TECHNICAL REPORT

90-03

Uncertainties in repository performance from spatial variability of hydraulic conductivities – statistical estimation and stochastic simulation using PROPER

Lars Lovius[1], Sven Norman[1], Nils Kjellbert[2]

[1] Starprog AB
[2] SKB AB

February 1990

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

UNCERTAINTIES IN REPOSITORY PERFORMANCE FROM SPATIAL VARIABILITY OF HYDRAULIC CONDUCTIVITIES - STATISTICAL ESTIMATION AND STOCHASTIC SIMULATION USING PROPER

Lars Lovius[1], Sven Norman[1], Nils Kjellbert[2]

1 Starprog AB
2 SKB AB

February 1990

Uncertainties in Repository Performance from Spatial Variability of Hydraulic Conductivities

– Statistical Estimation and Stochastic Simulation Using PROPER.

Lars Lovius
Sven Norman
Starprog AB

Nils Kjellbert
SKB AB

February 1990

ABSTRACT

An assessment has been made of the impact of spatial variability on the performance of a KBS–3 type repository. The uncertainties in geohydrologically related performance measures have been investigated using conductivity data from one of the Swedish study sites. The analysis was carried out with the PROPER code and the FSCF10 submodel.

LIST OF CONTENTS

1 INTRODUCTION

1.1 Background

The purpose of the PROPER code package developed by the SKB is to provide the safety analyst with a computerized methodology that enables him/her to study the propagation of input data uncertainties in performance–assessment–related model calculations.

The PROPER Monitor is used to interconnect the desired submodels, selected from a library at runtime, and propagate the input parameter uncertainties to find the associated uncertainties in the results using Monte Carlo techniques. The final evaluation must be carried out using a PROPER post–processing code.

The Monte Carlo approach requires simple submodels and/or use of very fast numerical algorithms.

The finite element geohydrology code FSCF10 (= Flow of Slightly Compressible Fluids) has been specially designed by Carol Braester of the Technion University, Haifa, and Roger Thunvik of the Royal Institute of Technology, Stockholm, as a PROPER submodel. It is capable of treating 2–D and axi–symmetric 3–D groundwater flow problems.

1.2 Purpose and Scope of Study

It is known that fractured rock displays great spatial heterogeneity and variability as to its properties, such as hydraulic conductivity. Those properties are furthermore "known" only at a limited number of points in space.

The purpose and scope of this study are to try to find out whether the spatial variability and uncertainty are important from the safety point of view, or if they just average out. For this end, the FSCF10 submodel has been supplemented with the routines necessary to carry out a stochastic simulation. To assess the implication as to the safety of a KBS–3 design repository, a set of safety–related geohydrological performance measures were formulated.

The course of the analysis was:

- create statistical model for hydraulic conductivity based on the stochastic process concept,
- estimate spatial trend and covariance,
- generate conductivity fields with the estimated trend and covariance,
- collect statistics for performance measures.

2 THE GEOHYDROLOGICAL MODEL

The kind of block of undisturbed rock envisaged for the future repository for spent fuel was used as a basis for the modelling and assessment. The block is assumed to be surrounded by fracture zones possessing a high hydraulic conductivity, see Figure 1. The factors producing the hydraulic gradients driving the groundwater flow are associated with the topography.

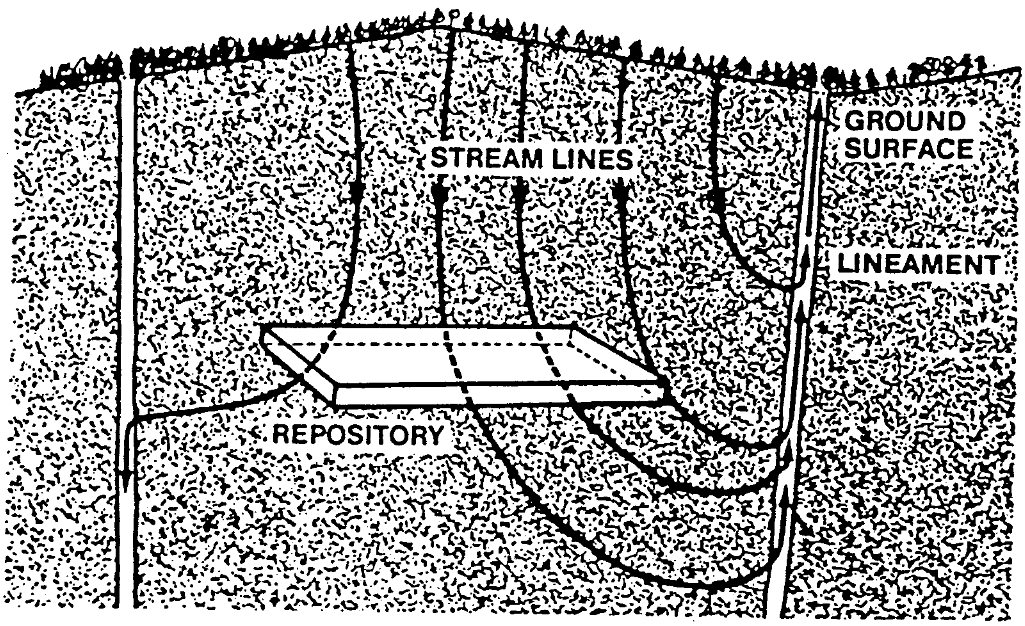


Figure 1

2.1 The Geometrical Model

The above situation is modelled with the 3–D axisymmetric finite element solver FSCF10 developed especially as a PROPER submodel by Braester and Thunvik which is used with the PROPER package. The geometry of the model is shown in Figure 2. The KBS–3 repository is placed at a depth of 500 m in a cylindrical rock block surrounded by fracture zones 100 m from the outermost part of the repository. It is the hill (which is marked with an "a" in the figure) that generates the local gradients causing the groundwater flow in the model. The height of the hill is adjusted to give reasonable values of the annual recharge (see Subsection 2.4.5). Besides the height of the hill the amount of waterflow through the model depends on the conductivity field. Each ring element in the mesh (Figure 3) is assigned a conductivity value generated in the actual realization (see chapter 4).

