +$$

$$T_2(i_1,\, i_2-1,\, i_3)\, h(i_1,\, i_2-1,\, i_3) + T_2(i_1,\, i_2,\, i_3)\, h(i_1,\, i_2+1,\, i_3) +$$

$$T_3(i_1,\, i_2,\, i_3-1)\, h(i_1,\, i_2,\, i_3-1) + T_3(i_1,\, i_2,\, i_3)\, h(i_1,\, i_2,\, i_3+1) -$$

$$D(i_1,\, i_2,\, i_3)\, h(i_1,\, i_2,\, i_3) = 0 \qquad 3.3$$

where

$$\begin{cases} T_1(i_1, i_2, i_3) = \dfrac{s_2 s_3}{s_1} K_1(i_1, i_2, i_3) \\ T_2(i_1, i_2, i_3) = \dfrac{s_1 s_3}{s_2} K_2(i_1, i_2, i_3) \\ T_3(i_1, i_2, i_3) = \dfrac{s_1 s_2}{s_3} K_3(i_1, i_2, i_3) \end{cases} \qquad 3.4$$

and

$$D(i_1,\, i_2,\, i_3) = T_1(i_1-1,\, i_2,\, i_3) + T_1(i_1,\, i_2,\, i_3) +$$

$$T_2(i_1,\, i_2-1,\, i_3) + T_2(i_1,\, i_2,\, i_3) + T_3(i_1,\, i_2,\, i_3-1) + T_3(i_1,\, i_2,\, i_3) \qquad 3.5$$

Some further information on this topic is contained in section 7.1 These equations are then solved employing a conjugate gradient algorithm, see for instance [Golub and van Loan, p 353 - 380]].

4. ESTIMATION AND CONVERGENCE OF STATISTICS

In this section we shall describe the way ensemble quantities, such as expectation and semivariogram functions, associated with the simulated fields are estimated.

In principle one could regard nodal values of different involved fields as separate stochastic variables. One could then estimate various ensemble moments from a set of independent realizations. In this chapter, let us introduce the notations $X(i_1, i_2, i_3)$, $Y(i_1, i_2, i_3)$ as dummies for two, possibly identical, fields of interest. We could imagine X being hydraulic head and Y being conductivity or for that matter they both being equal to hydraulic head. Then we write the estimates as for instance

$$E[X(i_1, i_2, i_3)] \approx \frac{1}{N}\sum_{j=1}^{N} X^{(j)}(i_1, i_2, i_3) \qquad 4.1$$

or

$$E\left[\tfrac{1}{2}\left(Y(i'_1, i'_2, i'_3) - X(i_1, i_2, i_3)\right)^2\right] \approx$$

$$\tfrac{1}{2}\frac{1}{N}\sum_{j=1}^{N}\left(Y^{(j)}(i'_1, i'_2, i'_3) - X^{(j)}(i_1, i_2, i_3)\right)^2 \qquad 4.2$$

where superindex (j) signifies the j:th simulation. The advantage is that no stationarity assumptions are needed and that the terms in the series are independent.

The obvious disadvantage is that a very large number of realizations are needed to obtain accurate values. This can be compared with the situation of spatial inference where there is only one realization and the ensemble averages of the above type is replaced by spatial averages

$$E[X(\xi)] \approx \frac{1}{N}\sum_{j=1}^{N} X(\xi_j)$$

or

$$E\left[\tfrac{1}{2}\left(Y(\xi + \psi) - X(\xi)\right)^2\right] \approx \tfrac{1}{2}\frac{1}{NL}\sum_{j=1}^{N}\sum_{l=1}^{L}\left(Y(\xi_j + \psi_l) - X(\xi_j)\right)^2.$$

This requires however some assumptions of stationarity. In the first formula we require that $E[X(\xi_j)]$ is approximately constant independent of j and also in order that this estimate should be a good one, that the sample domain, i.e. the domain enclosing the ξ_j:s, should be large in comparison with the correlation scale of $X(\xi)$.

In the second case we must require that the term $\left(Y(\xi_j + \psi_l) - X(\xi_j)\right)^2$ has an approximately constant expectation value, independent of j and l and also, as above, that the sample domain containing the pairs (ξ_j, ψ_l) should be large in comparison to the correlation scale of $\left(Y(\xi_j + \psi_l) - X(\xi_j)\right)^2$.

Summing up, what is needed in order to obtain spatial averages which are good approximations of their ensemble counterparts is that the quantities in question has expectation functions that are approximately constant over several correlation scales.

In order to get the the best out of two worlds the approach here is to mix the two recognizing that even if the condition for obtaining a good spatial average estimate is not fulfilled with regard to sufficiently large sample domains spatial averaging can be used to reduce the variance of the summands in the formulas like 4.1 and 4.2 and thus reduce the number of iterations needed to get a "sufficiently small " variance of the output statistic. That is we replace these formulas with

$$E[X(P)] \approx \frac{1}{N|P|}\sum_{j=1}^{N}\sum_{k\in P} X^{(j)}(\xi_k)$$

and

$$E\left[\tfrac{1}{2}(Y(P+L)-X(P))^2\right] \approx$$

$$\tfrac{1}{2}\frac{1}{N|P||L|}\sum_{j=1}^{N}\sum_{\xi_k\in P}\sum_{\psi_l\in L}\left(Y^{(j)}(\xi_k+\psi_l)-X^{(j)}(\xi_k)\right)^2. \qquad 4.3$$

Here P is a spatial averaging domain, L is an averaging domain in the lag space for which we will use the term lagclass and in agreement with the previously introduced notation $|\cdot|$ signifies the number of points or lags in the averaging domains. However we retain the requirement that the involved expectation functions should be approximately constant over the averaging domains

To exemplify the influence of increasingly larger spatial averaging domains P we show the development of the maximal difference between successive estimates of the semivariogram function for hydraulic conductivity i.e.

$$\max_i\left|\Gamma^{j+1}(P,L_i)-\Gamma^{j}(P,L_i)\right|$$

where $\Gamma^j(P, L_i)$ is the semivariogram estimate using j Monte Carlo iterations, L_i is a lagclass, and P is set of nodes (see 4.3). This is done in a situation where we know that the expectation value function of the conductivity is constant and gives an intuitive image of how the fluctuation of the statistic diminishes.

The diagrams plot

$$\log\frac{\max_i\left|\Gamma^{j+1}(P,L_i)-\Gamma^{j}(P,L_i)\right|}{V}$$

versus log(j). Here V is the known value of the variance of the hydraulic conductivity.

Covariance model:	Spherical
Range parameter, a (see section 2.2.2.):	1. 0
Bandwidth factor, f (see section 6.2.)	1. 0
Bandwidth, T (see section 6.2.):	0. 33
R (see section 2.2.):	20
Number of icosahedron sets (see section 2.3.) :	1
Number of random lines (see section 2.3.) :	85
Spatial sample size(for head), | P | :	27, 729, 3969
Spatial sample size(for head), | L |	1
Number of realizations:	100

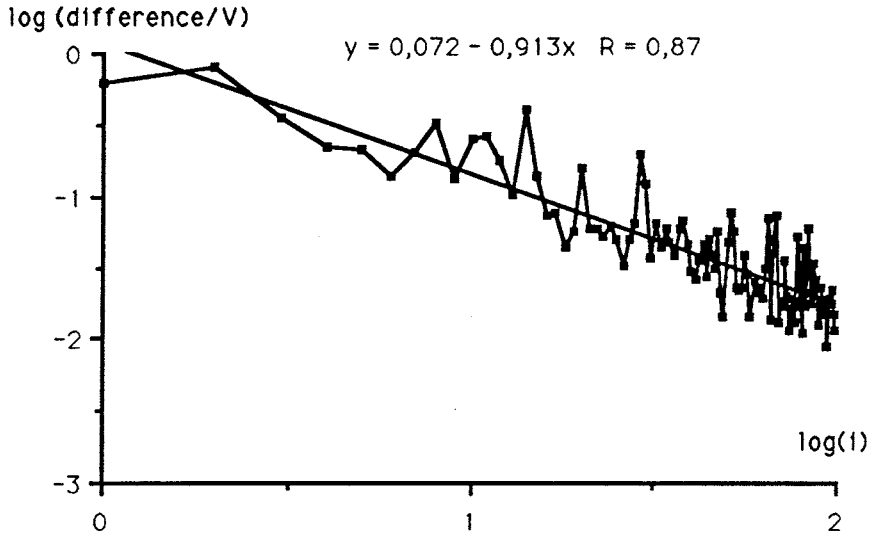


Fig. 4.1. Successive differences, | P | = 27 (points)

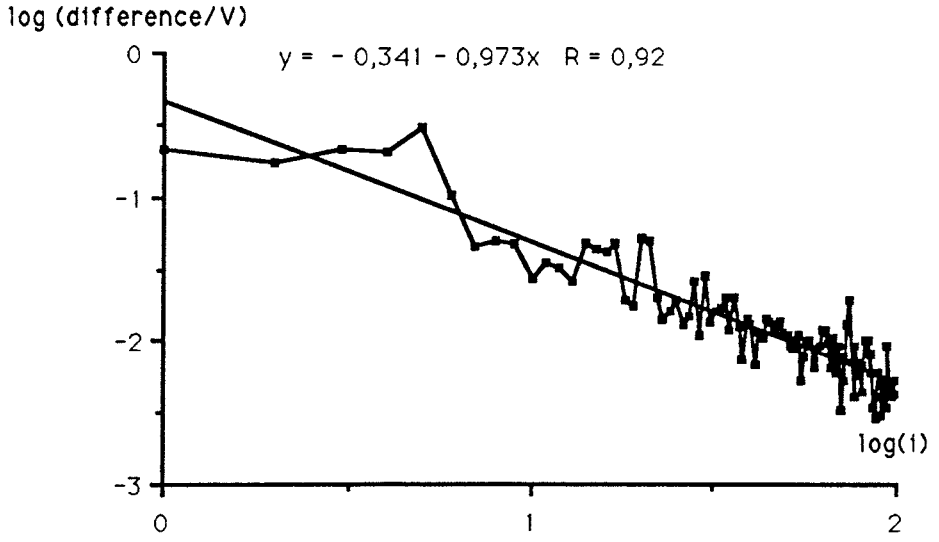


Fig. 4.2. Successive differences, | P | = 729 (points)

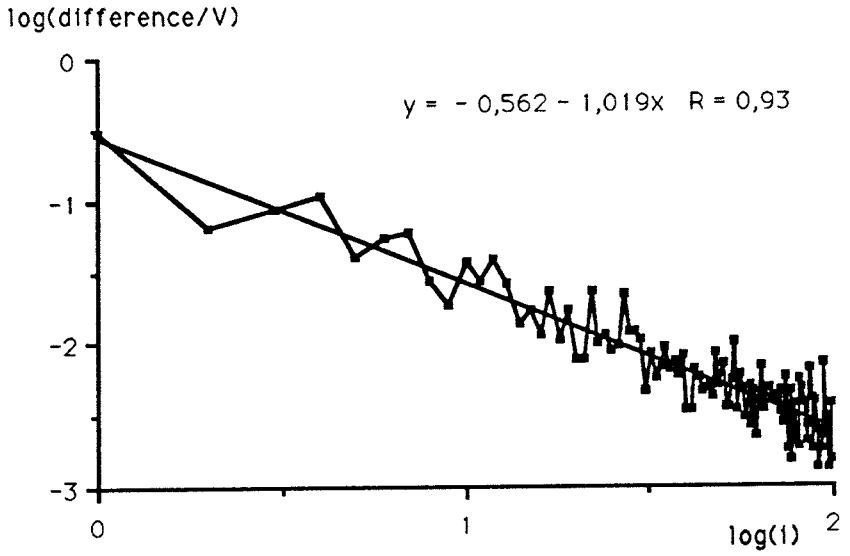


Fig. 4.3. Successive differences, | P | = 3969 points

Note in this connection that the convergence studied here is in terms of maximum norms over the whole domain of definition of the involved semivariogram functions. This is very tough requirements and it is clear that it is the lag class with the least number of lags that determines the number of iterations needed to get "sufficiently small " variance of the output statistic. Thus the governing principles of lag class divisions should be

- As already pointed out the semivariogram function should be approximately constant over a lag class. If we choose too large lagclasses our estimate, which is like the average of the semivariogram functions over a lag class, may deviate substantially from point values of the semivariogram function inside the lag class.
- The number of lags should be approximately constant when comparing between different lag classes in order to achieve a uniform convergence. That is one should avoid to have lagclasses containing very few lags in comparison to others.

5. PERTURBATION SOLUTION

In the following we will develop the first order perturbation solution to the stochastic hydrology equation

$$\nabla K \nabla H = 0, \qquad 5.1$$

which is the same as equation 3.1, in a domain Ω with K (conductivity) being a known stochastic process and H (hydraulic head) being a stochastical process unknown apart from on the boundary of Ω, $\partial\Omega$. This will be done following the references [Gelhar and Axness, 1983], [Bakr et al, 1978] and [Neuman, Winter and Newman, 1987] closely and treating the stochastical processes rather freely. In fact without motivation we will

- assume that our processes are differentiable,
- interchange the order of expectation and differentiation and

For more precise formulations in these matters the reader is referred to [Doob, 1953].

We start by rewriting the equation in terms of the log conductivity Y=ln(K)[5] as

$$\nabla^2 H + \nabla Y \nabla H = 0 \qquad 5.2$$

and then dividing the stochastic functions as

$$Y = \langle Y \rangle + Y'$$

$$H = \langle H \rangle + H'$$

where angular brackets indicates expectation value.

Introducing this in 5.2 we have

$$\nabla^2\langle H\rangle + \nabla^2 H' + \nabla\langle Y\rangle\nabla\langle H\rangle + \nabla\langle Y\rangle\nabla H' + \nabla Y'\nabla\langle H\rangle + \nabla Y'\nabla H' = 0.$$

Now we want to approximate this equation in the case of the primed quantities being small by neglecting products of primed quantities to obtain

$$\nabla^2\langle H\rangle + \nabla^2 H' + \nabla\langle Y\rangle\nabla\langle H\rangle + \nabla\langle Y\rangle\nabla H' + \nabla Y'\nabla\langle H\rangle = 0 \qquad 5.3$$

and as a consequence obtain, by taking the expectation of the above equation,

$$\nabla^2\langle H\rangle + \nabla\langle Y\rangle\nabla\langle H\rangle = 0. \qquad 5.4$$

By using

$$K = \exp(\langle Y\rangle + Y') = -\exp(\langle Y\rangle)\left(1 + Y' + \frac{1}{2!}Y'^2 + \ldots\right)$$

and the expectation of the first order approximation of the expression above

$$\langle K\rangle = \exp(\langle Y\rangle)$$

we can write this analogously with 5.1 as

---

[5]Note the temporary redefinition of Y from log(K) to ln(K).

$$\nabla\langle K\rangle\nabla\langle H\rangle = 0. \qquad 5.5$$

The mean removed form of 5.3 is

$$\nabla^2 H' + \nabla\langle Y\rangle\nabla H' + \nabla Y'\nabla\langle H\rangle = 0 \qquad 5.6$$

or even better

$$\nabla\langle K\rangle\nabla H' = -\langle K\rangle\nabla\langle H\rangle\nabla Y'$$

where $\langle H\rangle$ can be determined from 5.5 and thus is regarded as known.

This is a much simpler problem than the corresponding equation 5.1 since the uncertainty is only present in the right hand side. Thus the solution can be expressed with the aid of a Greens function and the moments of H follow as closed expressions involving the moments of Y. Let us also point out in this connection that the previous first order perturbation approach is equivalent to calculate

$$\frac{dH}{dY}(\langle Y\rangle)\,dY$$

where the derivative involved are a so called Frechet derivative. This is true since the approximations are consistently made to the first order. This is not respected in [Gelhar and Axness, 1983] where the development is an inconsistent mix between first and second order approximations. This is correctly pointed out in [Neuman et al, 1987].

We now insert the extra assumption that Y is second order stationary and in particular that the expectation of Y is constant. This assumption does two things. First it simplifies the appearance of the equations 5.5 and 5.6 even further. They are reduced to

$$\nabla^2\langle H\rangle = 0 \qquad 5.7$$

and

$$\nabla^2 H' + \nabla Y'\nabla\langle H\rangle = 0 \qquad 5.8$$

respectively. Secondly, adding the assumption that the studied domain Ω is the whole three dimensional space, $\mathbf{R}^3$, we infer that Y´ has a spectral representation and thus we are in a position to solve the equation 5.8 by spectral techniques. In order to do so we must unfortunately introduce some more assumptions and these are of a worse kind than the previous ones. The reason for this is that they are assumptions on the solution. This kind of assumptions, especially of stationarity, is quite common in the field of stochastic partial differential equation and are known as dishonest hypotheses, see [Beran]. Thus the first dishonest assumption is to assume that H´ is stationary and the second that $\nabla\langle H\rangle$ is a constant. The second assumption can be derived from another assumption namely that $\langle H\rangle$ equals a linear function at infinity. This follows from 5.7.