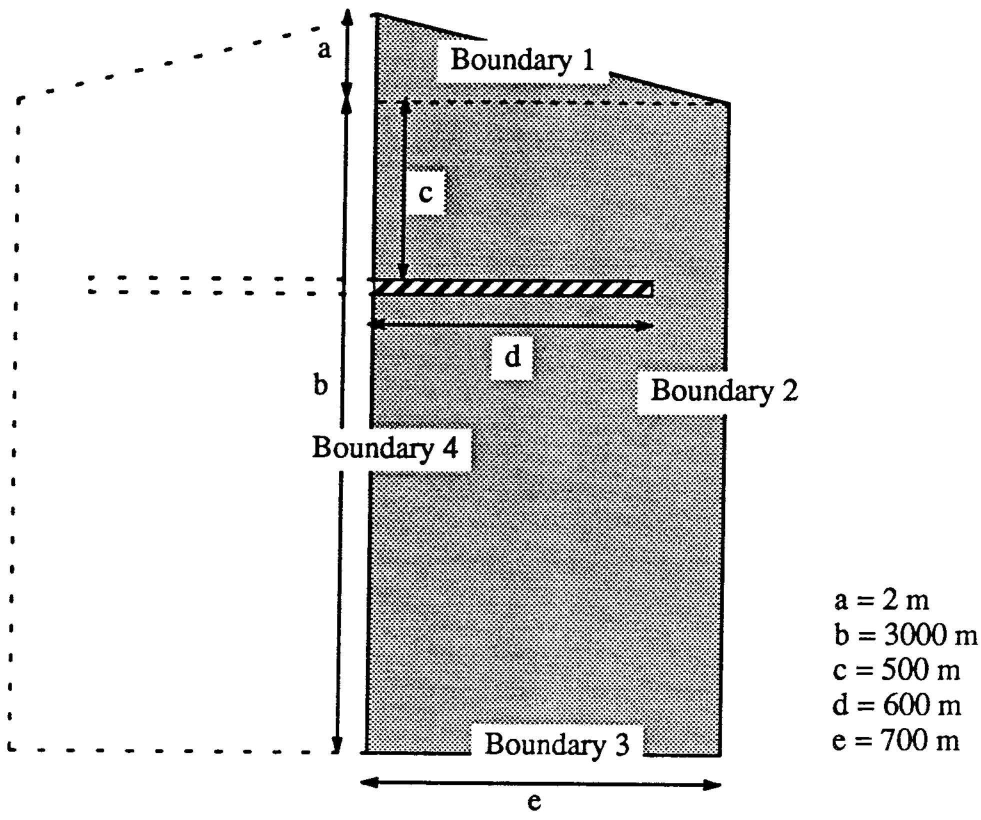


Figure 2

2.2 The Mathematical Model

FSCF10 uses the Representative Elementary Volume (REV) or continuous porous medium formulation.

Heating effects due to the radioactive decay of the waste are not taken into account in this model. Assuming stationary and incompressible flow the following model is used:

$$\nabla \cdot \mathbf{u} = 0 \qquad \mathbf{u} = -K\nabla h \tag{1}$$

where K is the conductivity and h is the head (potential) for incompressible flow defined as

$$h = \frac{p - p_0}{\varrho g} - (z - z_0) \tag{2}$$

where p is the pressure, p_0 is a reference pressure, ϱ is the water density, g is the acceleration of gravity, z is the depth coordinate (increasing downwards) and z_0 is a reference level.

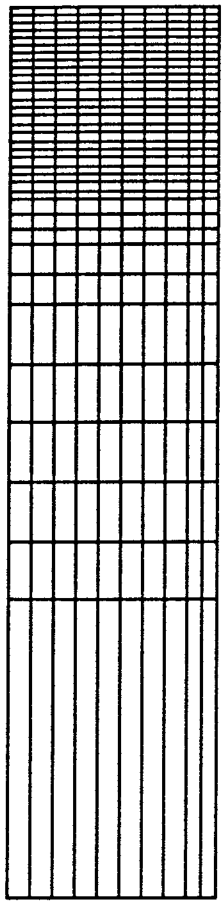


Figure 3, the mesh

2.3 Boundary Conditions

Referring to Figure 2 the following boundary conditions are applied

- **Boundary 1:** p is assigned a constant value p_0.
- **Boundary 2:** $p = p_0 + \varrho g z$, corresponds to the pressure from a column of incompressible water.
- **Boundary 3:** Isolated, $\frac{\partial p}{\partial \mathbf{n}} = 0$, **n** is the surface normal vector.
- **Boundary 4:** $\frac{\partial p}{\partial \mathbf{n}} = 0$. No sources along $r = 0$

2.4 Geohydrological Performance Measures

2.4.1 Rationale

A number of problems are faced in trying to assess geohydrologically related safety factors:

- the geohydrology as such is only of secondary interest; the implications for the transport of radionuclides from the repository to the biosphere is the primary issue,
- the REV formulation does not provide an ideal description for the radionuclide transport; the relationships are unclear/unknown.

The problem is actually that of performing a safety assessment without radionuclide transport modelling, but based on REV geohydrological performance measures only, i e:

- a set of factors must be found having the greatest possible influence on the radionuclide transport,
- those factors must be general enough to permit comparison between different repository layouts.

The strategy adopted in the present study is as follows:

1. Assumption:
   there are **correlations** between REV–geohydrological parameters and radionuclide transport allowing for layout comparisons based on generic sites,
2. performance measures which are associated with the nearfield as well as the farfield are identified,
3. nearfield and farfield performance measures are separated as a first attempt,
4. if nearfield and farfield performance measures are spatially related, this must also be handled in a second attempt.

The following subsections suggest a solution.

2.4.2 Integrated Film Transfer Coefficient or Q_{eq}

The diffusion rate of a radionuclide per unit area from the repository at stationary conditions can be formulated as:

$$N_A = K_v(C_0 - 0) \qquad (3)$$

where $C_0 - 0$ is the difference between the concentration at the boundary of the engineered diffusion barrier and the concentration a large distance away. The mass transfer coefficient, K_v, includes the transport resistance due to the near–stagnancy of the slowly moving groundwater outside the barrier (Ref.1).

The mass transfer can be integrated over the entire repository surface area:

$$N = \left(\int K_v dA \right) \cdot (C_0 - 0) = Q_{eq} \cdot C_0 \qquad (\mathrm{m^3/\,year}) \qquad (4)$$

where Q_{eq} can be regarded as an equivalent groundwater flow rate ($\mathrm{m^3}$/ year) (Ref.2). Previous analyses have shown the major importance of the film resistance. Thus Q_{eq} associated with the film resistance only can be taken as an appropriate nearfield performance measure.

From a hydrological point of view it is the time of contact between a fluid particle and the repository/canister that determines the film resistance. In the model implementation it has been assumed that a fluid particle is in contact with one canister only, as it travels to the fracture zone, i.e. the nuclide concentration in the water already coming in contact with a canister is regarded as zero. The possibility that the water already is contaminated when it arrives at the canister is neglected.

The dimension of a KBS–3 canister is small compared to the scale of the flow field. This fact justifies the use of the very general formulation of penetration theory to compute Q_{eq} (Ref.3). The stationary transport is replaced with an "equivalent Soxhlet test" where the time–dependent equation is solved using a "period" determined by the undisturbed groundwater pore velocity.

$$Q_{eq} = \epsilon \sqrt{\frac{D}{\pi}} \cdot \sum_i \frac{A_i}{\sqrt{\Theta_i}} \tag{5}$$

Where:

- ϵ = porosity = 10^{-4},
- D = effective diffusion coefficient $6*10^{-2}$ m^2/yr,
- A_i = total area of the canisters in element i,
- Θ_i = contact time for a fluid particle that travels along the surface of a canister in element i.

The sum is taken over the elements where the canisters are located.

2.4.3 Mean Travel Distance Velocity or V_d

A performance measure associated with the farfield is also needed. Time is what allows the radionuclides to decay so groundwater travel times from the repository to the biosphere should be important. A performance measure could have the dimensionless form:

$$\frac{1}{A}\int e^{-t_w/\tau}\, dA \tag{6}$$

where t_w represents the groundwater travel time distributed over all the points on the surface of the repository, and where τ is some characteristic time associated with processes involved solution etc. It would be difficult, however, to represent all nuclides with one single τ.