Accepting these assumptions we have that, according to the spectral theorem for stationary processes [Yaglom, 1962], there exist complex random measures $dZ_Y(\xi)$ and $dZ_H(\xi)$ with the properties

$$E\left(dZ_H(M)\right) = 0,$$

$$E\left(dZ_H(M)\,dZ_H^*(M')\right) = 0 \qquad \text{if } M \cap M' = \varnothing,$$

$$E\left(dZ_Y(M)\right)=0,$$

$$E\left(dZ_Y(M)dZ_Y{}^{*}(M')\right)=0 \qquad \text{if } M\cap M'=\varnothing,$$

where superindex * signifies complex conjugation and moreover

$$Y'(x)=\int_{R^3} e^{ix\cdot\xi}\,dZ_Y(\xi)$$

$$H'(x)=\int_{R^3} e^{ix\cdot\xi}\,dZ_H(\xi).$$

It is true under fairly general conditions that

$$\nabla Y'(x)=\int_{R^3} i\xi e^{ix\cdot\xi}\,dZ_Y(\xi)$$

$$\nabla^2 H'(x)=\int_{R^3} -|\xi|^2 e^{ix\cdot\xi}\,dZ_H(\xi)$$

and thus that the equation 5.8 has as its spectral counterpart

$$\xi^2 dZ_H(\xi)=-i\xi\nabla\langle H\rangle dZ_Y(\xi)$$

or

$$dZ_H(\xi)=-\frac{i}{\xi^2}\xi\nabla\langle H\rangle dZ_Y(\xi). \qquad 5.9$$

where ξ is the wave vector and products involving it are scalar products. We stress that this presumes that $\nabla\langle H\rangle$ is a constant, that the primed quantities are stationary with zero expectation value and of course that the domain is all of R^3. From 5.9 we now obtain the relations between the spectral densities as

$$S_H(\xi)\,d\xi=\frac{1}{\xi^4}\left(\xi\nabla\langle H\rangle\right)^2 S_Y(\xi)\,d\xi \qquad 5.10$$

where the spectral densities are defined by

$$\begin{cases} S_H(\xi)d\xi = E\left[dZ_H\,dZ_H{}^{*}\right] \\ S_Y(\xi)d\xi = E\left[dZ_Y\,dZ_Y{}^{*}\right]. \end{cases}$$

In an analogous fashion we may also express the spectral density for the Darcy velocity in terms of the spectral density for the conductivity. Starting with Darcy's law we have

$$q=-\exp(\langle Y\rangle+Y')(\nabla\langle H\rangle+\nabla H')=$$

$$-\exp(\langle Y\rangle)\left(1+Y'+\tfrac{1}{2!}Y'^2+\ldots\right)(\nabla\langle H\rangle+\nabla H')$$

and neglecting terms of second order as above we obtain

$$q = -\exp(\langle Y \rangle)(\nabla\langle H \rangle + \nabla H' + Y'\nabla\langle H \rangle).$$

We get for the average

$$\langle q \rangle = -\exp(\langle Y \rangle)\nabla\langle H \rangle \qquad 5.11$$

and the average removed form becomes

$$q' = -\exp(\langle Y \rangle)(Y'\nabla\langle H \rangle + \nabla H'). \qquad 5.12$$

We also point out that 5.7 and 5.8 or 5.5 and 5.6 in the case of nonconstant expectation of the log conductivity follows from 5.11 and 5.12 if we invoke the perturbed form of the continuity equation for incompressible stationary flow i.e.

$$\nabla\langle q \rangle = 0$$

and

$$\nabla q' = 0.$$

The spectral relationship corresponding to 5.12 is deduced, again under the assumption that $\langle Y \rangle$ and $\nabla\langle H \rangle$ are constants to be

$$dZ_q(\xi) = -\exp(\langle Y \rangle)(dZ_Y(\xi)\nabla\langle H \rangle + i\xi dZ_H(\xi)) =$$

$$\langle q \rangle dZ_Y(\xi) - \exp(\langle Y \rangle)\, i\xi dZ_H(\xi)$$

where we also used 5.11 in the last equality. This is a vector relationship, in particular the $dZ_q(\xi)$ is a vector valued measure with three components. This results in a tensor relationship for the spectral densities after first using 5.9 and 5.11 to obtain

$$dZ_q(\xi) = \left(1 - \xi\frac{1}{\xi^2}\xi^T\right)\langle q \rangle dZ_Y(\xi) \qquad 5.13$$

we then have

$$S_q(\xi) = \left(1 - \xi\frac{1}{\xi^2}\xi^T\right)\langle q \rangle\langle q \rangle^T\left(1 - \xi\frac{1}{\xi^2}\xi^T\right)S_Y(\xi). \qquad 5.14$$

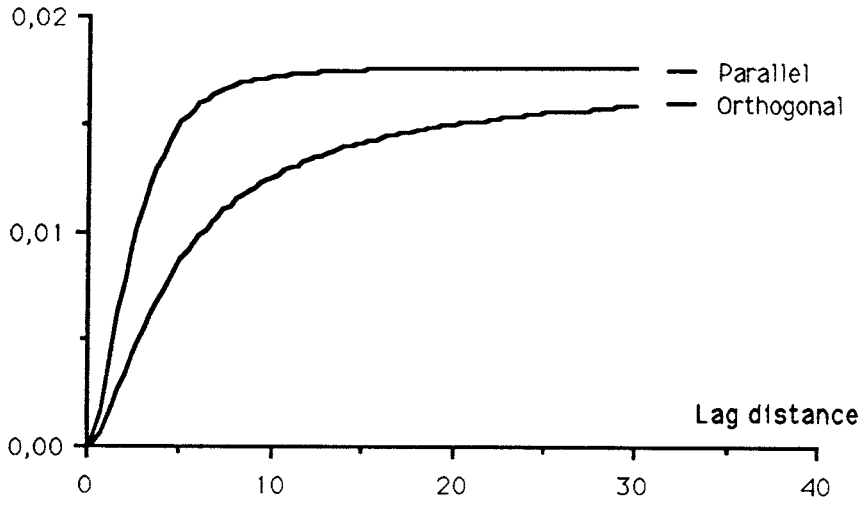


Fig. 5.1. Diagram showing the head semivariogram functions obtained from the analytical solution 5.15 in the direction parallel and orthogonal to the head gradient using the following parameters:

Covariance model:	Exponential
Range parameter, a (see 2.2.2.):	1
Variance of 10logK:	0. 01
Gradient, $\|\nabla H\|$:	1
Head variance, σ_H^2:	0. 0177

Now in order to use these results in verification purposes we want to translate the spectral relationships into relationships between the covariances. In [Bakr et al, 1978] the resulting covariance for the head is calculated given an exponential covariance model for the conductivity. Explicitly

$$R_Y(x) = \sigma_Y^2 e^{-|x|/\lambda}$$

results in a spectral density for the conductivity

$$S_Y(\xi) = \left(\frac{\sigma_Y}{\pi}\right)^2 \frac{\lambda^3}{\left(1+\lambda^2|\xi|^2\right)^2}$$

and thus from 5.10 we obtain the spectral density for the hydraulic head as

$$S_H(\xi) = -\left(\frac{\sigma_Y}{\pi}\right)^2 \frac{(\xi \nabla\langle H\rangle)^2}{|\xi|^4} \frac{\lambda^3}{\left(1+\lambda^2|\xi|^2\right)^2}$$

which finally leads to a covariance for the head as given in [Bakr et al, 1978].

$$C_H(r,\chi)=\frac{\|\nabla(H)\|^2\sigma_Y{}^2\lambda^2}{2}\left\{(-\sin^2\chi)\left(e^{-\rho}+(e^{-\rho}-1)\frac{2}{\rho}\right)+\right.$$

$$\left.(3\cos^2\chi-1)\left[(1-e^{-\rho})\frac{8}{\rho^3}-e^{-\rho}\left(1+\frac{4}{\rho}+\frac{8}{\rho^2}\right)\right]\right. \qquad 5.15$$

where we have introduced the auxiliary notation

$$\rho=\frac{r}{\lambda}.$$

We note in particular that letting r tend to zero we obtain the head variance as

$$\sigma_H{}^2=\frac{\|\nabla H\|\sigma_Y{}^2\lambda^2}{3}.$$

The graph of the semivariogram corresponding to this covariance function is shown in Fig. 5.1.

6. RESULTS OF STOCHASTIC SIMULATIONS.

6.1. One dimensional realizations

In order to test the line process generator described in section 2.2 and the different random number generators described in section 2.2.1 some runs were made for the spherical and exponential case and the covariances or equivalently the semivariograms were estimated on the lag distances kT, k = 1, 2, 3. . N_{ch}. Here T signifies the bandwidth and the integer N_{ch} is the approximative number of bands over which the line processes are correlated. The integer N_{ch} is given by

$$N_{ch} = \frac{2R}{N_q} + 1$$

in the spherical case and

$$N_{ch} = \frac{6R}{N_q} + 1$$

in the exponential case. We used the rule of thumb that the practical range of a process with an exponential type covariance is three times its range parameter, a. For notations see sections 2.2 and 2.2.2, in particular equations 2.2.4, 2.2.2.4 and 2.2.2.6.

The parameter choices involved when using the convolution method is described in section 2.2. In fact for the one dimensional processes the only interesting parameter is R which determines the accuracy of the approximation of the convolution integrals in equations 2.2.2.3 and 2.2.2.5. For this parameter [Journel and Huijbregts, 1978] recommend the value R = 20 in both the exponential and spherical cases and this value is consistently used in the following diagrams. The random numbers t_k is consistently chosen as uniformly distributed in the interval [-1/2, 1/2].

The diagrams on the following two pages 6.1.1 - 6.1.4 show the one dimensional semivariogram as a function of lag distance using a spherical model and the two different random number generators described in section 2.2.1. We note that the estimated semivariograms are very close to the exact ones and thus that the approximation of the convolution integral 2.2.2.3 is quite sufficient. It should even be possible to reduce the number of points in the approximation. Moreover we note that the random number generators 2.2.1.1 and 2.2.1.2 give similar result and this is a verification of the random number generator 2.2.1.1.

The next two pages, diagrams 6.1.5 - 6.1.8 show the same experiments but for the exponential case. Here we draw the same conclusions about the random number generator but we note that the agreement between the exact and estimated values has been substantially deteriorated in comparison to the spherical case in spite of the much larger number of points used in the approximation 2.2.2.5 as compared to 2.2.2.3. The reason for this is probably that the equidistant approximation used in the approximations of the convolution integrals is suited for the function given in 2.2.2.1 whereas it is somewhat unsuited for the function given in 2.2.2.2.

Covariance model:	Spherical
Range parameter, a (see section 2.2.2.)	1. 0
R (see section 2.2.2.)	20
Bandwidth, T (see section 2.2.)	0. 333
b (see section 2.2.)	0. 025
N_q (see section 2.2.)	3
Random number generator	2.2.1.2.
Random measure, t_k, (see section 2.2.)	U(-1/2, 1/2)
σ_t, see 2.2.	1/12

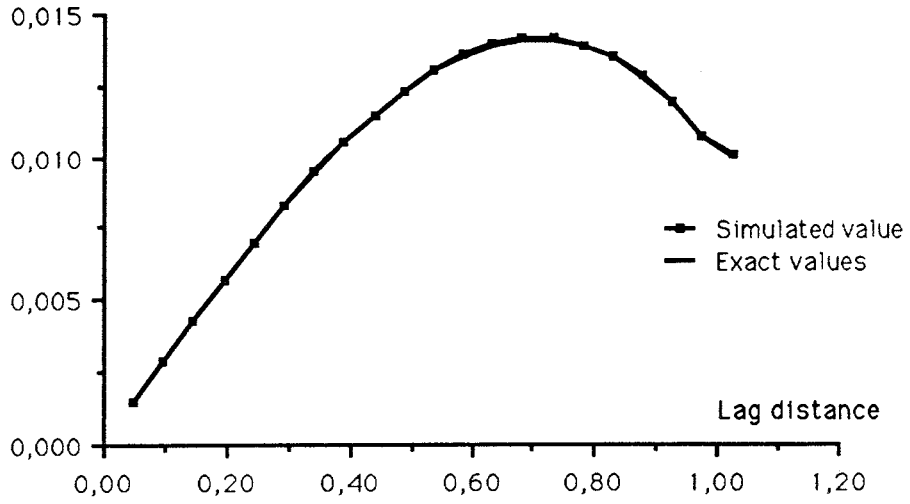


Fig. 6.1.1. One dimensional realizations in the spherical case with the random number generator 2.2.1.2. Exact and estimated values for the one dimensional semivariogram as a function of lag distance.

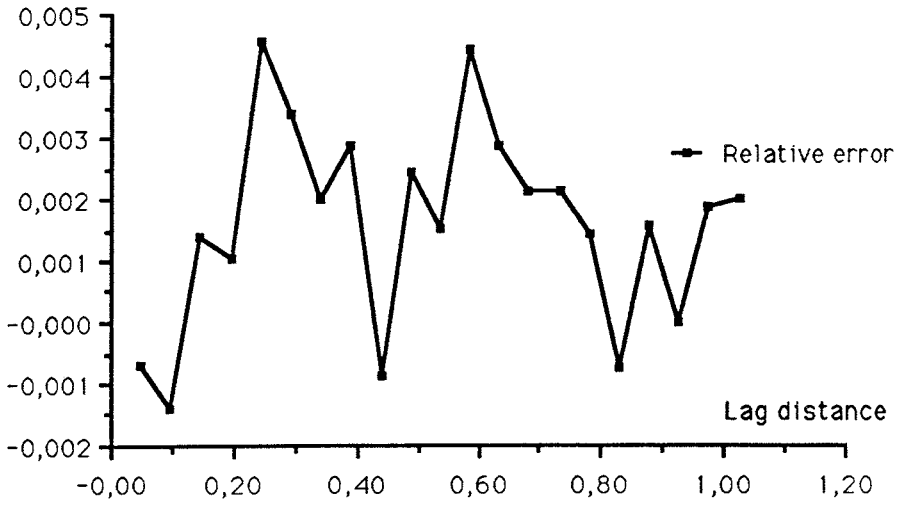


Fig. 6.1.2. One dimensional realizations in the spherical case with the random number generator 2.2.1.2. Relative error as a function of lag distance.

Covariance model:	Spherical
Range parameter, a (see section 2.2.2.)	1. 0
R (see section 2.2.2.)	20
Bandwidth, T (see section 2.2.)	0. 333
b (see section 2.2.)	0. 025
N_q (see section 2.2.)	3
Random number generator	2.2.1.1.
Random measure, t_k (see section 2.2.)	U(-1/2, 1/2)
σ_t (see section 2.2.)	1/12

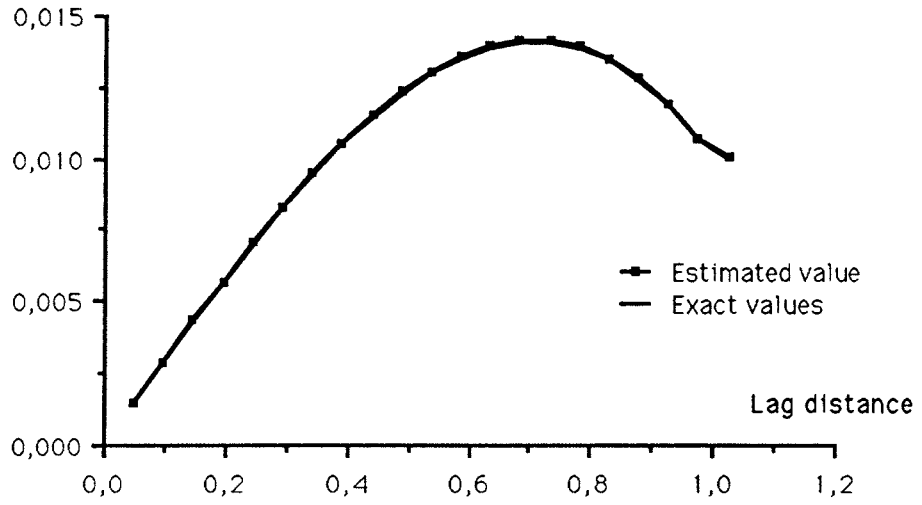


Fig. 6.1.3. One dimensional realizations in the spherical case with the random number generator 2.2.1.1. Exact and estimated values for the one dimensional semivariogram as a function of lag distance.

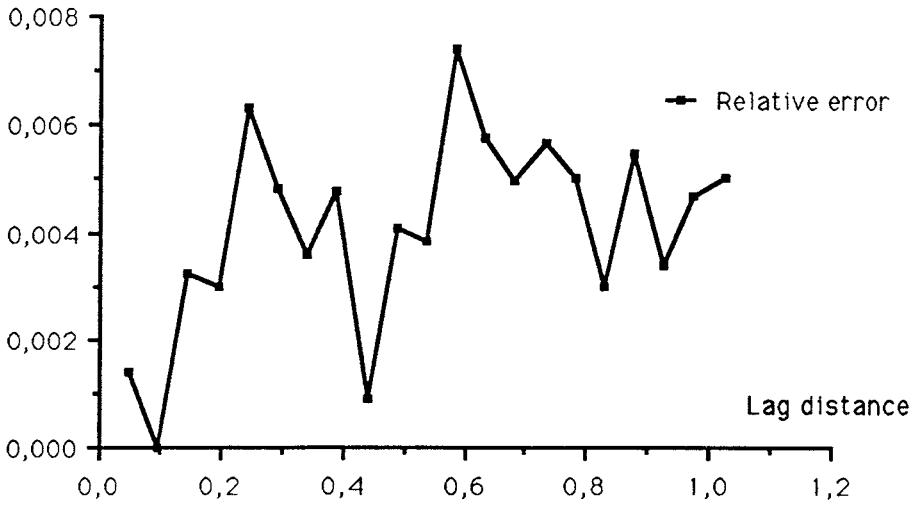


Fig. 6.1.4. One dimensional realizations in the spherical case with the random number generator 2.2.1.1. Relative error as a function of lag distance.

Covariance model:	Exponential
Range parameter, a (see section 2.2.2.)	1. 0
R (see section 2.2.2.)	20
Bandwidth, T (see section 2.2.)	0. 333
b (see section 2.2.)	0. 025
N_q (see section 2.2.)	3
Random number generator	2.2.1.2.
Random measure, t_k (see section 2.2.)	U(-1/2, 1/2)
σ_t (see section 2.2.)	1/12

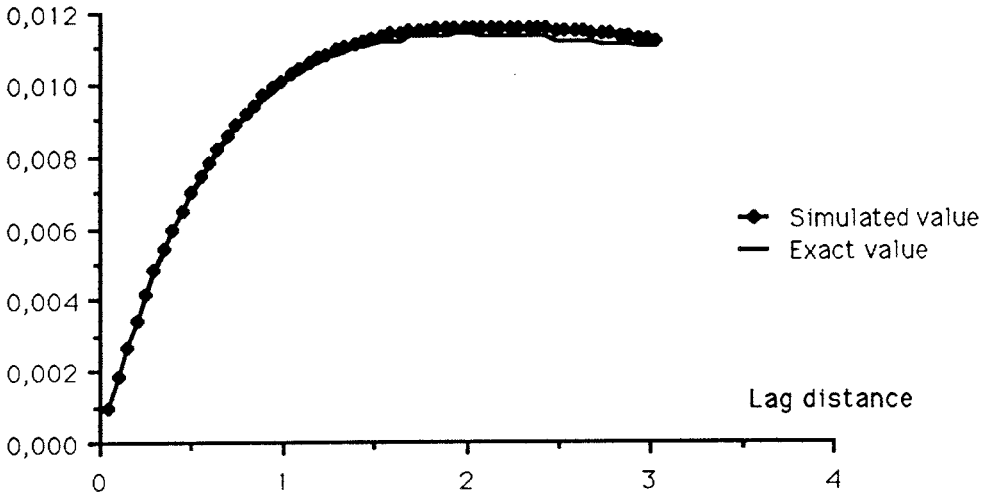


Fig. 6.1.5. One dimensional realizations in the exponential case with the random number generator 2.2.1.2. Exact and estimated values for the one dimensional semivariogram as a function of lag distance.