Another possible performance measure would be the average travel time over the repository:

$$\frac{1}{A}\int t_w\, dA \tag{7}$$

but very large times would tend to dominate completely and swamp the fast paths. A reasonable alternative seems to be:

$$V_d = \frac{1}{A}\int \frac{1}{t_w}\, dA \quad (1/\text{year}) \tag{8}$$

V_d is a kind of "travel distance velocity" that expresses the number of path lengths a water parcel travels per unit time averaged over the whole repository. In the computer code V_d is calculated as:

$$V_d = \frac{\sum_i A_i/t_{wi}}{\sum_i A_i} \tag{9}$$

t_{wi} = fluid particle travel time from element i of the repository to the fracture zone.

2.4.4 Combined Nearfield/Farfield Performance Measure or *NFFF*.

The two previous performance measures are correlated spatially via the locations of the different parts of the repository. A portion having a fast diffusion across the film will probably also "see" a short groundwater travel time. A performance measure that takes this fact into account is:

$$NFFF = \int \frac{Qeq}{t_w}\, dA \quad (\mathrm{m^3/\,year^2}) \tag{10}$$

This is implemented as:

$$NFFF = \epsilon \sqrt{\frac{D}{\pi}} \sum_i \frac{A_i/t_{wi}}{\sqrt{\Theta_i}} \tag{11}$$

2.4.5 Annual Groundwater Recharge

The recharge is calculated via the the flow through the right boundary of the rock block, that is the flow into the fracture zone, i.e.

$$Q_{rch} = \int_{Boundary2} \mathbf{v} \cdot \mathbf{n}\ ds/A_L = \frac{\sum_k v_{\perp k} A_k}{A_L} \tag{12}$$

$v_{\perp k}$ = recharge component perpendicular to Boundary 2 in Figure 2,

A_i = area towards Boundary 2 within element *i,*

A_L = total planar top area of the model.

The sum is taken over all elements constituting Boundary 2.

The hill generates the local gradient that drives the flow, so it is the height of this hill together with the realization of the conductivity field that determines the value of the recharge. The height of the hill is calibrated through a few realizations. The aim is that the recharge in the following simulations should vary between 50 and 500 mm per year. In the actual outcome of the simulations, some values are considerably higher than 500 mm/year however, see Figure 22.

It is difficult to distinguish between infiltration and what must actually be regarded as runoff in the modelling situation at hand; a large portion of the flow toward the fracture zone is rather superficial.

3 THE STATISTICAL CONDUCTIVITY MODEL

3.1 The Conductivity Data

The estimation procedures discussed below will be applied on conductivity data from five boreholes at the Klipperås study site, available in SKB's database GEO-TAB (see Appendix 1 for details). In these holes the conductivity is measured in 20 meter packed off sections[1]. Use is made of this regularity as will be described. Note that this regularity does not hold for the depth (i.e. the projection on the vertical coordinate) because of different angles of inclination, bending of holes etc.

Figures 4–6 show the common logarithm of the conductivity versus the depth. There is an overall decreasing trend for the conductivity with increasing depth, but the variation around any trend is huge. Fitting one trend per borehole these trends will be quite different. It also seems possible to discern subtrends within some of the holes, see Figures 14–16.

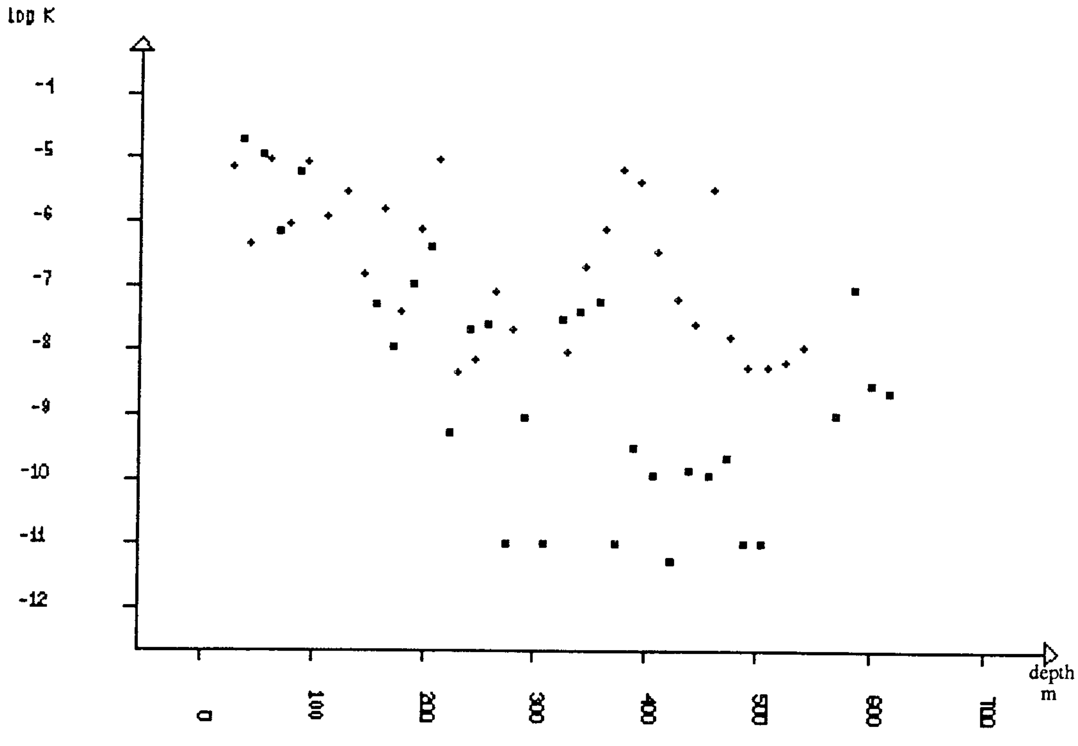


Figure 4, borehole 1 and 2

1. In one of the holes a 10 m section is followed by a 30 m section.

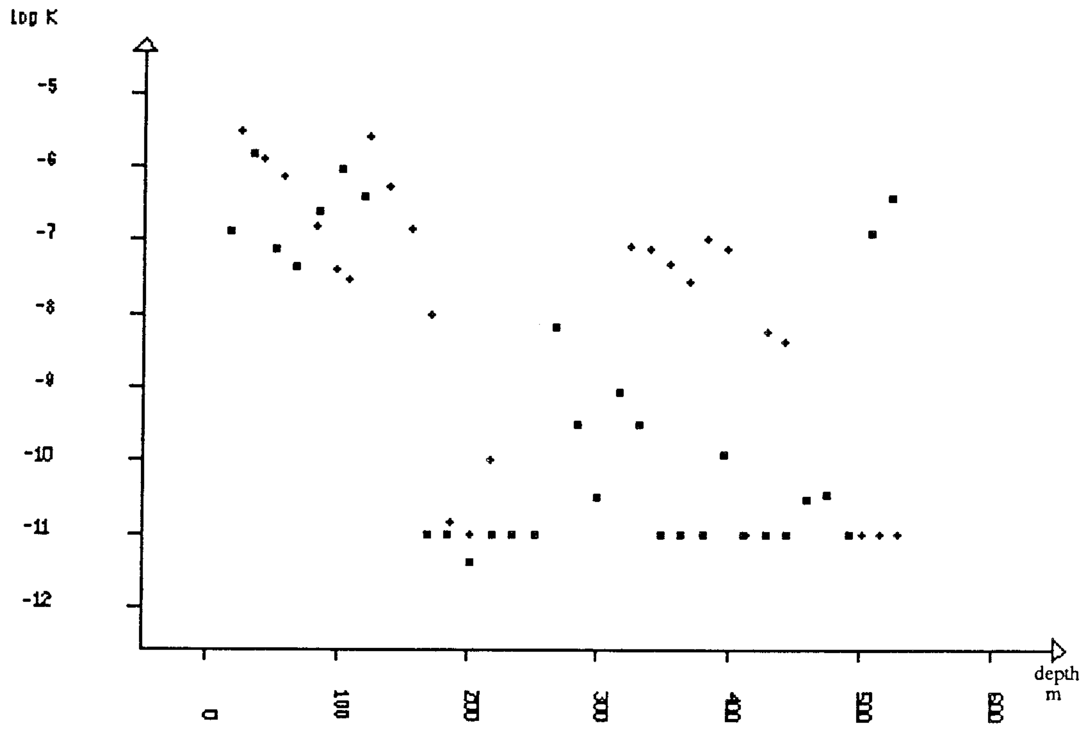


Figure 5, borehole 3 and 4

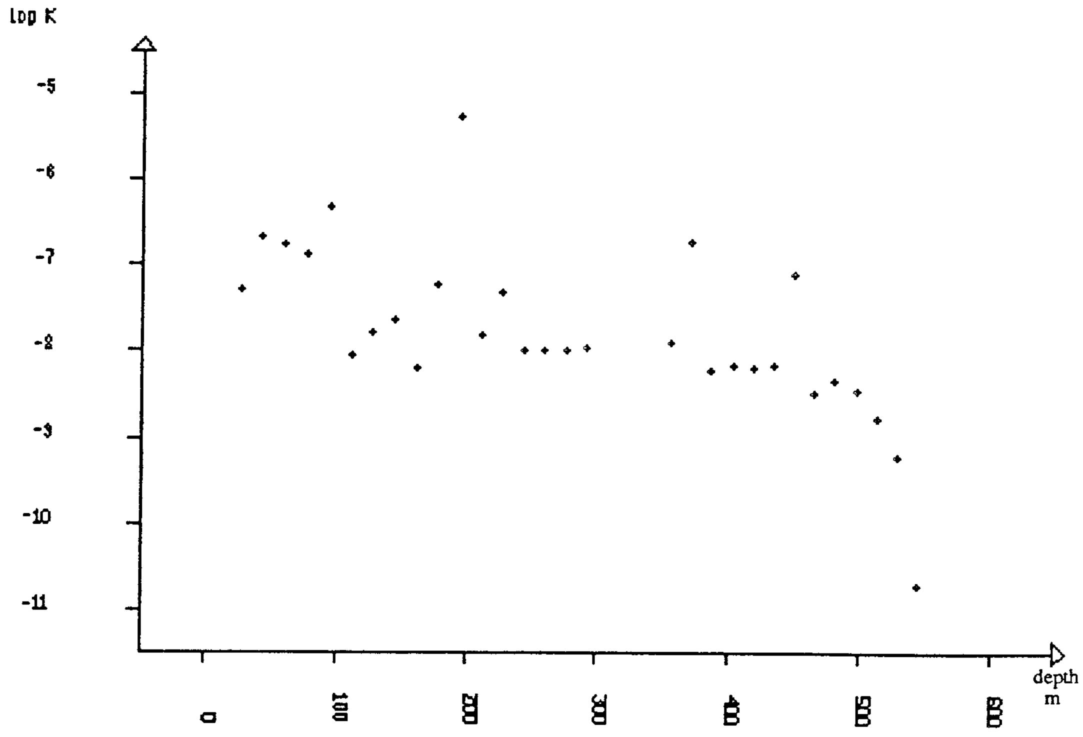


Figure 6, borehole 5

Besides few measuring points, large variation, very different trends for the holes and possible subtrends, there are two additional features that complicate the analysis.

1. **Censoring.** There is a measurement limit (K=10^{-11}m/s)[2]

2. **Missing values.** Our aim is to model good generic rock, therefore points within major fracture zones have been removed, thus destroying the desirable regularity. Those zones which have been located by other means than conductivity measurements are pointed out in Ref.4 .

3.2 Basic Model

Consider the spatial process

$$Y = \log K \tag{13}$$

Assume that

$$X = Y - E[Y] \tag{14}$$

is weakly stationary, isotropic and Gaussian. Besides simplifying the analysis, the small amount of data force us to assume weak stationarity. Here $E[\cdot]$ denotes the expectation value operator. The conductivity measurement was performed every 20 m along the boreholes so we assign every measuring point an integer j. That integer may belong to any of three sets

1. $j \in O$, the conductivity value is observed ($K_j > 10^{-11}$ m/s),
2. $j \in C$, the conductivity lies below the measurement limit,
3. $j \in M$, the information is missing, K_j is removed because the measurement section lies within a major fracture zone.

To simplify the notations below we define $D = O \bigcup C$.

The rock volume and thus the data set is further divided into N parts Ω^i, $i = 1 \ldots N$ each containing m_i data points where n_i of them contain information (i.e. $\in D$)[3]. Approximate $E[Y]$ locally with a trend that varies linearly with the depth i.e. $e[Y] \approx \beta_0^i + \beta_1^i z$ in Ω^i. Now, regard

$$\{x_j\} = \bigcup_{j \in \Omega^i} \{y_j - (\beta_0^i + \beta_1^i z_j)\} \tag{15}$$

as a sample of X . In what follows we denote stochastical variables and processes with capital letters and observations of stochastical variables (i.e. numerical values) with lower case ditto.

2. Two points with values below this limit are detected (see Appendix 1).
3. This also results in the division of O, C, M and D as $\bigcup_i O^i$ $\bigcup_i C^i$ $\bigcup_i M^i$ and $\bigcup_i D^i$ respectively

The problem is now:

a) Estimate the trend parameters β_0^i and β_1^i for each i,

b) Estimate the covariance for X.

Three cases are explored:

1. $N = 1$, just one regression line is fitted to all data,
2. $N = 5$, one regression line per borehole,
3. $N = 9$, one regression line per subtrend.

To simplify the expressions the index i will be dropped if there is no risk of confusion.

Note:

- The covariance between X_j and $X_{j+\tau}$ denoted $r(X_j, X_{j+\tau})$ is regarded as zero if X_j and $X_{j+\tau}$ belong to different holes. The distance $d_{i\ i+\tau}$ between the points are considered as infinite.