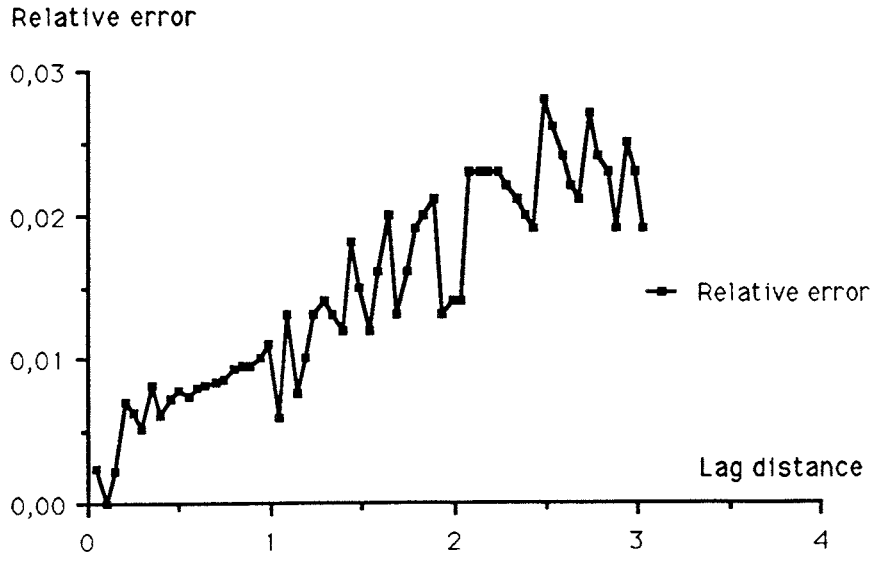


Fig. 6.1.6. One dimensional realizations in the exponential case with the random number generator 2.2.1.2. Relative error as a function of lag distance.

Covariance model:	Exponential
Range parameter, a (see section 2.2.2.)	1. 0
R (see section 2.2.2.)	20
Bandwidth, T (see section 2.2.)	0. 333
b (see section 2.2.)	0. 025
N_q (see section 2.2.)	3
Random number generator	2.2.1.1.
Random measure, t_k (see section 2.2.)	U(-1/2, 1/2)
σ_t (see section 2.2.)	1/12

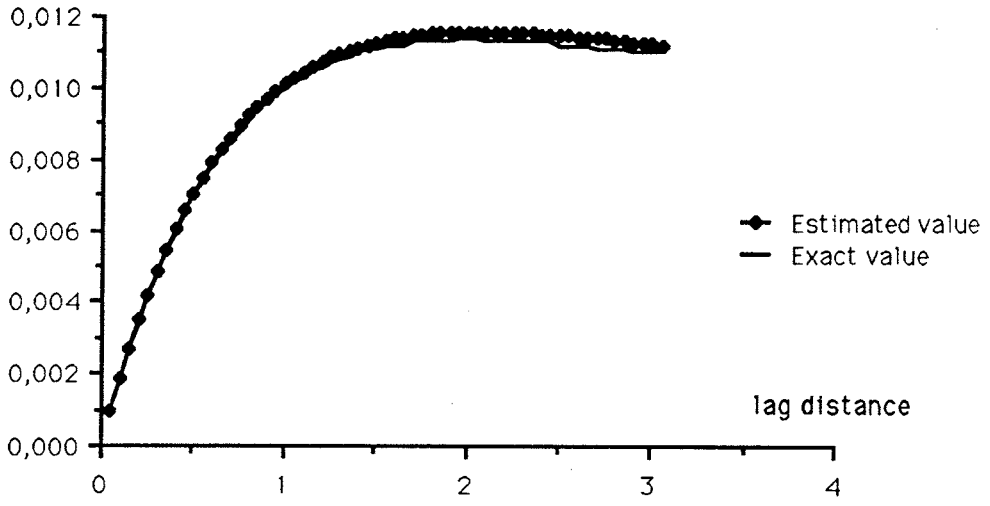


Fig. 6.1.7. One dimensional realizations in the exponential case with the random number generator 2.2.1.1. Exact and estimated values for the one dimensional semivariogram as a function of lag distance.

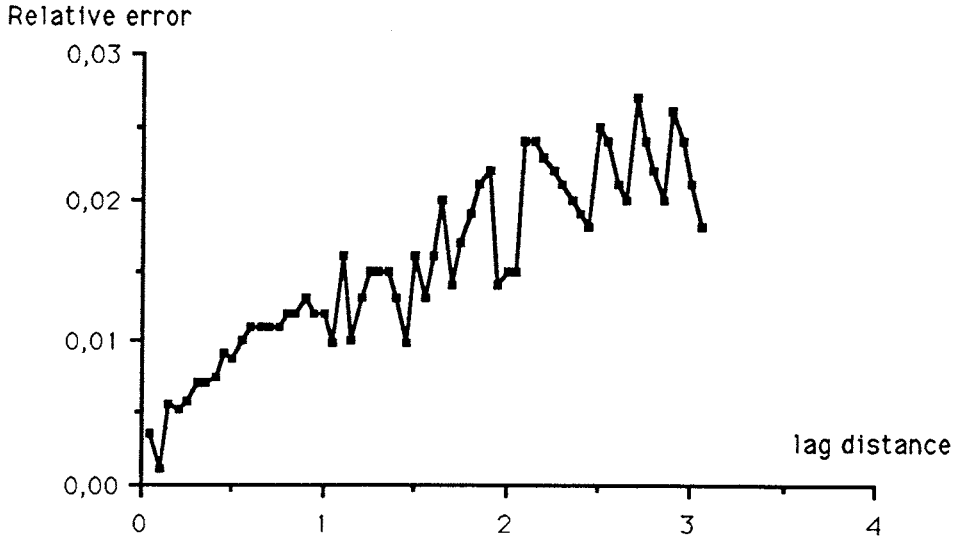


Fig. 6.1.8. One dimensional realizations in the exponential case with the random number generator 2.2.1.1. Relative error as a function of lag distance.

6.2. Three dimensional realizations

The goal of the analysis is twofold

- Show that we can reproduce the semivariograms given as input to the simulations by performing estimation of the output according to chapter 4.
- Make suggestions for choice of parameters.

Regarding the second point above this report has a serious drawback since no result is proved to be statistically significant. Then there is always the standard problem with varying parameters i.e. the number of cases grows very rapidly.

The three dimensional conductivity realizations in this section were all performed on a prismatic mesh with step vectors s_j primarily equal to 1/3, see chapter 3. The resulting covariance functions were estimated in two directions parallel to the step vectors. We have found that typically the covariances are too small as compared with the desired values. This deficiency can be diminished by decreasing the bandwidth, T, but only up to a certain point. The bandwidth is usually recommended to be chosen as the minimal distance between simulation points [Journel and Huijbregts, 1978] i.e. in a prismatic mesh

$$T = f \min (s_1, s_2, s_3)$$

where we introduced a bandwidth factor f less than one in order to have some flexibility.

A characteristic situation is shown in the diagrams 6.2.1 - 6.2.2. The covariance function in the 1 - direction shows to low values for the lags approaching the range. The too high values seen for the larger ranges are less serious since the convergence there is not so good due to a smaller number of samples. As is readily seen the semivariogram in the 2 - direction shown in diagrams 6.2.3 - 6.2.4 looks better. More important is the fact that it looks significantly different from the first one. However, here the word significantly does not attach to it any formal statistical meaning as it should. That is, ideally we should be able to produce the formal statement "the semivariograms in the different directions are significantly different on the ε -level".

The generated semivariograms are thus apparently anisotropic. This can only be due to poor convergence as discussed in in section 2.3. We note also that this takes place in spite of the large number of realizations and the good looking convergence curves, see diagram 4.3. This is the disadvantage with using random lines.

On the other hand the main advantage with random lines is that one avoids linelike patterns however this is dependent only on the number of lines as noted in section 2.3. From this stems the idea already mentioned in section 2.3 to use several icosahedron sets all subjected to a random rotation characterized by a rotation vector chosen from a uniform distribution i.e.

$$\omega_j = 2\pi U_j \qquad j = 1,2,3$$

where U_j is uniformly distributed over [0, 1] and ω_j j = 1, 2, 3 are the components of the random rotation vector. Results obtained from testing this idea is shown in the diagrams 6.2.5 - 6.2.8 These show some clear improvements in particular the apparent anisotropy seems to have vanished. However, nothing has been done yet in order to see if this improvement is significant.

For the exponential covariance model we present the example shown in the diagrams 6.2.9 - 6.2.12 which have been performed using the line generation advocated above, that is four random icosahedron sets were used Otherwise the difference with the previous cases are that the prismatic mesh in this case has unit step length, i.e. $s_j = 1$ $j = 1, 2, 3$ and that the spatial sample size is much smaller than in the previous section.

Finally in diagram 6.2.13 a generated histogram is shown meant to indicate the Gaussian behaviour of the simulated fields .

Covariance model:	Spherical
Range parameter, a (see section 2.2.2.):	4. 0
Bandwidth factor, f:	0. 2
Bandwidth, T:	0. 066
R (see section 2.2.2.):	60
Number of icosahedron sets (see section 2.3.) :	1
Number of random lines (see section 2.3.) :	85
(Maximal) Spatial sample size, | P | (see chapter 4.) :	3969
Spatial sample size, |L| (see chapter 4.)	1
Number of realizations:	100

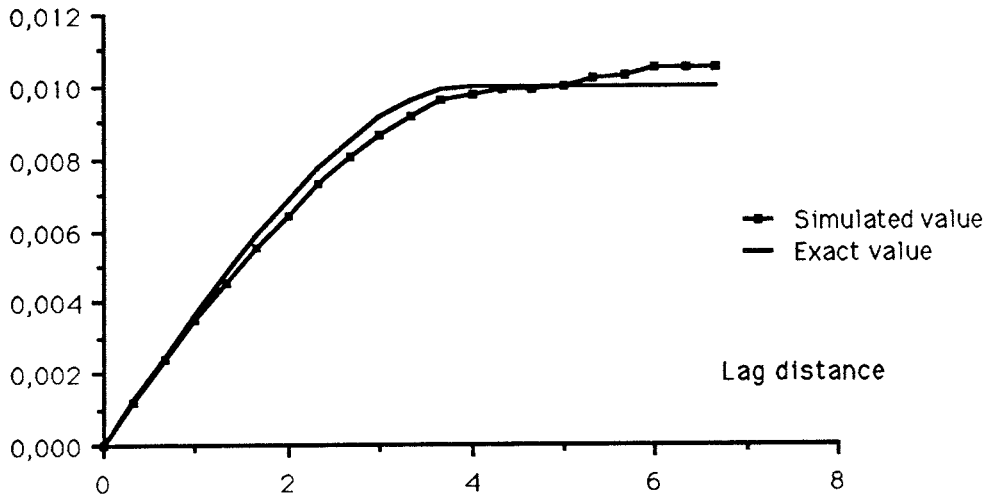


Fig. 6.2.1. Three dimensional realizations in the spherical case. Exact and simulated values of the semivariogram for the conductivity in the 1-direction.

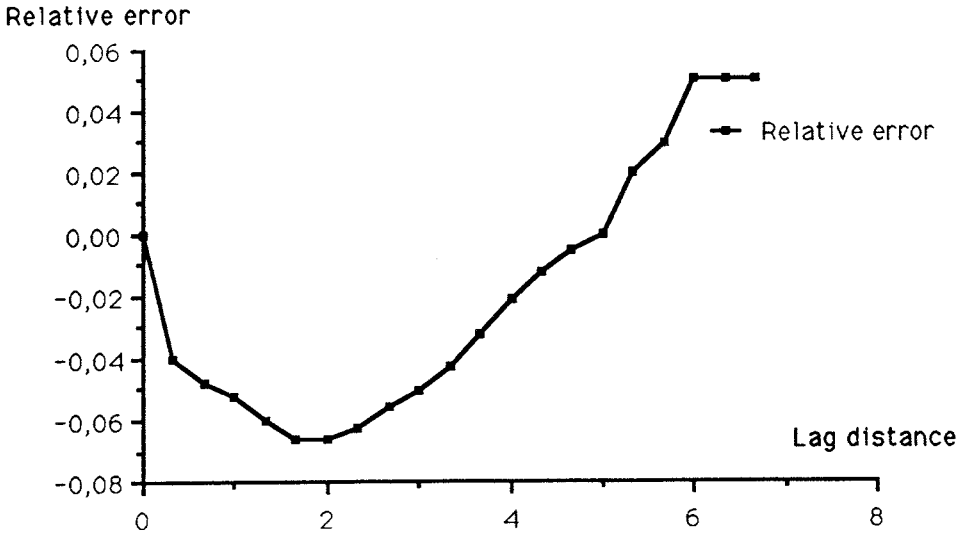


Fig. 6.2.2. Three dimensional realizations in the spherical case. Relative errors of the semivariogram for the conductivity in the 1- direction.

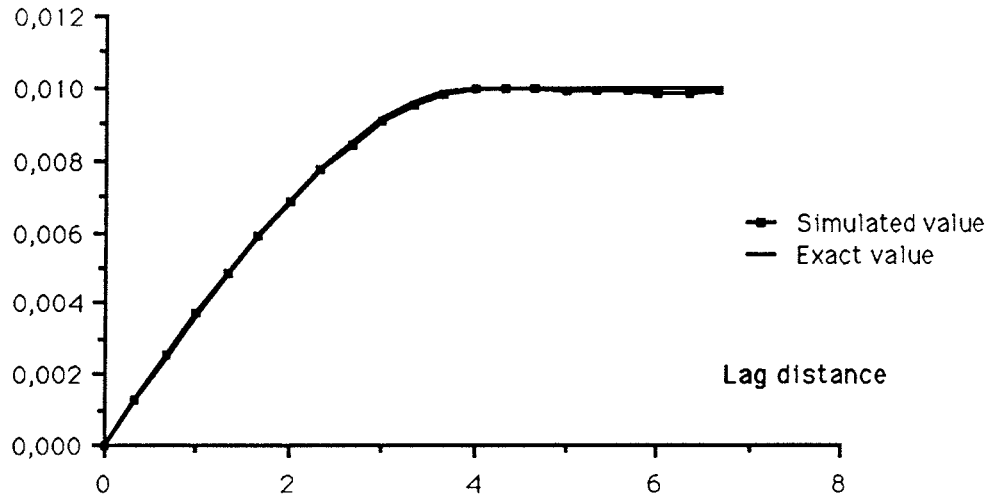


Fig. 6.2.3. Three dimensional realizations in the spherical case. Exact and simulated values of the semivariogram for the conductivity in the 2-direction.

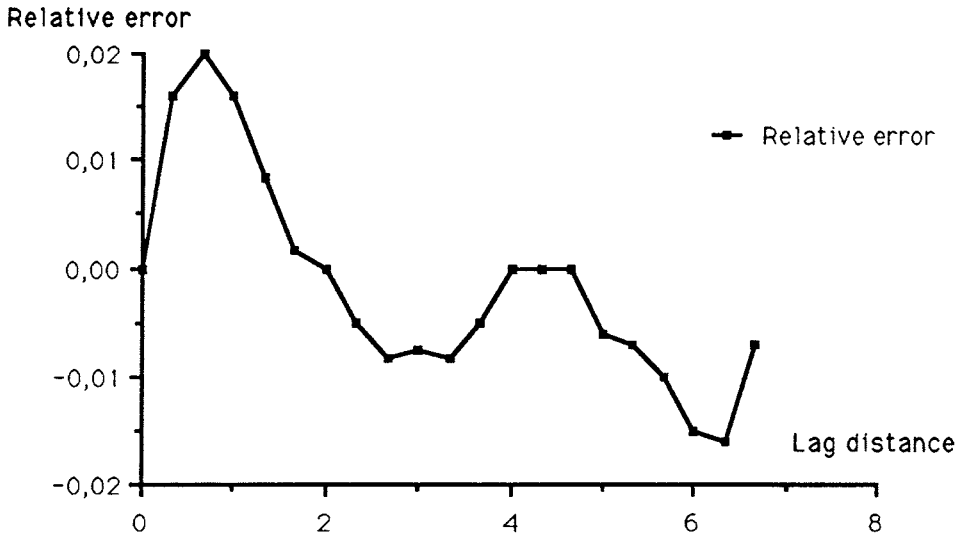


Fig. 6.2.4. Three dimensional realizations in the spherical case. Relative errors of the semivariogram for the conductivity in the 2- direction.

Covariance model:	Spherical
Range parameter, a (see section 2.2.2.):	4. 0
Bandwidth factor, f:	0. 2
Bandwidth, T:	0. 066
R (see section 2.2.2.):	60
Number of icosahedron sets (see section 2.3.) :	4
Number of random lines (see section 2.3.) :	0
(Maximal) Spatial sample size, | P | (see chapter 4.):	3969
Spatial sample size, |L| (see chapter 4.)	1
Number of realizations:	100

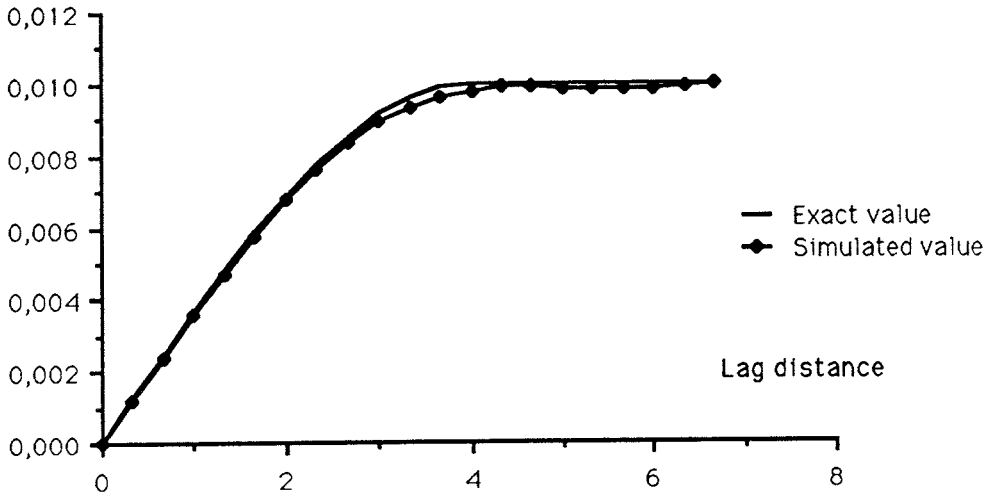


Fig. 6.2.5. Three dimensional realizations in the spherical case using four icosahedron sets. Exact and simulated values of the semivariogram for the conductivity in the 1- direction.