- The set M is defined to simplify the notation in what follows, so that the 20 m regularity for points that belong to the same hole is preserved (e.g. so it is clear that the residual X_{j+3} lies three 20m lags beyond X_j if it exists and if X_{j+3} and X_j belong to the same borehole). In the following regression expression points corresponding to $j \in M$ should be ignored. In the covariance estimation residuals in these points are set to zero (i.e. $x_j = 0$ if $j \in M$).

3.3 The Trend Estimation

Two methods for the regression analysis are used, iterative generalized least square estimation (IGLSE) and maximum likelihood estimation (MLE).

3.3.1 Iterative Generalized Least Square Estimation (IGLSE)

In the ordinary least square estimation (LSE) one is searching for the choice of the parameters β_0 and β_1 that minimize the euclidian norm of the residuals, i.e. the parameters are assigned values that satisfy

$$\frac{\partial\left[\sum_j x_j^2\right]}{\partial\beta_0^i}=0 \qquad \frac{\partial\left[\sum_j x_j^2\right]}{\partial\beta_1^i}=0 \qquad i=1,,\ldots,N. \tag{16}$$

The equations are solved for each value of i ,which will be dropped henceforth in this section.

Now write equation (15) in matrix form as

$$\mathbf{x}=\mathbf{y}-Z\mathbf{b} \tag{17}$$

where

$$\mathbf{x}=\{x_j\} \qquad \mathbf{y}=\{y_j\} \qquad Z=\{1,z_j\} \qquad \mathbf{b}=\begin{bmatrix}\beta_0\\ \beta_1\end{bmatrix} \tag{18}$$

and the boldface letters represent column vectors and Z denotes a $m_i \times 2$ matrix.

Then, with these notation the solution to (16) is given by the solution of

$$Z^T Z\mathbf{b}=Z^T\mathbf{y} \tag{19}$$

This analysis assumes, from a stochastical point of view, that the components of $\mathbf{x}$ are independent, and have equal variances. If the components do not satisfy this the modified form:

$$Z^T V^{-1} Z\mathbf{b}=Z^T V^{-1}\mathbf{y} \tag{20}$$

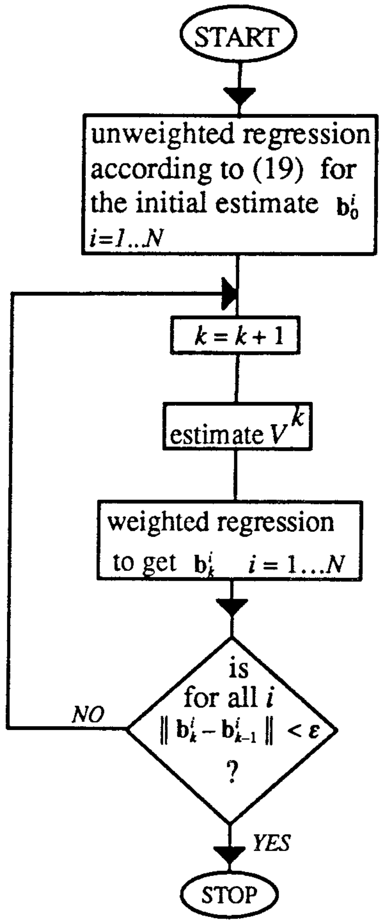
has to be used where V is the covariance matrix of X i.e. the component in row k column l of V is

$$v_{kl}=E[X_k\ X_l]\approx r(|k-l|). \tag{21}$$

Due to the stationarity assumption the covariances are only dependent on the difference between the indices i.e. the lag. Here and in what follows $r(|k-l|)$ signifies an estimate of the covariance $E[X_k\ X_l]$. This estimate is calculated using (33) below.

The expression (20) is in fact the LSE of the transformed expression[4] $\mathbf{F}=V^{-1/2}\mathbf{X}$. The neat calculation $E[\mathbf{F}\mathbf{F}^T]=V^{-1/2}E[\mathbf{X}\mathbf{X}^T]V^{-1/2}=I$ shows that the transformed vector $\mathbf{F}$ has independent coefficients with unit variance (I denotes the unit matrix). This method (20) is called weighted regression (see Ref.5).

Since we do no know the covariance matrix when the regression parameters are computed, the following iterative method is used:



This method is referred to in the literature as Iterative Generalizes Least Square Estimation (Ref.6)

3.3.2 Maximum Likelihood Estimation (MLE)

This is a way to estimate both the trend(s) and the covariance in the same stroke. The idea is to maximize the probability of the observation over some parameter space. In our case the probability of the observation may be written

$$L = \prod_{i=1}^{N} \left\{ \prod_{j \in C^i} \int_{-\infty}^{\varepsilon-(\beta_0^i+\beta_1^i z_j)} dt_j \right\} f_n(\mathbf{t})|_{t_j = x_j, j \in O^i} \qquad (22)$$

The expression in braces is a multidimensional integral operator on components corresponding to censored values, ε is the lowest measurable value of $\log K$, $f_n(\mathbf{t})$ is the probability function for the stochastic residual vector $\{\mathbf{X}\}_{j \in D}$ of dimension n

$$n = \sum_{i}^{N} n_i \qquad (23)$$

4. Here X denotes the stochastical vector that is observed in (18).

Since $\mathbf{X}$ was assumed Gaussian

$$f_n(\mathbf{t}) = \frac{1}{\sqrt{(2\pi)^n \det(V)}} \exp\left(-\frac{\mathbf{t}^T V^{-1} \mathbf{t}}{2} \right)$$
$$V = \left\{ E[X_k X_l] \right\}_{k,l \in D} \tag{24}$$

The parameters which we are to maximize over are the trend parameters $\{\beta_0^i, \beta_1^i\}_{i=1}^N$ and also the parameters describing the covariance. For instance an exponential covariance model is parametrized by r_0 and d_0 as

$$r(z_j - z_k) = r_0 \exp\left(-\frac{d_{jk}}{d_0} \right) \tag{25}$$

d_{ij} denotes the distance between two points in the borehole ($d_{jk} = 20 \cdot |j-k|$ metres). We see that r_0 can be identified with the variance of the process (c.f Subsection 3.4.1)

Due to the assumption that the residuals of different holes are uncorrelated V obtains a the block diagonal structure and (22) may be rewritten as

$$L = \prod_{i=1}^{N} L^i \ ,$$
$$L^i = \left\{ \prod_{j \in C^i} \int_{-\infty}^{\varepsilon - (\beta_0^i + \beta_1^i z_j)} dt_j \right\} f_{n_i}(\mathbf{t})|_{t_j = x_j, j \in O^i} \tag{26}$$

where now

$$f_{n_i}(\mathbf{t}) = \frac{1}{\sqrt{(2\pi)^{n_i} \det(V^i)}} \exp\left(-\frac{\mathbf{t}^T (V^i)^{-1} \mathbf{t}}{2} \right)$$
$$V = \left\{ E[X_k X_l] \right\}_{k,l \in D^i} \tag{27}$$

In spite of the apparent simplicity this approach includes calculating N multidimensional integrals as well as inverting N covariance matrices in each step of the optimization. Since the dimension of these integrals equals the number of censored values for each trend, which is large at the Klipperås site, the straight forward approach to try to maximize this function seems unfeasible.