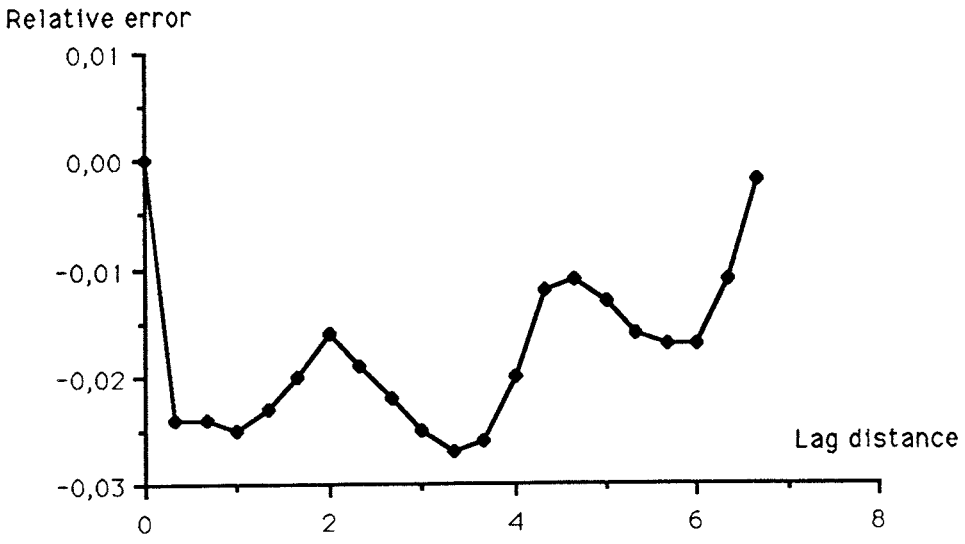


Fig. 6.2.6. Three dimensional realizations in the spherical case using four icosahedron sets. Relative errors of the semivariogram for the conductivity in the 1- direction.

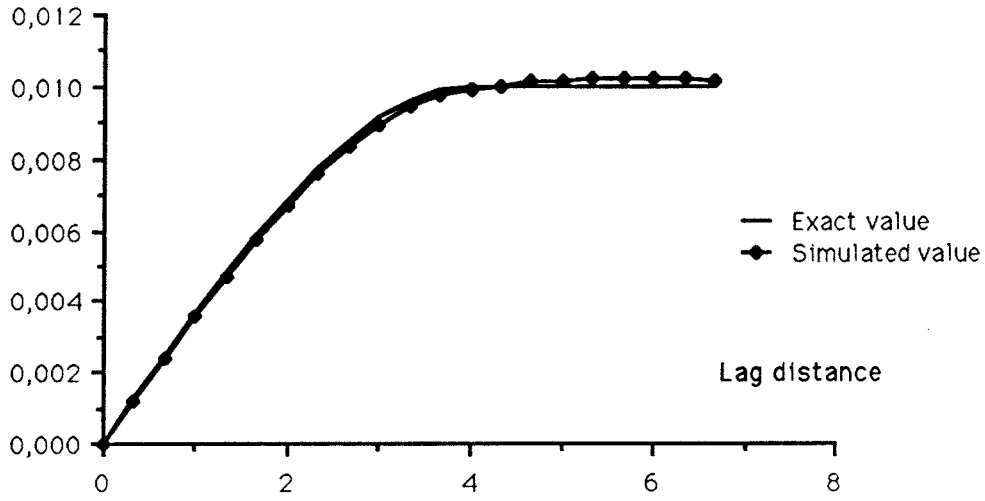


Fig. 6.2.7. Three dimensional realizations in the spherical case using four icosahedron sets. Exact and simulated values of the semivariogram for the conductivity in the 2- direction.

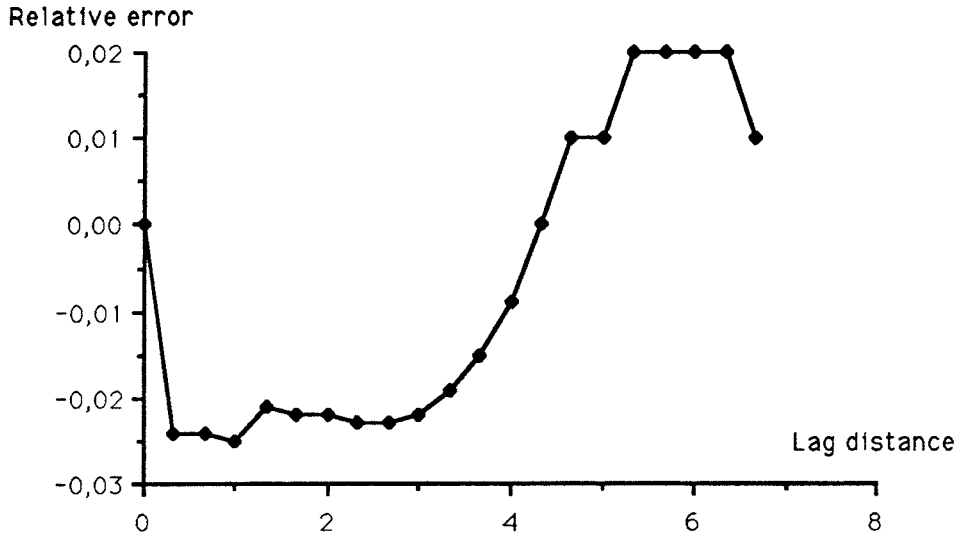


Fig. 6.2.8. Three dimensional realizations in the spherical case using four icosahedron sets. Relative errors of the semivariogram for the conductivity in the 2- direction.

Covariance model:	Exponential
Range parameter, a (see section 2.2.2.):	1. 0
Bandwidth factor, f:	0. 2
Bandwidth, T:	0. 01
R (see section 2.2.):	20
Number of icosahedron sets (see section 2.3.):	4
Number of random lines (see section 2.3.):	0
Spatial sample size, |P| (see section 4.):	121
Spatial sample size, |L| (see section 4.)	1
Number of realizations:	100

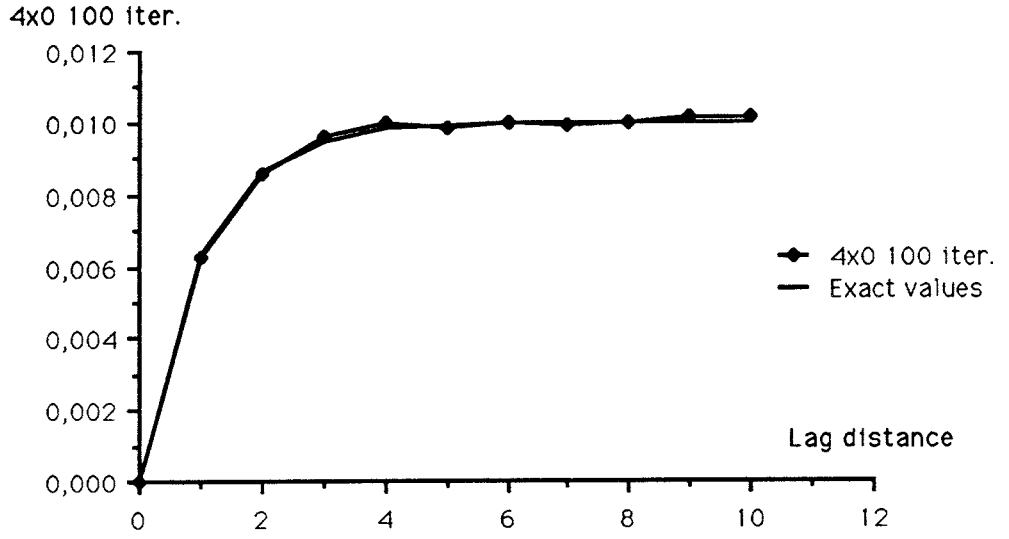


Fig. 6.2.9. Three dimensional realizations in the exponential case using four icosahedron sets. Exact and simulated values of the semivariogram for the conductivity in the 1- direction.

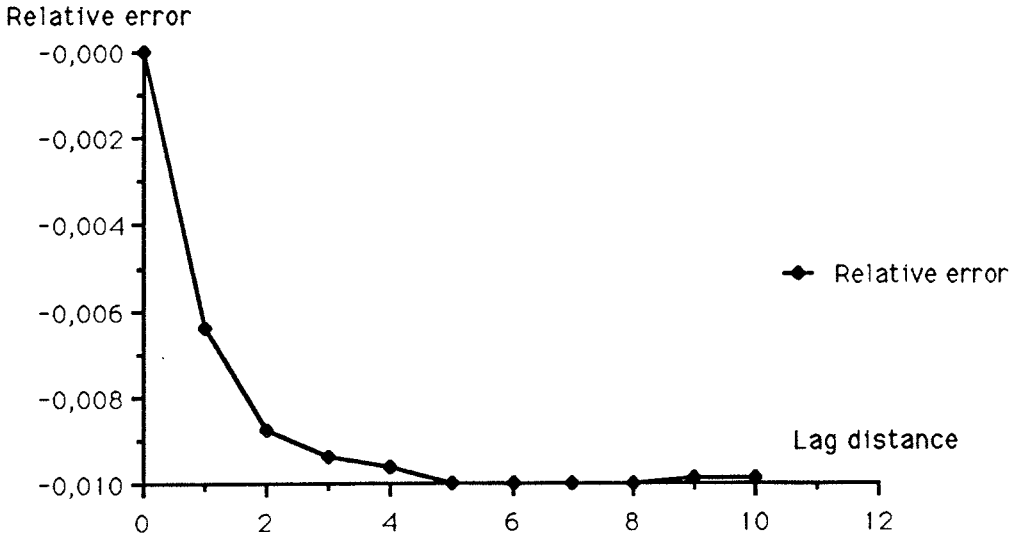


Fig. 6.2.10. Three dimensional realizations in the exponential case using four icosahedron sets. Relative errors of the semivariogram for the conductivity in the 1- direction.

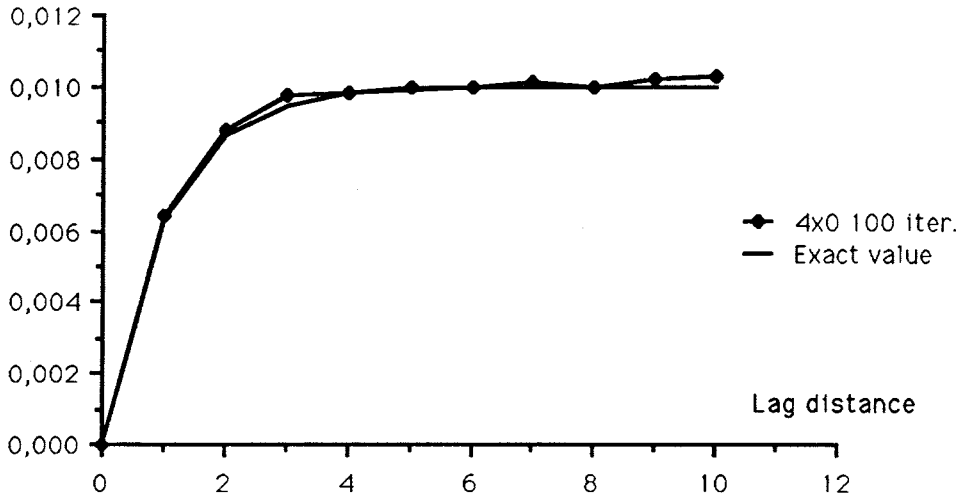


Fig. 6.2.11. Three dimensional realizations in the exponential case.using four icosahedron sets. Exact and simulated values of the semivariogram for the conductivity in the 2- direction.

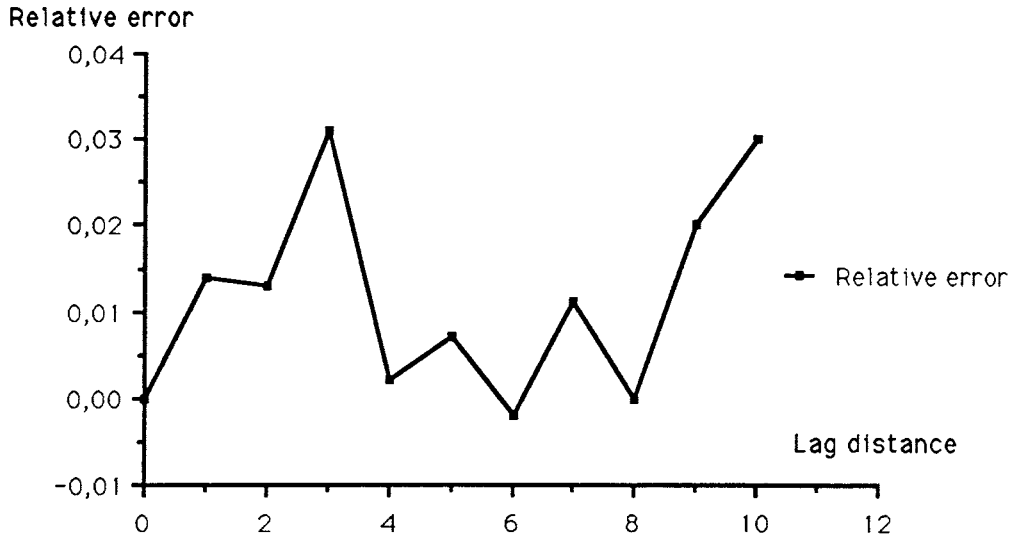


Fig. 6.2.12. Three dimensional realizations in the exponential case.using four icosahedron sets. Relative errors of the semivariogram for the conductivity in the 2- direction.

Covariance model:	Spherical
Variance	0. 01
Range parameter, a (see section 2.2.2.):	1. 0
Bandwidth factor, f:	0. 2
Bandwidth, T:	0. 066
R (see section 2.2.):	60
Number of icosahedron sets (see section 2.3.) :	1
Number of random lines (see section 2.3.) :	85
Number of samples	441

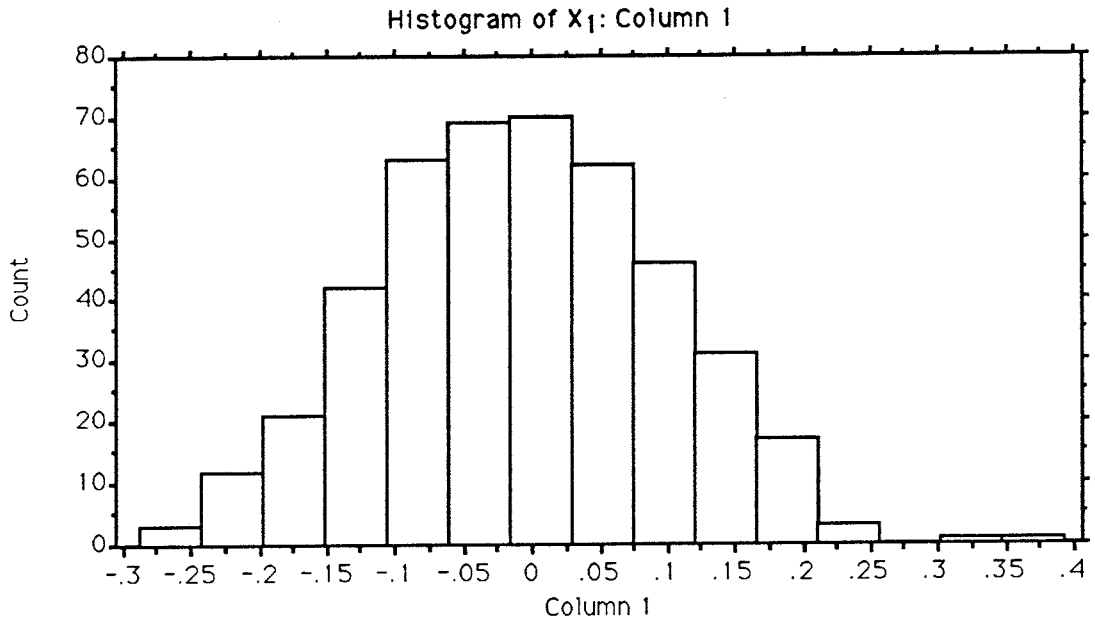


Fig. 6.2.13. Showing a histogram generated from actual simulations.of conductivity values.

6.3. Comparison with the perturbation solution

Approximating the whole three dimensional space with the following finite domain:

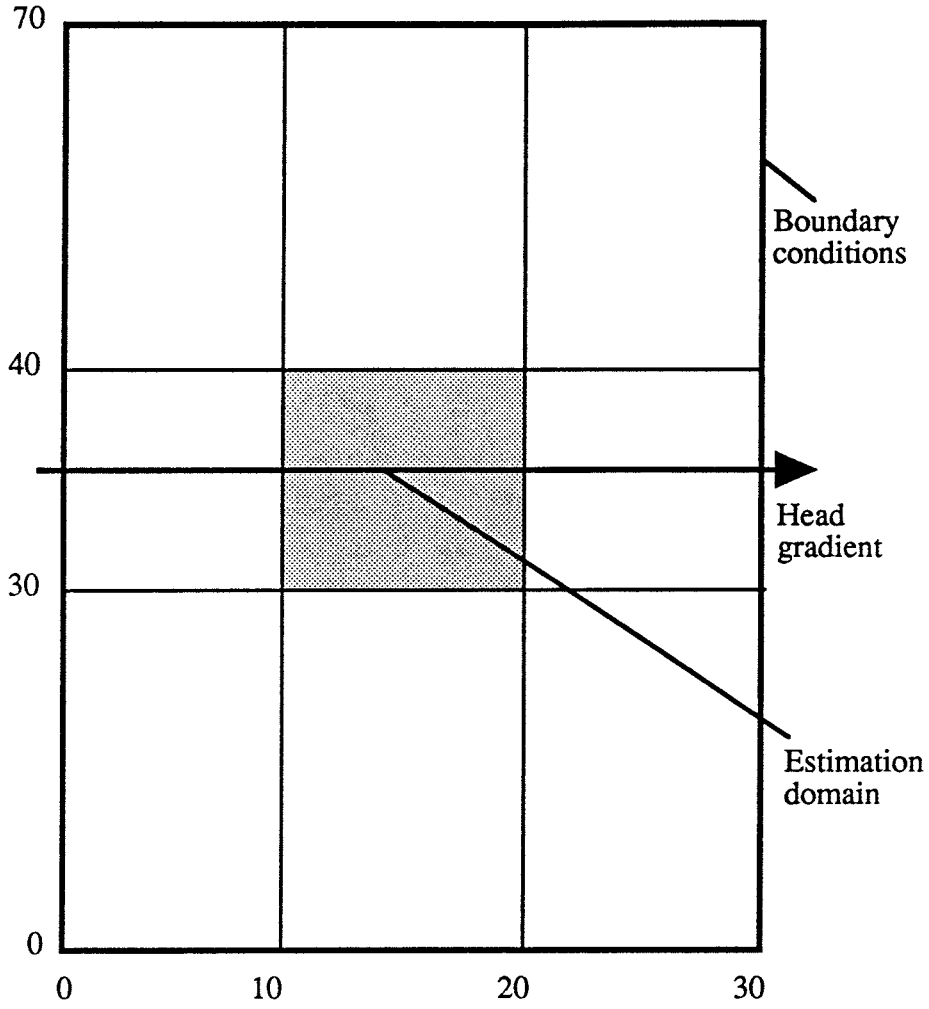


Fig. 6.3.1. Showing a cross section of the three dimensional computational domain used for comparison with perturbation solution.

All the points used in semivariogram estimation, see 4.3, is confined to the so called estimation domain which is located roughly one correlation scale away from the boundary. This is inferred from a look at Fig. 5.1.

On the boundary the head is fixed according to

$$h(x) = g \cdot x + c.$$

A large number of realizations were made on this configuration and slightly modified ones. A typical example is shown in diagrams 6.3.2 - 6.3.5 The prismatic mesh used had the basis vectors $s_j = 1, j = 1, 2, 3$.

The diagrams 6.3.2 - 6.3.3 show the exact semivariograms and the Chebyshev confidence intervals on the 0. 95 -level for the heads in the direction of the gradient and orthogonal to it. Note that the confidence intervals are entirely based on treating the statistic corresponding to 4.3 as a sequence of nonbiased independent spatial averages. Diagrams 6.3.4 - 6.3.5 show the same thing for the conductivities. These are reproduced

here because of the Chebyshev confidence intervals which are absent in Figs. 6.2.1 - 6.2.12 and also because of the large number of realizations in these computations.

As the reader easily sees the results for the head semivariograms embodied in Figs. 6.3.2 and 6.3.3 are not quite satisfactory. Instead we note the following flaws:

1. Too high semivariogram values for the short lags in the direction of the applied gradient.
2. Too high semivariogram values for lags orthogonal to the direction of the applied head gradient.
3. Too low semivariogram values for large lags in the direction of the applied head gradient.
4. The convergence is bad for large lags in the direction of the applied head gradient.

Discussing these points one at a time:

Point 1 is probably due to a too coarse mesh. The mesh used (31x71x71=138069 head nodes) contains three nodes per effective correlation length for the conductivity. Experiments have been performed with six nodes per effective correlation length for the conductivity in the direction of the head gradient and this produces a marked improvement for the short lags in the direction of the applied gradient.

Point 2 is not affected by the refinement mentioned above. The obvious test is to refine the mesh in the directions orthogonal to the applied gradient directions as well. This has not been done primarily because the total number of nodes increases as the square of the number of nodes in the directions orthogonal to the head gradient.

Point 3 is probably due to the finite approximation of the whole space. Tests that have been performed increasing the computational domain seem to validate this.

There is not much to do about point 4.

Thus we see that in the search of better results we are led to put larger and larger demands on our computer resources by increasing the mesh density and the size of the computational domain. The above example is, however, already on the computational limit.

Covariance model:	Exponential
Range parameter, a (see section 2.2.2.):	1
Bandwidth factor, f:	0. 2
Bandwidth, T:	0. 2
R (see section 2.2.):	20
Number of icosahedron sets (see section 2.3.) :	4
Number of random lines (see section 2.3.) :	0
Spatial sample size(for head), I P I (see chapter 4.):	121
Spatial sample size(for head), I L I (see chapter 4.):	1
Number of realizations:	400

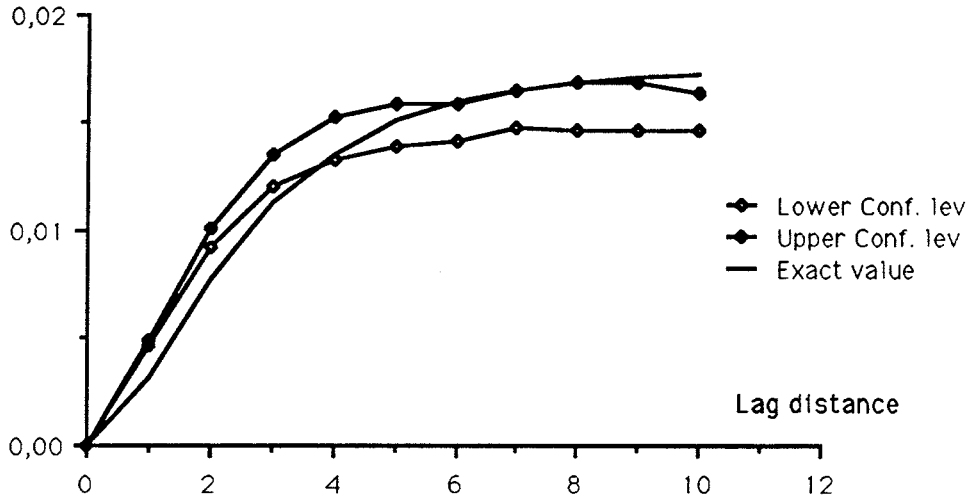


Fig. 6.3.2. Exact semivariogram and the Chebyshev confidence intervals on the 0. 95 -level for the heads in the direction parallel to the applied head gradient.

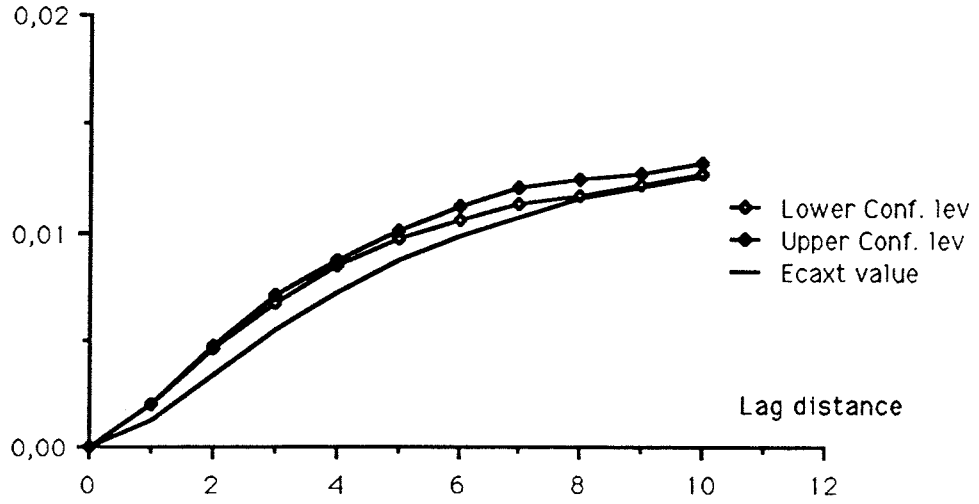


Fig. 6.3.3. Exact semivariogram and the Chebyshev confidence intervals on the 0. 95 -level for the heads in the direction orthogonal to the applied head gradient.

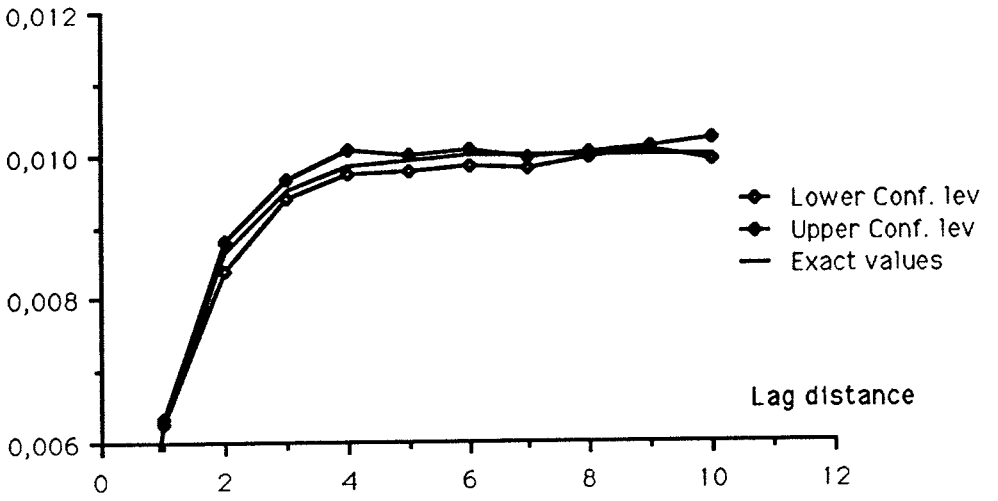


Fig. 6.3.4. Exact semivariogram and the Chebyshev confidence intervals on the 0. 95 -level for the conductivities in the direction of the gradient. The diagram shows only the upper part since all curves coincide on the lower.

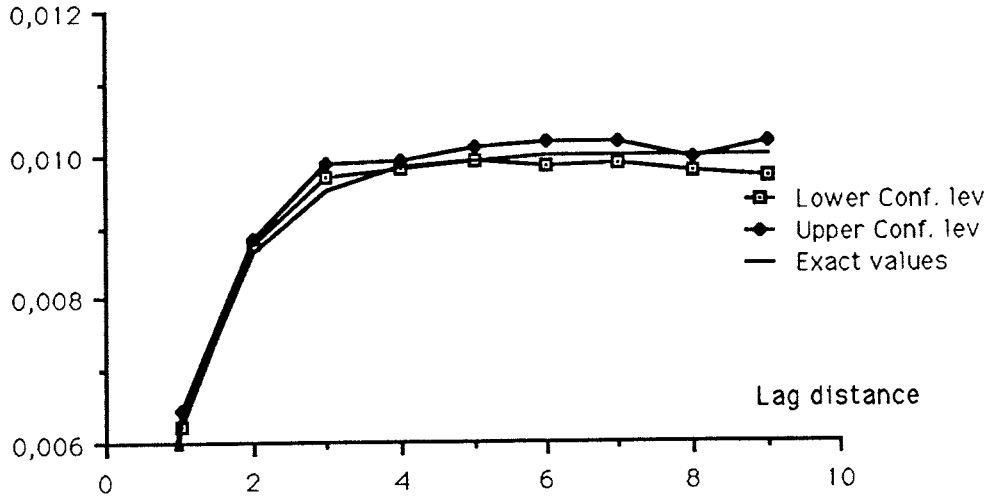


Fig. 6.3.5. Exact semivariogram and the Chebyshev confidence intervals on the 0. 95 -level for the conductivities in the direction orthogonal to gradient. The diagram shows only the upper part since all curves coincide on the lower.

7. RESULTS OF SIMULATION, COMPARISON WITH A HYDROCOIN EXAMPLE

[Hydrocoin, 1988] was an international effort for comparison and verification of different groundwater computer codes. This chapter describes how a verification of the finite difference solver, described in chapter 3, and the particle tracking algorithm in HYDRASTAR was performed joining that comparison. The two dimensional model in [Hydrocoin, 1988], level 1, case 2 as shown in Fig. 7.1 is a two dimensional rock mass intersected by two narrow and highly conductive fracture zones. This is not the kind of problem for which HYDRASTAR was designed, however, it is possible to use HYDRASTAR anyway and to represent the fracture zones as deterministic trends.

To avoid the use of excessively many nodes the zones must be taken into account indirectly. Here is described the methodology used to include small scale structures such as the highly conductive fracture zones.

7.1. Modelling of conductivity

The numerical method described in chapter 3 requires the approximative evaluation of

$$\int_S K(\xi)\nabla\langle h\rangle^f(\xi)dS$$

where S denotes one face of a mass balance element. Ordinarily then this is approximated as

$$\frac{|S|}{\delta_{AB}}K(x_{AB})\left[\langle h\rangle^f(x_A)-\langle h\rangle^f(x_B)\right]$$

where

$\|S\|$	is the area of a face of the mass balance element,
δ_{AB}	is the distance between the neighboring nodes A and B,
$K(x_{AB})$	is the conductivity of the face S of the mass balance element and
$\langle h\rangle^f(x_A)$, $\langle h\rangle^f(x_B)$	are the nodal values in node A and B respectively of the hydraulic head.

This approximation leads directly to the equations 3.3 to 3.5.

If a zone cuts through the face the natural analog of the above is

$$\frac{|S|}{\delta_{AB}}K_{eff}(x_{AB})\left[\langle h\rangle^f(x_A)-\langle h\rangle^f(x_B)\right]$$

with the effective conductivity calculated from the assumption that the gradient is constant over the interface S as

$$K_{eff}(x_{AB})=\frac{|S_{zone}|K_{zone}+|S-S_{zone}|K_{rock}}{|S|}.$$

Here

$\|S_{zone}\|$	is the area of the intersection of the fracture zone and the face of the mass balance element,

$\|S-S_{zone}\|$	is the remaining area of the face of the mass balance element,
K_{rock}	is the conductivity of the rockmass
K_{zone}	is the conductivity of the fracture zones.

7.2. Boundary conditions

The primary input data are, referring to Fig. 7.1,

Maximal height of the domain	1100	m
Minimal height of the domain	1000	m
Width of the domain	1600	m
Conductivity for the fracture zones	10^{-4}	m/s
Conductivity for the rock mass	10^{-8}	m/s,

for full details on the problem the reader is referred to [Hydrocoin, 1988].

The boundary conditions are of the noflow type on all side boundaries and with hydrostatic pressure i.e.

$$(h)^s = z \qquad 7.2.1$$

on the top boundary. Since HYDRASTAR is restricted to a prismatic mesh and a computational domain in the form of a three dimensional parallelepiped this is approximated by taking a parallelepiped extending from the bottom of the model to the low points of the upper boundary where the boundary condition 7.2.1 is required to hold. This is equivalent with neglecting the potential drop caused by flow in the upper parts of the model and will manifest itself by slightly too high head values in the results.

7.3. Results

Calculations on two different meshes were made. Mesh 1 consists of 704 nodes, with elemental size 50x51 m, and should be considered coarse according to the definitions in [Hydrocoin, 1988]. Mesh 2 consists of 2880 nodes, elemental size 25x24 m and is classified as a medium mesh. This classification is based on comparison with the other finite difference codes participating in [Hydrocoin, 1988].

In mesh 1 the discretisation caused a smoothing of the top boundary head values (Fig. 7.3.1). The medium mesh reproduces these values better, which gives results with a larger span between the highest and lowest head value for each z, see Figs. 7.3.1 to 7.3.3.

In general the results are in good agreement with the other teams' results, except for a small offset, around +3 m, in the head values. See Figs. 7.3.2 to 7.3.5. The offset is due to our approximation of the top boundary values. This could be overcome by either choosing a more sophisticated way of representing the top boundary or simply using other teams' calculated head values as our boundary values.

In Figs. 7.3.2 and 7.3.3 the head values are presented for a z - level near but not exactly on the -200 m and -800 m level, since it is not possible to find a HYDRASTAR grid that at the same time represents the boundary conditions correctly and has calculation nodes at the -200 m and -800 m levels.

Fig. 7.3.6 shows qualitative representations of the equipotential lines for the two mesh densities.

In Fig. 7.3.7 the stream lines from 4 tracers are illustrated. The ability to resolve the fracture zones depends on the size of the mesh as can be seen from the figure. However our results are consistent with the ones given in [Hydrocoin, 1988] in that trajectory 2 i.e. the trajectory started at the lowest leftmost position exits through the rightmost fracture zone, see Fig. 7.3.9.

Finally Fig. 7.3.8 shows the accumulated travel distance as a function of time for trajectory 2. These graphs are in good agreement with [Hydrocoin, 1988], see Fig. 7.3.9.

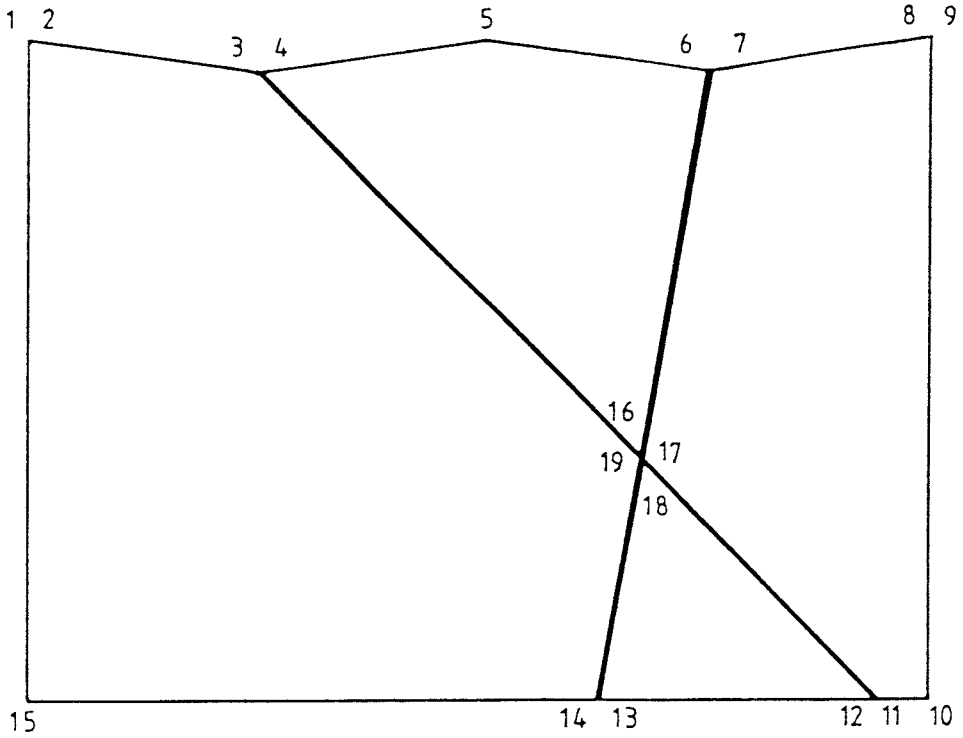
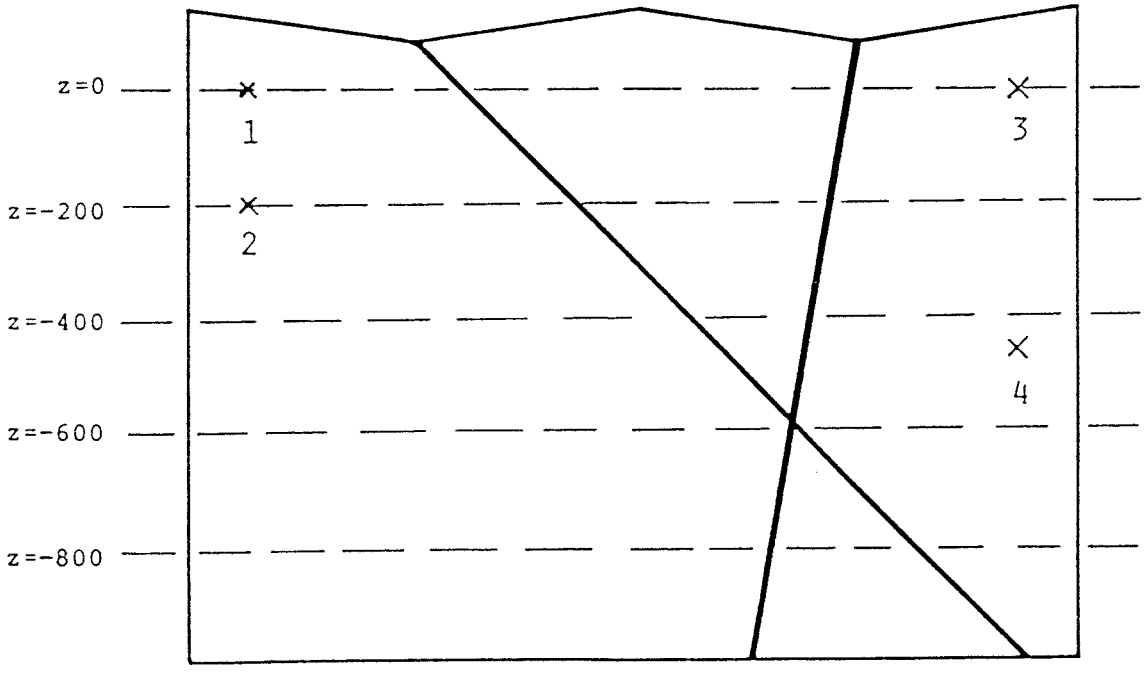


Fig. 7.1. Geometry of the modelled domain from [Hydrocoin, 1988].