However, if our primary interest is to evaluate the trends and we are willing to neglect the influence on these from the covariances (it seems to be small (see Section 5.2)) , the set of parameters are reduced, but more important the integrals split. In fact when assuming independent equally distributed components X_j the covariance matrix becomes $r_0 I$ and thus (22) becomes

$$L\left(\{\beta_0^i,\beta_1^i\}_{i=1}^N,\sigma\right)=\prod_{i=1}^N\prod_{j\in O^i}\frac{1}{\sigma}\phi\left(\frac{y_j-\beta_0^i-\beta_1^i z_j}{\sigma}\right)\prod_{j\in C^i}\Phi\left(\frac{\varepsilon-\beta_0^i-\beta_1^i z_j}{\sigma}\right) \quad (28)$$

The expression splits into products. $\phi(\cdot)$ is the one–dimensional standardized distribution function $\phi(t)=\frac{1}{\sqrt{2\pi}}e^{-t^2/2}$ and Φ its cumulative counterpart,

$$\Phi(t)=\int_{-\infty}^{t}\phi(u)\,du.$$

Taking the logarithm of the likelihood function L in (28) and setting the partial derivatives with respect to β_0,β_1 and σ to zero, we get the following set of non-linear equations to solve for the parameters

$$\frac{\partial\log L}{\partial\beta_0^i}=\frac{1}{\sigma^2}\sum_{j\in O^i}(y_j-\beta_0^i-\beta_1^i z_j)-\frac{1}{\sigma}\sum_{j\in C^i}\frac{\phi[(\varepsilon-\beta_0^i-\beta_1^i z_j)/\sigma]}{\Phi[(\varepsilon-\beta_0^i-\beta_1^i z_j)/\sigma]}\quad i=1,\ldots,N$$

$$\frac{\partial\log L}{\partial\beta_1^i}=\frac{1}{\sigma^2}\sum_{j\in O^i}z_j(y_j-\beta_0^i-\beta_1^i z_j)-\frac{1}{\sigma}\sum_{j\in C^i}z_j\frac{\phi[(\varepsilon-\beta_0^i-\beta_1^i z_j)/\sigma]}{\Phi[(\varepsilon-\beta_0^i-\beta_1^i z_j)/\sigma]}\quad i=1,\ldots,N \quad (29)$$

$$\frac{\partial\log L}{\partial\sigma}=-\frac{n}{\sigma}+\frac{1}{\sigma^3}\sum_{i=1}^N\sum_{j\in O^i}(y_j-\beta_0^i-\beta_1^i z_j)^2-\frac{1}{\sigma}\sum_{i=1}^N\sum_{j\in C^i}\frac{\left[(\varepsilon-\beta_0^i-\beta_1^i z_j)/\sigma\right]\cdot\phi\left[(\varepsilon-\beta_0^i-\beta_1^i z_j)/\sigma\right]}{\Phi\left[(\varepsilon-\beta_0^i-\beta_1^i z_j)/\sigma\right]}.$$

where n is the total number of observed points. If we for example want to estimate five trends we get a set of eleven equations to solve (five for β_0^i, five for β_1^i and one for σ). We solve the equations using an iterative method suggested by Sampford and Taylor described in Ref.7

If there are no censored values the equations (29) become linear and identical to the LSE equations (17)

3.3.3 Properties of the Regression Methods

One major problem with the IGLSE as here described compared with the MLE is the correct handling of the censored values. Simple substitution of the measurement limit leads to biased estimates (Ref.8 –10 discuss this problem and present some ideas on how to deal with them). On the other hand it is practicable to take the correlation between the values into consideration in the IGLSE.

In the MLE it is important that the assumed distribution function does not deviate too much from the real underlying one and that the ratio n_i/σ is big (not necessarily the case for our data), otherwise a large bias is likely to appear in the regression analysis (Ref.8).

Both of the methods discussed in this section have been applied to our data (see chapter 5).

3.4 Estimation of the Covariance Function

In the literature covariance estimation is discussed when values are missing or censored, however we do not know any text suggesting estimators for samples when both of these features are present.

3.4.1 Fitting a Function

Two methods for pointwise estimation of the covariances for lags that are multiples of 20 m are presented below. A continuous function is then fitted for use in a generator that supplies the FSCF10 model with a stochastic conductivity field (see chapter 4). It can easily be shown that only positive definite functions is qualified for this purpose. We use the truncated exponentially decreasing function

$$r(d) = \begin{cases} r(0)\exp(-d/d_0), & d \leq d_{cut} \\ 0, & d > d_{cut} \end{cases} \tag{30}$$

where d is the distance between the points under consideration. Thus the correlation between the K–values in the model are assumed to be isotropic. The cutoff parameter d_{cut} is used speed up the generator that will be discussed below. The variance $r(0)$ is taken from the regression analysis, the constant d_0 is fitted by eye to expression (30) for the discrete values of $r(\tau); \tau = 0,1,2,3,$ i.e. for the lags 0,20,40,60 m. Only these smallest lags were used, because the covariance estimators get less reliable for greater lags due to smaller amount of data participating in the estimation (c.f the results in chapter 5). The value of d_0 was also estimated by the MLE on the logarithm of (30). This logarithmic transformation gives a curve that fits well for low values of r at the expense of a poor fit for greater r which is not desirable. There are certainly methods for the optimal fitting of the function (30) based on some condition (e.g. minimum euclidian norm of the residual) but for our present purpose and with respect to the uncertainties in the pointwise covariance estimation (see chapter 5) the simple fitting by eye will do.

3.4.2 The Classical Estimator

Parzen (Ref.11) views a series with missing values x (from a covariance estimation point of view) as the result of amplitude modulating an imagined series w (with no missing values) with an indicator series a i.e.

$$x_j = a_j w_j; \qquad a_j = \begin{cases} 0 & j \in M \\ 1 & j \in D \end{cases} \tag{31}$$

where M denotes the set of missing values and D denotes the complement of the former (Section 3.2). Assume, for the moment, that the series w is available. Then the classical covariance estimator

$$r(\tau) = \frac{1}{m} \sum_{i=1}^{N} \sum_{j=1}^{m_i - \tau} (w_j^i - \langle w^i \rangle)(w_{j+\tau}^i - \langle w^i \rangle) \tag{32}$$

can be used. Here $\langle w^i \rangle$ denotes the mean value $\sum_{j=1}^{m_i} (w_j^i)/m_i$ for trend i which in the case of LSE equals zero, m_i denotes the number of points corresponding to trend i. The use of $m = \sum_{i=1}^{N} m_i$ instead of $\sum_{i=1}^{N} m_i - \tau$ (the total number of participating pairs with the lag τ) in the denominator gives the estimator a slight bias, on the other hand the mean square error in synthetic testing becomes lower (Ref.12) and it guarantees a positive definite covariance matrix.

When, as in our case, only the amplitude modulated sequence x is detectable (e.g. the series contains missing values) Dunsmuir (Ref.13) suggests

$$r(\tau) = \frac{1}{n} \sum_{i=1}^{N} \sum_{j=1}^{m_i - \tau} (x_j^i - \langle x^i \rangle)(x_{j+\tau}^i - \langle x^i \rangle) a_j a_{j+\tau} \tag{33}$$

($n = \sum_{i=1}^{N} n_i$) which is based on Parzen's ideas. As one can see it is very similar to the classical estimator (32). Here we substitute the censored values by the measurement limit, which gives a biased estimator.