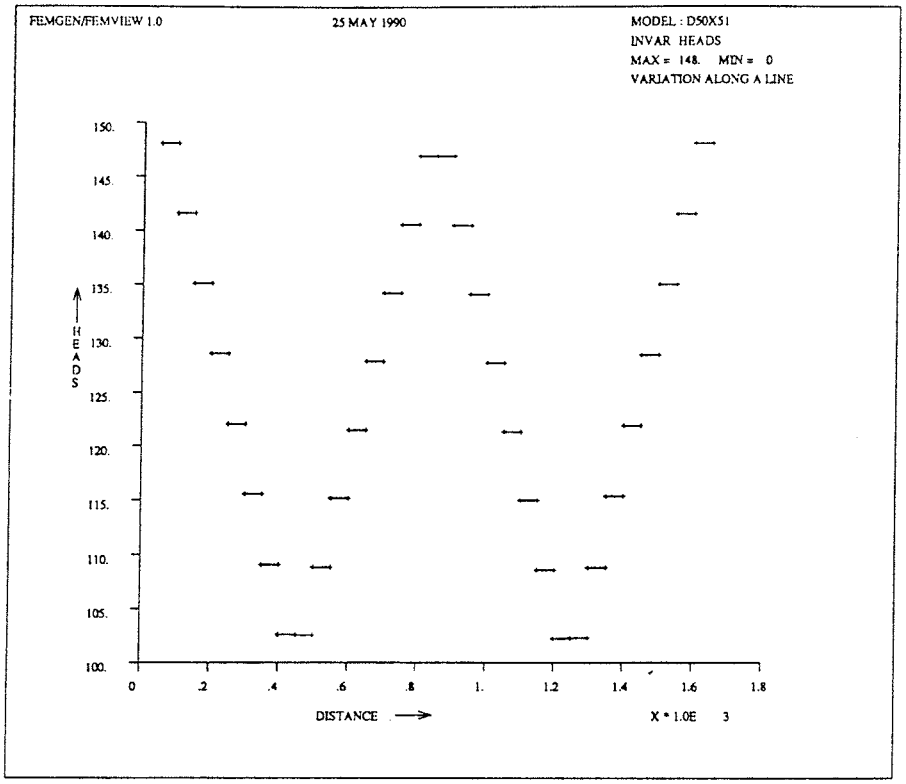
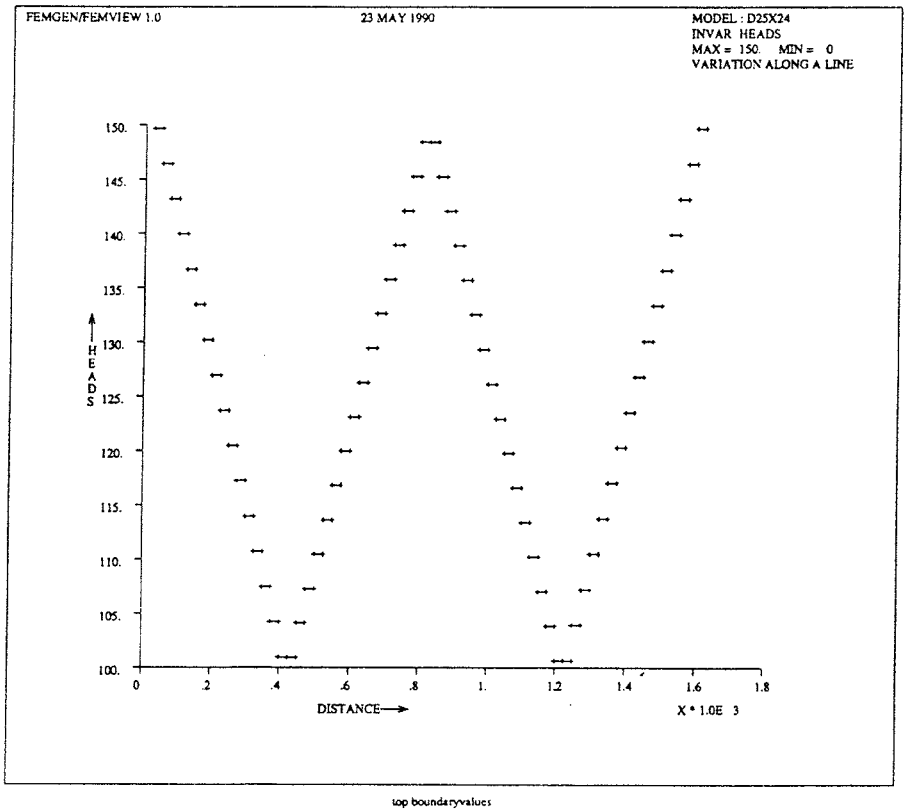


Fig. 7.3.1. Top boundary values from HYDRASTAR for mesh 1 (above) and mesh 2 (below).



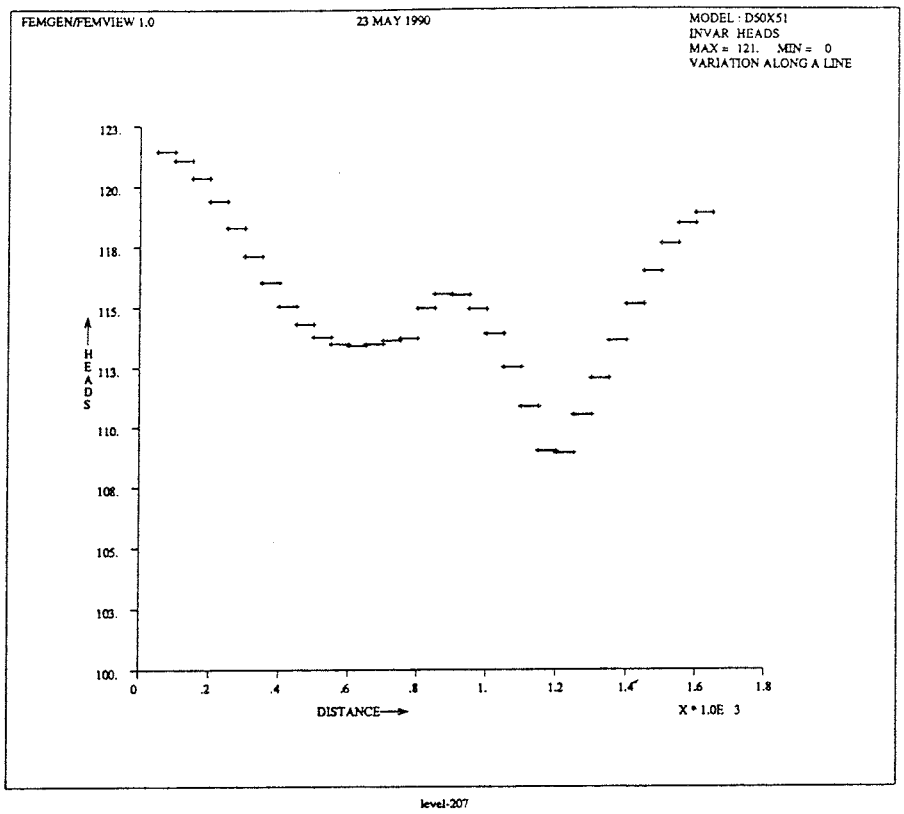
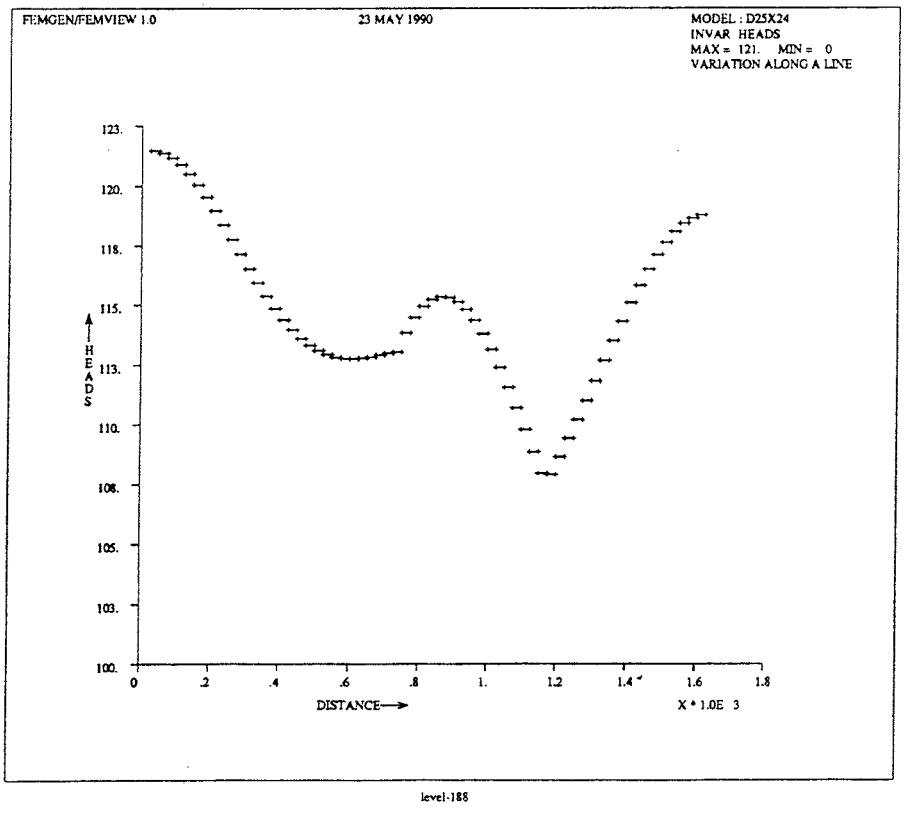


Fig. 7.3.2. Head values from HYDRASTAR at level -207 m for mesh 1 (above) and at level -188 m for mesh 2 (below).



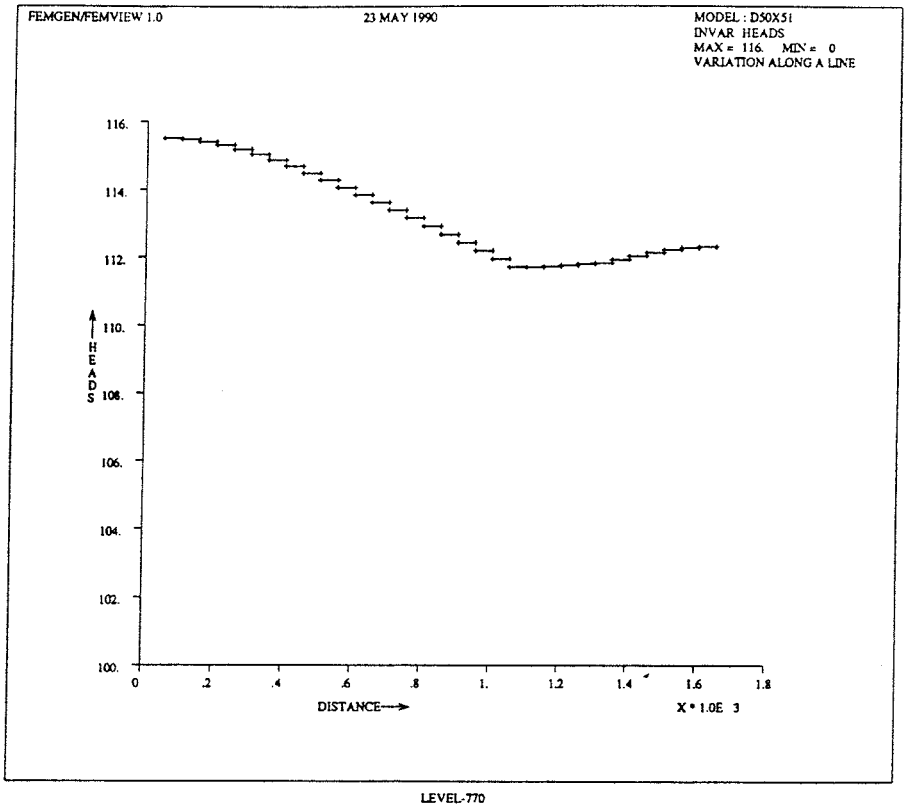
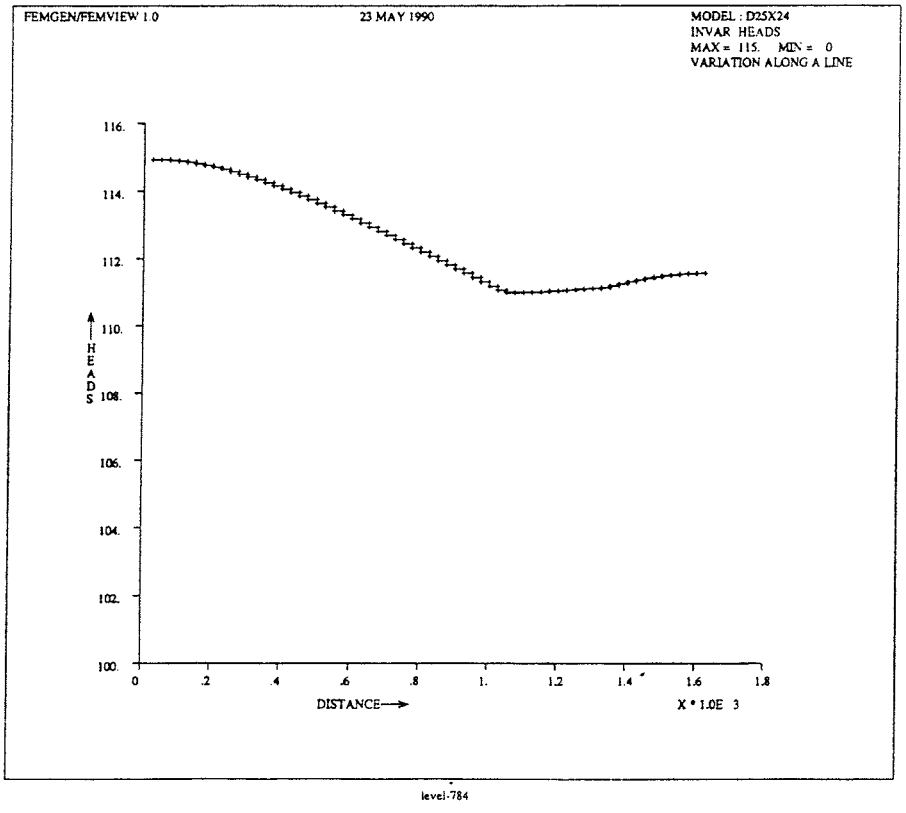


Fig. 7.3.3. Head values from HYDRASTAR at level -821 m for mesh 1 (above) and at level -808 m for mesh 2 (below).



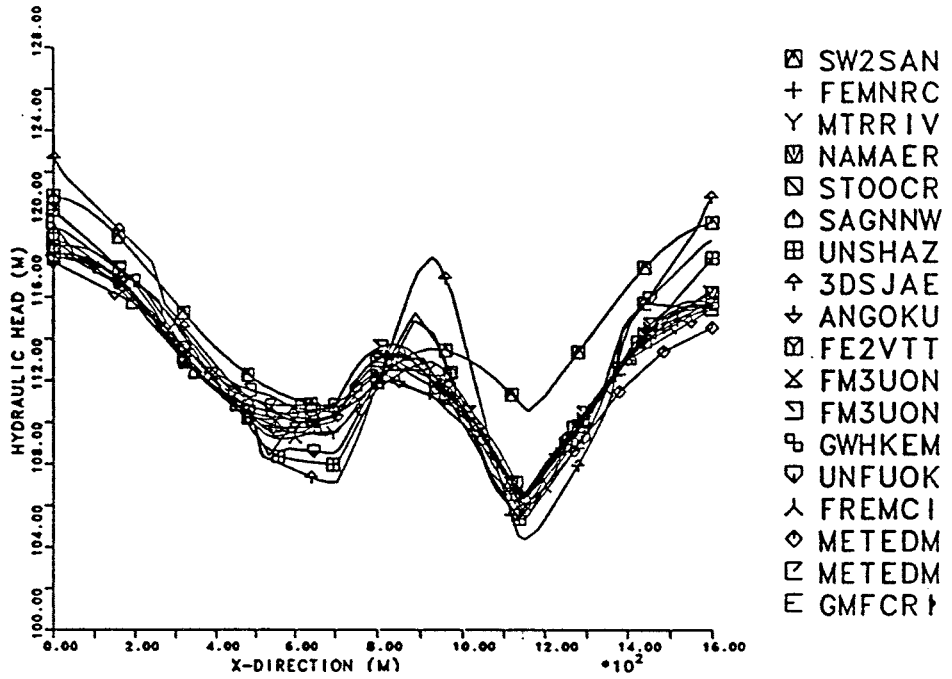
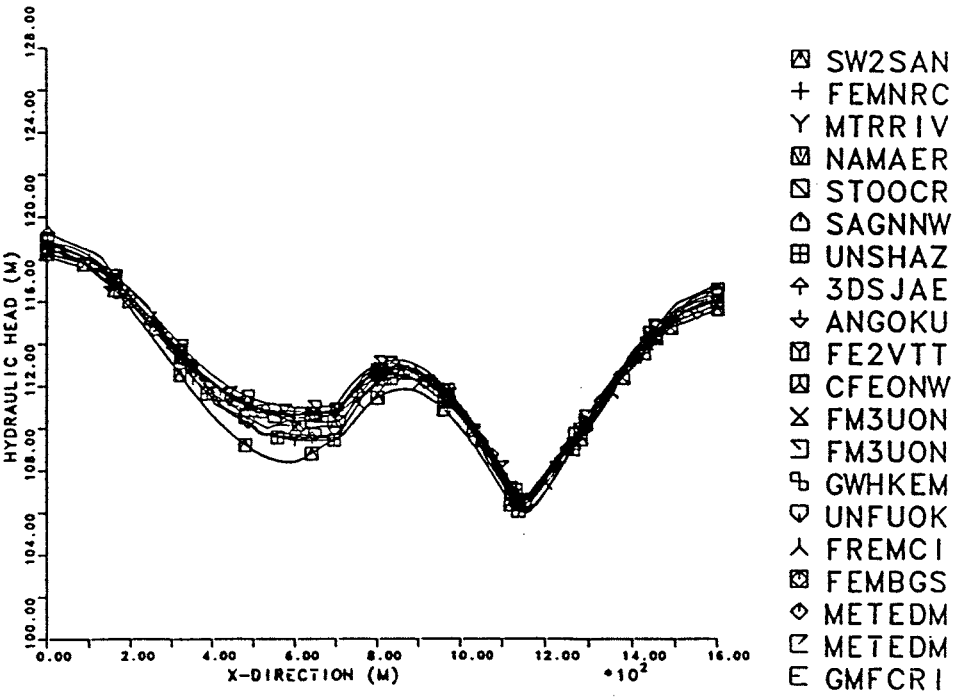


Fig. 7.3.4. Head values from [Hydrocoin, 1988] at a level of -200 m with coarse meshes (above) and medium meshes (below).