3.4.3 The Robinson Estimator

Robinson (Ref.14) proposes an alternative estimator when some of the data are censored. Define the stochastic variable

$$W_j = \frac{X_j}{\sigma} \tag{34}$$

Now let δ_τ^{rs} be the moment:

$$\delta_\tau^{rs} = \int_0^\infty \int_0^\infty t_1^r t_2^s \, f_\tau(t_1, t_2) \; dt_1 dt_2 \tag{35}$$

where $f_\tau(t_1, t_2)$ is the bivariate distribution function:

$$f(t_1, t_2) = \frac{1}{2\pi \det V} \exp\left\{ -\frac{(t_1, t_2)\; V^{-1} (t_1, t_2)^T}{2} \right\} \tag{36}$$

$$V = \begin{bmatrix} 1 & r_W(\tau) \\ r_W(\tau) & 1 \end{bmatrix} \tag{37}$$

Here r and s are positive integers, and r_W is the covariance for W. This is consistent with the assumption that X is Gaussian.

The parameter σ in (34) is not known. The MLE as well as the IGLSE provides an estimate $\hat{\sigma}$ of the standard deviation. Robinson proposes (no trend is present in his case) the estimator

$$\hat{\sigma}^2 = \frac{\sum_j x_j^2 \; I(x_j > 0)}{\sum_j I(x_j > 0)} \tag{38}$$

where $I(\cdot)$ is an indicator function which is equal to one if the condition in the argument is fulfilled. As an estimator of (35) in the discrete case with T points Robinson suggests

$$\hat{\delta}_\tau^{rs} = \frac{1}{T-\tau} \sum_{j=1}^{T-\tau} w_j^r w_{j+\tau}^s \; I(w_j \geq 0, w_{j+\tau} \geq 0) \tag{39}$$

In our case with missing values we use

$$\hat{\delta}_\tau^{rs} = \frac{1}{l_\tau} \sum_{i=1}^{N} \sum_{j=1}^{m_i-\tau} (w_j^i)^r (w_{j+\tau}^i)^s \; I(w_j^i \geq 0, w_{j+\tau}^i \geq 0, \;\; j, \; j+\tau \in D) \tag{40}$$

where l_τ is the total number of existing pairs with the lag τ.

Now, there are formulas that relate $r_W(\tau)$ to δ_τ^{rs} for different values of r and s given the assumption of normality. In our case we use $r = s = 1$, $r_W(\tau)$ is then given by the implicit equation

$$\delta_\tau^{11} = \frac{1}{2\pi} \left\{ r_W(\tau)\left(\frac{\pi}{2} + \arcsin r_W(\tau)\right) + \sqrt{1 - r_W(\tau)^2} \right\} \tag{41}$$

Other value on r and s may also be used giving other relations between $r_W(\tau)$ and δ_τ^{rs}. We have tested a few of them but the former seems to perform best.

Two ways to obtain the covariances are then possible, either the straightforward

$$r(\tau) = \hat{\sigma} \cdot r_W(\tau) \tag{42}$$

or one inspired by the discussion in previous subsection

$$r(\tau) = \frac{l_\tau^i}{n_i} \cdot \hat{\sigma} \cdot r_W(\tau) \tag{43}$$

4 SIMULATING THE CONDUCTIVITY FIELD

When the parameters of the regression model and the covariance function is determined, the next step is to use this information to simulate the conductivity field for the FSCF10 model.

4.1 Generating the Conductivity Field from Statistical Data

In Ref.15 the following multivariate normal generator is proposed. Suppose we have m points in the FSCF10 model where we desire to generate a value for the conductivity. Since we have assumed X to be Gaussian the vector $\mathbf{X} = (X_1, X_2, \ldots, X_m)$ becomes multivariate normally distributed with a covariance matrix the components of which depend on the distance between the points according to (30)

$$V = \begin{bmatrix} r(d_{11}) \ldots & \vdots & \vdots & \vdots & \ldots\, r(d_{1m}) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & r(d_{i-1i-1}) & r(d_{i-1i}) & r(d_{i-1i+1}) & \vdots \\ \vdots & \ldots\, r(d_{i\;i-1}) & r(d_{i\;i}) & r(d_{i\;i+1}) \ldots & \vdots \\ \vdots & r(d_{i+1i-1}) & r(d_{i+1i}) & r(d_{i+1i+1}) & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ r(d_{m1}) \ldots & \vdots & \vdots & \vdots & \ldots\, r(d_{mm}) \end{bmatrix} \tag{44}$$

As a first step, we compute a lower triangular matrix L by Cholesky factorization, that is $V = LL^T$. Then

1. A vector $\mathbf{Z}$ with m independent standard normal components (i.e. $E(Z_j) = 0$ and $v(Z_j) = 1$) are generated using the PROPER random number generator.
2. The dependent vector is computed as $\mathbf{X} = L\mathbf{Z}$.
3. The desired conductivity is now calculated from (15) given any desired trend parameters for the site studied c.f Section 5.4.

The easy calculation

$$V = E[\mathbf{X}\mathbf{X}^T] = E[(L\mathbf{Z})(L\mathbf{Z})^T] = E[L\mathbf{Z}\mathbf{Z}^T L^T] = LE[\mathbf{Z}\mathbf{Z}^T]L^T = LL^T \tag{45}$$

shows that the generator works.

4.2 Problems Faced in the FSCF10 Modelling

There are two major problems that have to be solved before the rock can be simulated in the FSCF10 program on the basis of the statistical information extracted from the covariance and trend estimators.

1. **3D–2D**: A 2–d model is used for a true 3–d phenomenon.
2. **scale**: The spatial scale of the FSCF10 finite elements does not correspond to the length of the sections of measurement for the conductivity data.

4.2.1 3D–2D

The numerical model used in this study is three–dimensional axisymmetric, which is adequate in homogeneous formations. However when aiming towards simulating the response of real heterogeneous systems this restriction is disturbing. The only way to give some justification to the use of this model is to work with φ –averages i.e. all our state variables are considered as being averages over the angular coordinate in a cylindrical coordinate system. For example denoting the real three–dimensional head h_3 our notation h in two dimensions is defined by

$$h(r,z) = \frac{1}{2\pi} \int_0^{2\pi} h_3(r,\varphi,z)\ d\varphi \tag{46}$$

It is rather easy to show that the ordinary (two dimensional) continuity equation holds for the φ –averaged Darcy velocity i.e.

$$\nabla \cdot \int \frac{v(r,\varphi,z)}{2\pi}\ d\varphi = 0 \tag{47}$$

but when turning to Darcy's law in particular and conductivity in general the situation becomes more cumbersome. If we start by accepting the three dimensional form of Darcy's law i.e.

$$u_3 = -K_3 \nabla h_3 \tag{48}$$

where the subscript "3" is used temporarily to distinguish two– and three–dimensional fields it is a natural question to ask whether one can average such a relation and obtain as an approximation

$$< u_3 >_\varphi = -< K_3 >_\varphi \cdot \nabla < h_3 >_\varphi \tag{49}$$

This is what one obtains if one performs a standard perturbation analysis of (48) invoking the assumption

$$| < (K_3 - < K_3 >_\varphi)\ \nabla(h_3 - < h_3 >_\varphi) >_\varphi | << | < K_3 >_\varphi \nabla < h_3 >_\varphi | \tag{50}$$

Now two objections can be raised against this approach, one theoretical and one practical. The theoretical one is easy enough to explain since it is merely that (50) does not seem to hold in view of the fact that the variance of $\log K$ is about two. The practical one is that if we accept (50) and want to utilize our statistical information of K_3 we need to infer the stochastical quantities of $< K_3 >_\varphi$ from those of K_3 . This is straightforward for the stochastic moments but not at all easy for the distribution of $< K_3 >_\varphi$.