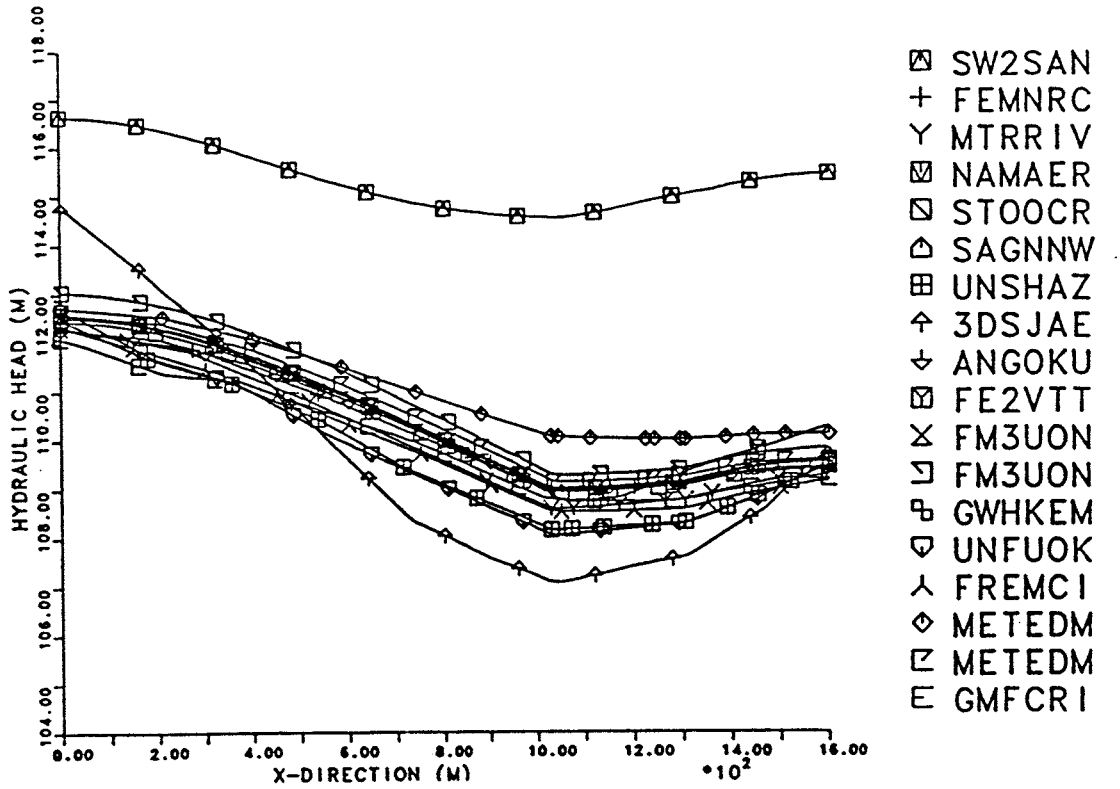
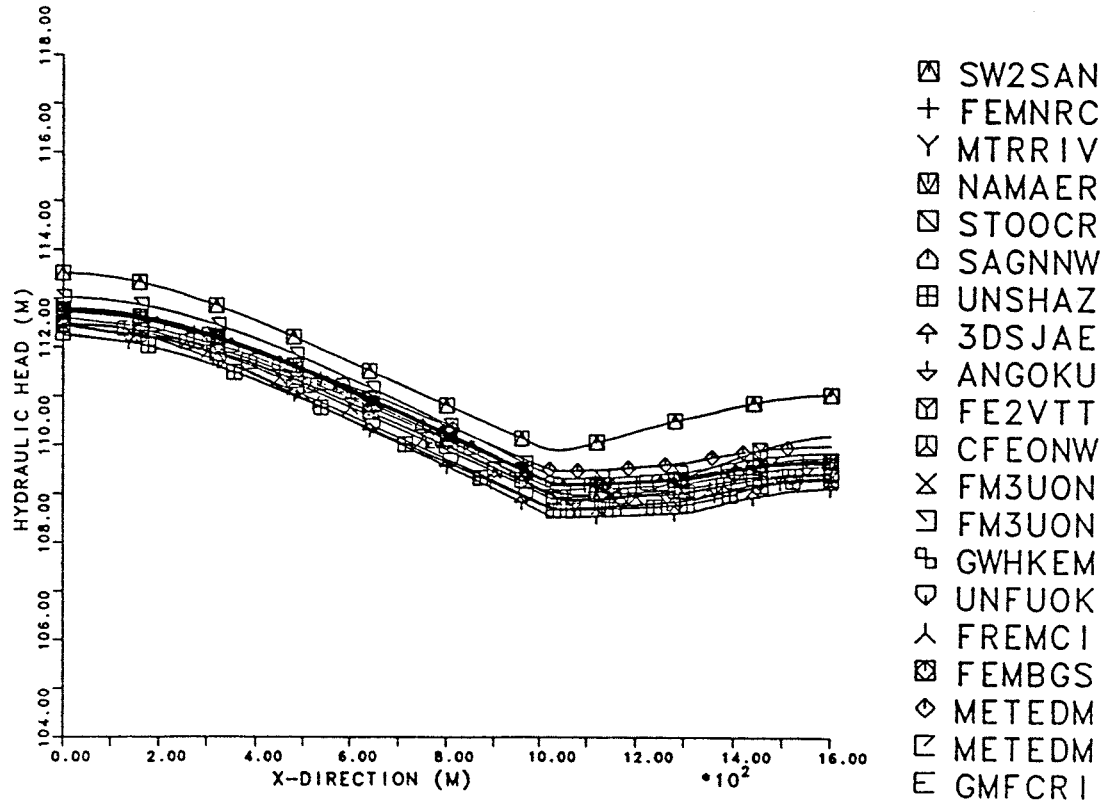


Fig. 7.3.5. Head values from [Hydrocoin, 1988] at a level of -800 m with coarse meshes (above) and medium meshes (below).



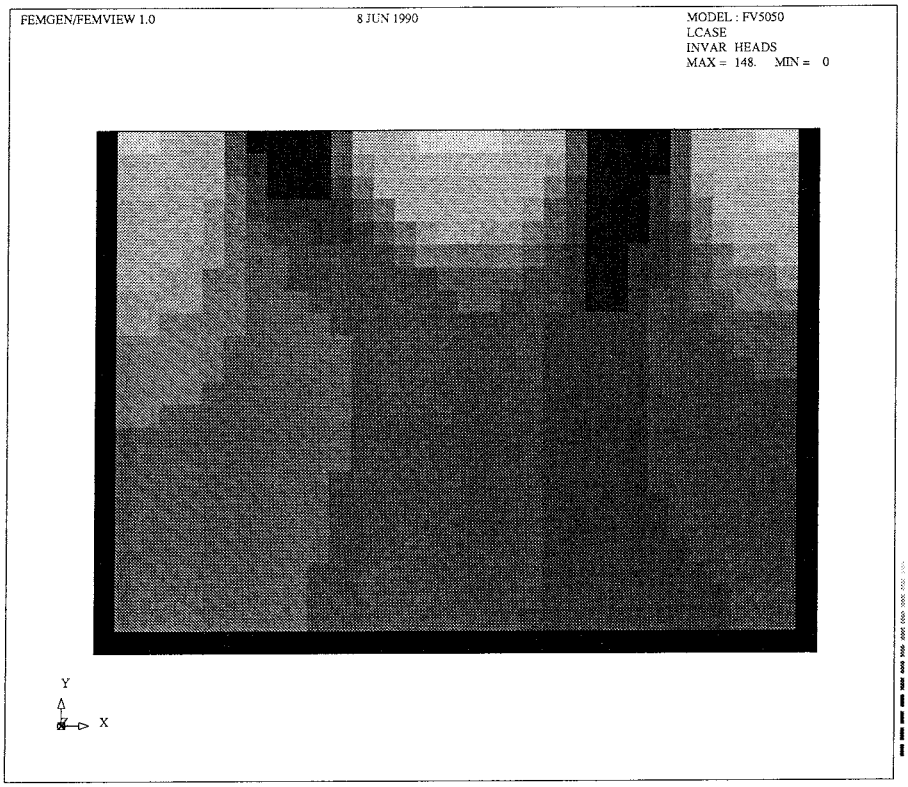


Fig. 7.3.6. Qualitative representation of the equipotential lines from HYDRASTAR and for mesh 1 (above) and mesh2 (below).

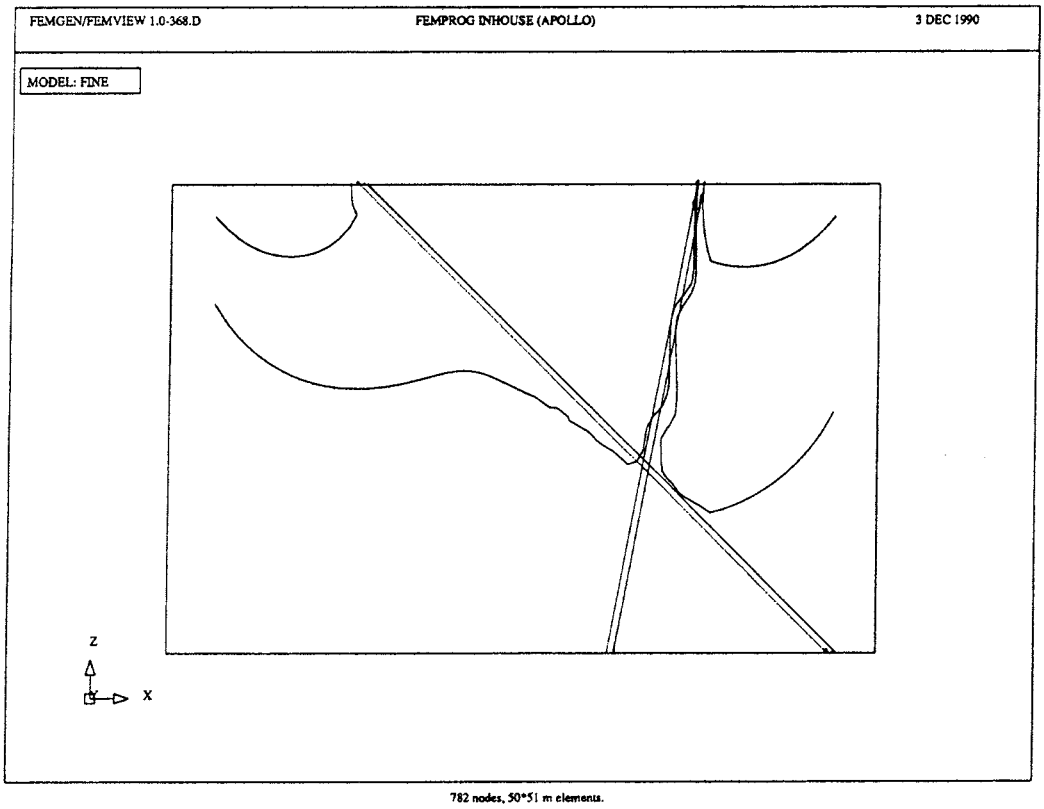
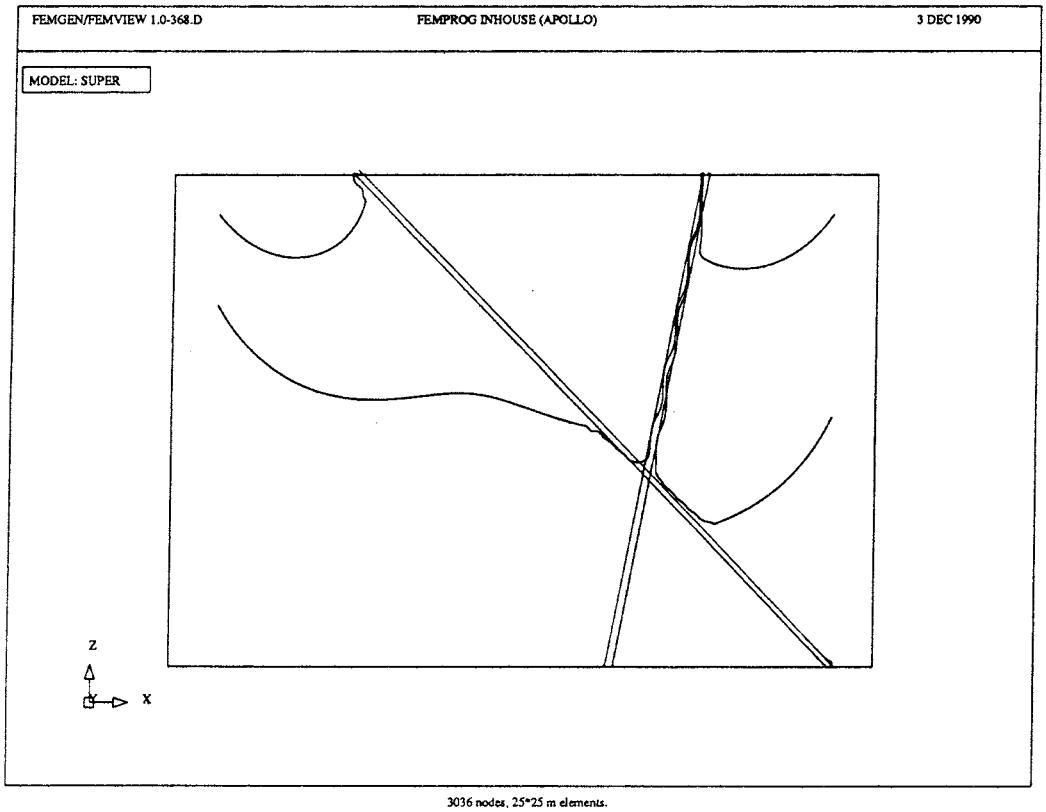


Fig. 7.3.7. Stream lines from HYDRASTAR for mesh1 (above) and mesh2 (below).



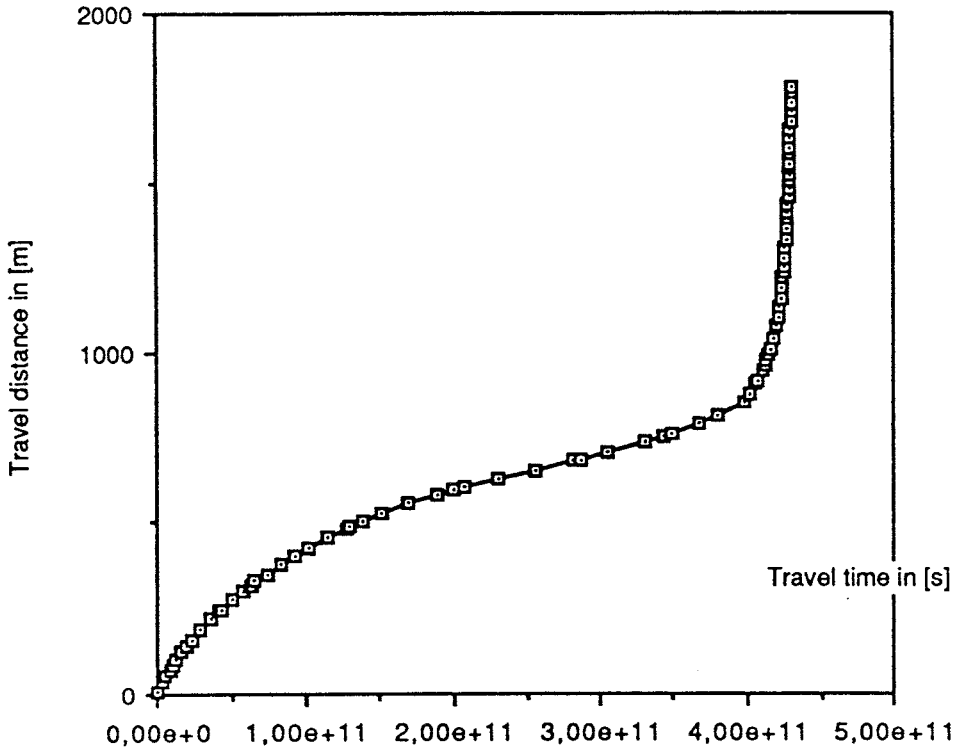
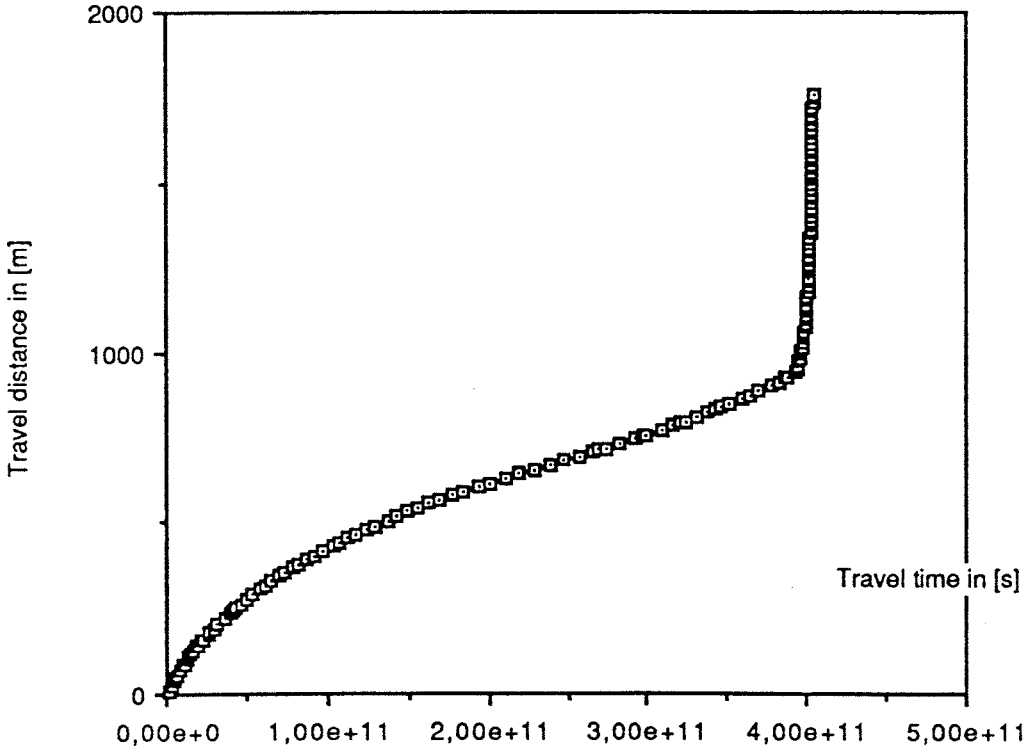


Fig. 7.3.8. Travel distance versus time from HYDRASTAR and for trajectory 2 for mesh1 (above) and mesh2 (below).



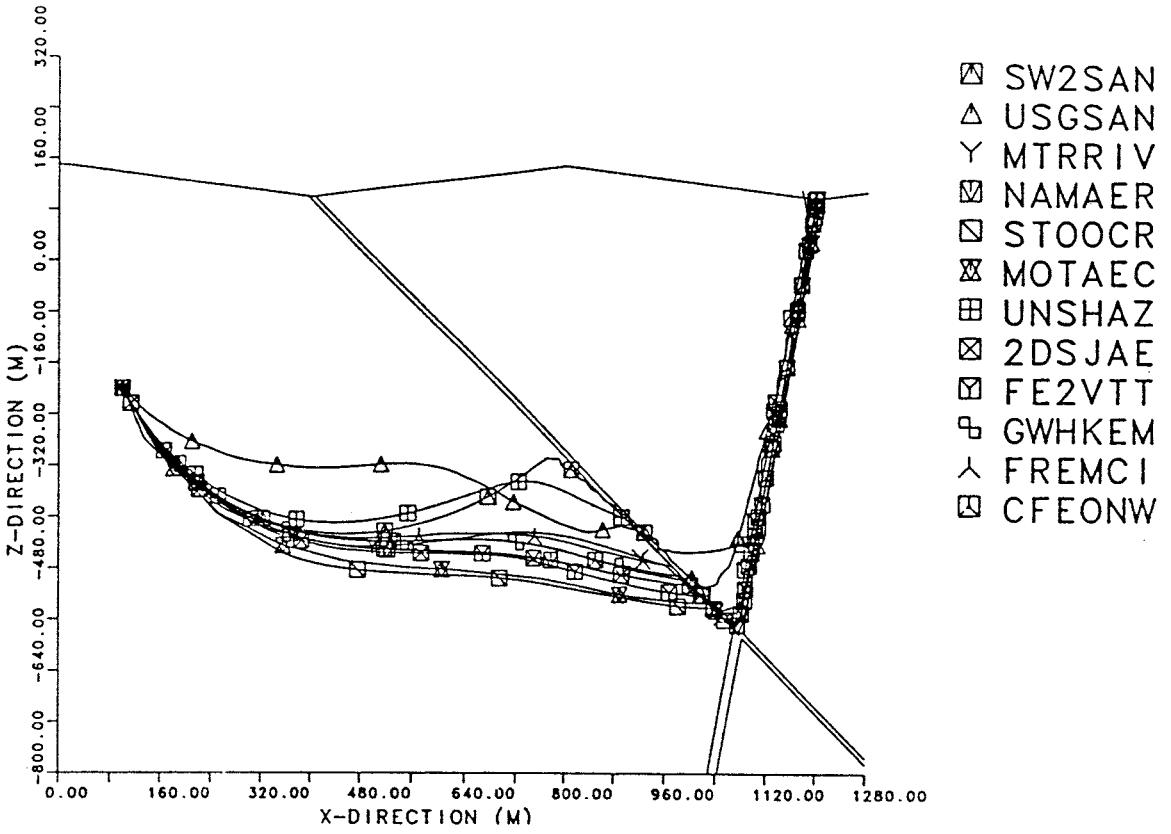
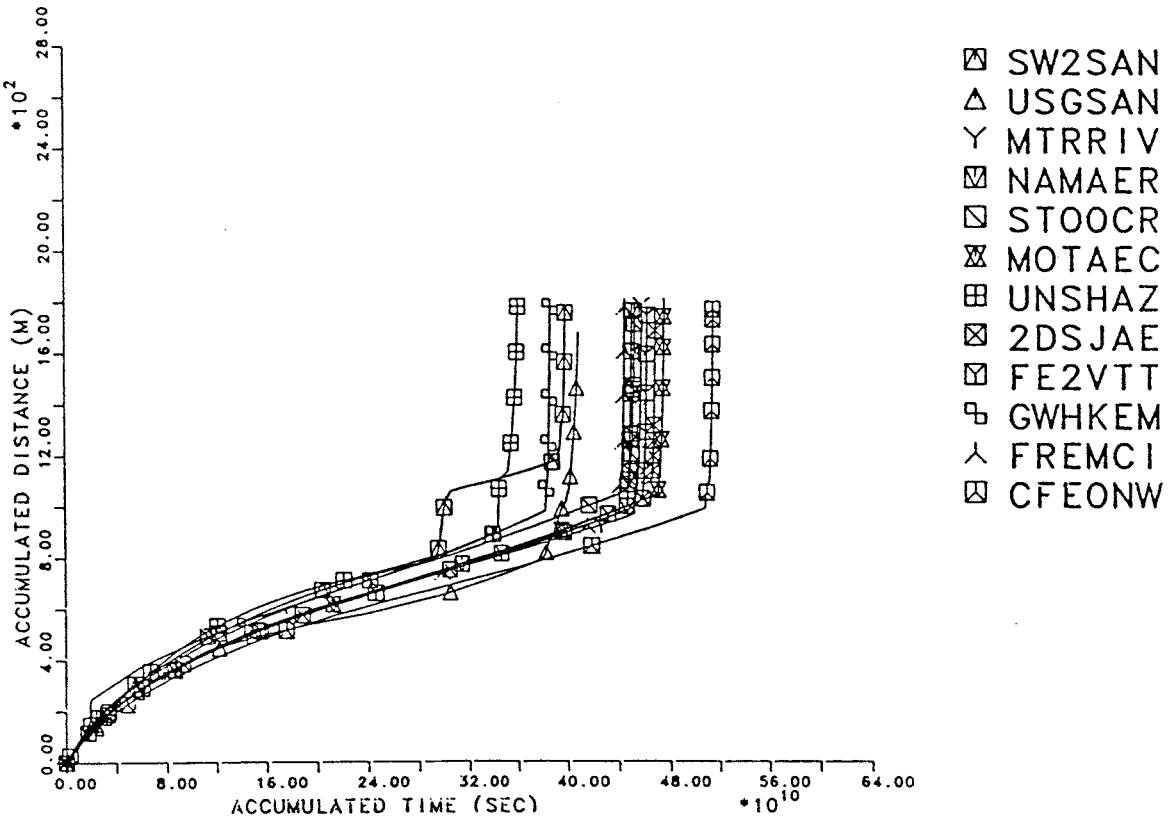


Fig. 7.3.9. Stream lines (above) and travel distance versus time (below) for trajectory 2 from [Hydrocoin, 1988].



8. ACKNOWLEDGEMENTS

The author wishes to thank

Mr Per Hörnfeldt for his devoted work with the HYDROCOIN comparison including coding and result generation. He also wrote parts of the corresponding chapter,

Mr Lars Lovius for his contribution to the development of the stream lines / particle tracking routines and the FEMVIEW interface,

Mr. Nils Kjellbert and Mr. Anders Ström for their constant support and encouragement during the project, they also proof-read the entire manuscript,

Mr Daniel Sundström for his assistance in planning the project,

Dr. Jesper Oppelstrup for his invaluable advice.

9. REFERENCES

[Bakr et al, 1978] Bakr, A. A. , Gelhar, L. W. Gutjahr, A. L. and MacMillan, J. R. , Stochastic Analysis of Spatial Variability in Subsurface Flow 1. Comparison of One- and Three-Dimensional Flows, Water Resources Research, Vol. 14, No 2, pp 263 -271, 1978.

[Beran, 1968] Beran, M. J. , Statistical Continuum Theories, Wiley Interscience, New York, 1968.

[Brately, Fox and Schrage, 1987] Bratley, P. , Fox, B. L. , Schrage, L. E. , A Guide to Simulation, 2nd Ed. , Springer, New York, 1987.

[Coveyou and MacPherson, 1967] Coveyou, R. L. , MacPherson, R. D. , Fourier Analysis of Uniform Random Number Generators, J. of the Assoc. for Computing Machinery, Vol. 14, No 1, Jan. 1967, pp 100 -116.

[Dagan, 1985] Dagan, G. Stochastic Modelling of Groundwater Flow by Unconditional and Conditional Probabilities: The inverse problem, Water resources Research, Vol. 21, No 1 pp 65 -72, 1985.

[Doob, 1953] Doob, J. L. , Stochastic processes, John Wiley and Sons, Inc. New York 1953.

[Fisz, 1963] Fisz, M. , Probability Theory and Mathematical Statistics, Wiley 1963.

[Gelhar and Axness, 1983] Gelhar, L. W. and Axness C. L. , Three-dimensional Stochastic Analysis of Macrodispersion in Aquifers, Water Resources Research, Vol. 19. No. 1, pp 161-180, 1983

[Golub and van Loan] Golub, G. van Loan, Matrix Computations, 2:d edition, Hopkins, July 1989.

[Hydrocoin, 1988] The international Hydrocoin project: Groundwater hydrology modelling strategies for performance assessment of nuclear waste disposal. , Level 1: Code verification. OECD, 1988.

[Journel and Huijbregts, 1978] Journel, A. Huijbregts, Ch. J. , Mining Geostatistics, Academic Press, 1989.

[Lovius et al, 1990] Lovius, L. , Norman, S. , Kjellbert, N. , Uncertainties in repository performance from spatial variability of hydraulic conductivities - statistical estimation and stochastic simulation using PROPER, SKB-TR, 90-03, Feb. 1990.

[Norman, 1991] SKB-TR, in progress.

[de Marsily, 1986] de Marsily, G. , Quantitative Hydrogeology, Academic Press, 1986.

[Matheron, 1973] Matheron, G. , The intrinsic random functions and their applications. , Adv. Appl. Prob. 5, 439 - 468.

[McGrath and Irving, 1975] McGrath, E. J. , Irving D. C. , Techniques for Efficient Monte Carlo Simulation ORNL-RSIC-38, Vol. 2, Oak Ridge National Laboratory,.1975

[Neuman et al, 1987] Neuman, S. P. , Winter, C. L. , Newman, C. M. , Stochastic Theory of Field Scale Fickian Dispersion in Anisotropic Porous Media, Water Resources research, Vol 23, No 3. pp 453-466, 1987.

[Neuman, 1988] Neuman S. P. , A proposed conceptual framework and methods for investigating flow and transport in Swedish crystalline rock, SKB AR 88-37, September 1988.

[Pörn, 1986] Pörn, K. , Testing of the Random Number Generator used in PROPER, SKB Memo 1986.

[Tompson, Ababou and Gelhar, 1989] Tompson, A. F. B. , Ababou, R. , Gelhar, L. W. , Implementation of the Three Dimensional Turning Bands Generator, Water Resour. Res. , Vol. 25, No. 10, pp 2227-2234, Oct. 1989.

[Yaglom, 1962] Yaglom, A. M. , Stationary random functions, Prentice-Hall, 1962. (Dover Publications 1973)

List of SKB reports

Annual Reports

1977-78
TR 121
KBS Technical Reports 1 – 120
Summaries
Stockholm, May 1979

1979
TR 79-28
The KBS Annual Report 1979
KBS Technical Reports 79-01 – 79-27
Summaries
Stockholm, March 1980

1980
TR 80-26
The KBS Annual Report 1980
KBS Technical Reports 80-01 – 80-25
Summaries
Stockholm, March 1981

1981
TR 81-17
The KBS Annual Report 1981
KBS Technical Reports 81-01 – 81-16
Summaries
Stockholm, April 1982

1982
TR 82-28
The KBS Annual Report 1982
KBS Technical Reports 82-01 – 82-27
Summaries
Stockholm, July 1983

1983
TR 83-77
The KBS Annual Report 1983
KBS Technical Reports 83-01 – 83-76
Summaries
Stockholm, June 1984

1984
TR 85-01
Annual Research and Development Report 1984
Including Summaries of Technical Reports Issued during 1984. (Technical Reports 84-01 – 84-19)
Stockholm, June 1985

1985
TR 85-20
Annual Research and Development Report 1985
Including Summaries of Technical Reports Issued during 1985. (Technical Reports 85-01 – 85-19)
Stockholm, May 1986

1986
TR 86-31
SKB Annual Report 1986
Including Summaries of Technical Reports Issued during 1986
Stockholm, May 1987

1987
TR 87-33
SKB Annual Report 1987
Including Summaries of Technical Reports Issued during 1987
Stockholm, May 1988

1988
TR 88-32
SKB Annual Report 1988
Including Summaries of Technical Reports Issued during 1988
Stockholm, May 1989

1989
TR 89-40
SKB Annual Report 1989
Including Summaries of Technical Reports Issued during 1989
Stockholm, May 1990

Technical Reports
List of SKB Technical Reports 1991

TR 91-01
Description of geological data in SKB's database GEOTAB
Version 2
Stefan Sehlstedt, Tomas Stark
SGAB, Luleå
January 1991

TR 91-02
Description of geophysical data in SKB database GEOTAB
Version 2
Stefan Sehlstedt
SGAB, Luleå
January 1991

TR 91-03
1. The application of PIE techniques to the study of the corrosion of spent oxide fuel in deep-rock ground waters
2. Spent fuel degradation
R S Forsyth
Studsvik Nuclear
January 1991

TR 91-04
Plutonium solubilities
I Puigdomènech[1], J Bruno[2]
[1]Enviromental Services, Studsvik Nuclear, Nyköping, Sweden
[2]MBT Tecnologia Ambiental, CENT, Cerdanyola, Spain
February 1991

TR 91-05
Description of tracer data in the SKB database GEOTAB
SGAB, Luleå
April, 1991

TR 91-06
Description of background data in the SKB database GEOTAB
Version 2
Ebbe Eriksson, Stefan Sehlstedt
SGAB, Luleå
March 1991

TR 91-07
Description of hydrogeological data in the SKB's database GEOTAB
Version 2
Margareta Gerlach[1], Bengt Gentzschein[2]
[1]SGAB, Luleå
[2]SGAB, Uppsala
April 1991

TR 91-08
Overview of geologic and geohydrologic conditions at the Finnsjön site and its surroundings
Kaj Ahlbom[1], Sven Tirén[2]
[1]Conterra AB
[2]Sveriges Geologiska AB
January 1991

TR 91-09
Long term sampling and measuring program. Joint report for 1987, 1988 and 1989. Within the project: Fallout studies in the Gideå and Finnsjö areas after the Chernobyl accident in 1986
Thomas Ittner
SGAB, Uppsala
December 1990

TR 91-10
Sealing of rock joints by induced calcite precipitation. A case study from Bergeforsen hydro power plant
Eva Hakami[1], Anders Ekstav[2], Ulf Qvarfort[2]
[1]Vattenfall HydroPower AB
[2]Golder Geosystem AB
January 1991

TR 91-11
Impact from the disturbed zone on nuclide migration – a radioactive waste repository study
Akke Bengtsson[1], Bertil Grundfelt[1], Anders Markström[1], Anders Rasmuson[2]
[1]KEMAKTA Konsult AB
[2]Chalmers Institute of Technology
January 1991

TR 91-12
Numerical groundwater flow calculations at the Finnsjön site
Björn Lindbom, Anders Boghammar, Hans Lindberg, Jan Bjelkås
KEMAKTA Consultants Co, Stockholm
February 1991

TR 91-13
Discrete fracture modelling of the Finnsjön rock mass
Phase 1 feasibility study
J E Geier, C-L Axelsson
Golder Geosystem AB, Uppsala
March 1991

TR 91-14
Channel widths
Kai Palmqvist, Marianne Lindström
BERGAB-Berggeologiska Undersökningar AB
February 1991

TR 91-15
Uraninite alteration in an oxidizing environment and its relevance to the disposal of spent nuclear fuel
Robert Finch, Rodney Ewing
Department of Geology, University of New Mexico
December 1990

TR 91-16
Porosity, sorption and diffusivity data compiled for the SKB 91 study
Fredrik Brandberg, Kristina Skagius
Kemakta Consultants Co, Stockholm
April 1991

TR 91-17
Seismically deformed sediments in the Lansjärv area, Northern Sweden
Robert Lagerbäck
May 1991

TR 91-18
Numerical inversion of Laplace transforms using integration and convergence acceleration
Sven-Åke Gustafson
Rogaland University, Stavanger, Norway
May 1991

TR 91-19
NEAR21 - A near field radionuclide migration code for use with the PROPER package
Sven Norman[1], Nils Kjellbert[2]
[1]Starprog AB
[2]SKB AB
April 1991

TR 91-20
Äspö Hard Rock Laboratory. Overview of the investigations 1986-1990
R Stanfors, M Erlström, I Markström
June 1991

TR 91-21
Äspö Hard Rock Laboratory. Field investigation methodology and instruments used in the pre-investigation phase, 1986-1990
K-E Almén, O Zellman
June 1991

TR 91-22
Äspö Hard Rock Laboratory. Evaluation and conceptual modelling based on the pre-investigations 1986-1990
P Wikberg, G Gustafson, I Rhén, R Stanfors
June 1991

TR 91-23
Äspö Hard Rock Laboratory. Predictions prior to excavation and the process of their validation
Gunnar Gustafson, Magnus Liedholm, Ingvar Rhén, Roy Stanfors, Peter Wikberg
June 1991

TR 91-24
Hydrogeological conditions in the Finnsjön area. Compilation of data and conceptual model
Jan-Erik Andersson, Rune Nordqvist, Göran Nyberg, John Smellie, Sven Tirén
February 1991

TR 91-25
The role of the disturbed rock zone in radioactive waste repository safety and performance assesment. A topical discussion and international overview.
Anders Winberg
June 1991

TR 91-26
Testing of parameter averaging techniques for far-field migration calculations using FARF31 with varying velocity.
Akke Bengtsson[1], Anders Boghammar[1], Bertil Grundfelt[1], Anders Rasmuson[2]
[1]KEMAKTA Consultants Co
[2]Chalmers Institute of Technology