Having pointed out these difficulties with a formal approach we now shift our point of view. The basis of the analysis is the measurements of the conductivity. These are based on assumptions of homogeneity and isotropy and an approximative solution to the resulting hydrology equation. The result of this approximative analysis is known as Moyes formula. The calculated conductivity values are interpreted as point values or effective values for some support (surrounding volume). Now since one prior to the analysis has assumed homogeneity and isotropy the problem opens for an approach using cylindrical coordinates. Then the difference between 2D and 3D conductivity is academic. Hence, any value of K resulting from such an analysis could equally be interpreted as a two dimensional ("ring") conductivity. To connect to Moyes formula we refer to the findings of Braester and Thunvik (Ref.16) who showed that the difference between Moyes formula and a correct numerical cylindrical analysis is small. Thus we may interpret the results from Moyes formula as an effective ring conductivity for some ring centred at the measurement section.

Implicitly the above discussion assumes that effective ring conductivities exist i.e. that a φ –averaged form of Darcys law holds

$$< v >_{\varphi} = -K_2 \nabla < h >_{\varphi} \tag{51}$$

However, this assumption, even for large rings is a compelled one to study the present problem with a 2–D axisymmetric solver.

Finally, in order to wrap things up, we must include two more assumptions:

a) The distribution of ring conductivity is independent of the diameter of the ring.

b) The covariance of the ring conductivities is isotropic.

The assumption a) is contained in the statement that if the rings are identified by cylindrical coordinates (r, z), the process $K_2(r, z)$ is stationary. The assumption b) is written as

$$\begin{aligned} E[(K_2(r_1, z_1) - E[K_2(r_1, z_1)]) \cdot (K_2(r_2, z_2) - E[K_2(r_2, z_2)])] = \\ = C([(r_1 - r_2)^2 + (z_1 - z_2)^2]^{1/2}) \end{aligned} \tag{52}$$

The assumption b) gives us the possibility to estimate the covariance from our drill hole measurements. An assumption of this kind is difficult to avoid if one wants to estimate the covariance structure directly from measurements in boreholes with a uniform angle of inclination. The assumption a) is difficult to validate. One could view it as an assumption of self similarity i.e. that the rock responds in the same way in different scales.

4.2.2 Scale

Since the numerical method used is finite elements, the conductivity values to be simulated is thought of as the constant conductivity of the element. Hence even if these simulated values were independent an element size dependent covariance would result. For example, if the elements have a characteristic length a the induced covariance becomes something like

$$C_a(d) = \begin{cases} V\left(\frac{|a-d|}{a}\right) & d < a \\ 0 & d \geq a \end{cases} \tag{53}$$

where V is the variance of the random conductivity. In particular this implies the impossibility to study "white noise conductivity" with Monte Carlo methods.

Now if we want to simulate a random field with a given covariance it is crucial to select the mesh size so that the effect of the given covariance is studied and is not drowned in the mesh–induced covariance. One way of studying this is to use spectral analysis. Let us assume that we want to simulate a random field with a exponential covariance i.e.

$$C(d) = V\exp\left(-\frac{|d|}{d_0}\right) \tag{54}$$

Letting ω denote the wave vector the corresponding spectral density is

$$S(\omega) = \frac{V}{\pi^2}\frac{d_0^3}{(1+d_0^2|\omega|^2)^2} \tag{55}$$

Now if again the characteristic mesh size is a, for instance a could be the side in a square mesh, the Nyquist angular frequency is $2\pi/a$. The Nyquist frequency is the highest frequency that gives an unique trace on the mesh i.e. for any frequency higher than this there is also a lower one taking the same value over the mesh. This last effect is known as aliasing. So ideally we would like the spectral density to be zero above the Nyquist frequency but clearly this is to much to ask. We have to content ourselves with taking the mesh so small that the spectral density is small above the Nyquist value or that the "energy" of the field above the Nyquist frequency is small. For example, in our case a typical value of d_0 is 26 m and the maximum mesh length in the radial direction is 75 m then

$$\frac{S(\omega_N)}{S(0)} = \frac{1}{(1+d_0^2\omega_N^2)^2} \approx 0.03 \tag{56}$$

$$\frac{\int_{\omega_N}^{\infty} S(\omega)\, d\omega}{\int_0^{\infty} S(\omega)\, d\omega} = 1 - \frac{2}{\pi}(\arctan \omega_N d_0 - \frac{\omega_N d_0}{1+\omega_N^2 d_0^2}) \approx 0.363 \tag{57}$$

This seems reasonably small. However the influence on the calculated flow of the inability to resolve higher frequencies in the conductivity is a separate problem in its own right which we have not addressed so far.

When the extent of an element approximately corresponds to the scale of measurement, its conductivity is calculated based on the depth of its centre. As can be seen from Figure 3 and the size of the model given in chapter 2 some of the elements in the FSCF10 mesh have a side which length is considerably greater than 20 m e.g. the height of the elements deep down in the model. This is because the spatial variation of the solution there is expected to be small.

A solution for high elements deep down in the model is to subdivide them into a number of 20 m high elements (Figure 7) and simulate a conductivity for each of these subelements. An effective conductivity for the high elements is then calculated as the mean conductivity of the subelement. The validity of this method may be inferred as follows.

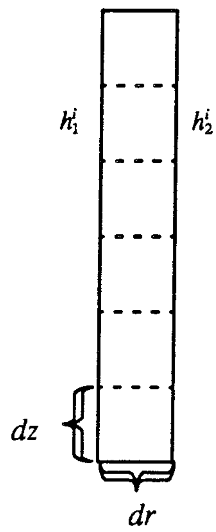


Figure 7

Suppose we have a high element that is divided into n subelements. Denote by v_i the Darcy velocity for element i, then the flow through the element is

$$q_i = dzv_i \approx -dzK_i(h_2^i - h_1^i)/dr \tag{58}$$

according to (1). The total flow for all the subelements is $q_{tot} = \sum_i v_i dz$. The average velocity over this large element is now

$$v_{tot} = \frac{q_{tot}}{\sum_i dz} = \frac{\sum_i dzK_i(h_1^i - h_2^i)/dr}{\sum_i dz} \approx$$
$$\approx ((<h_1> - <h_2>)/dr) \cdot \frac{dz}{\sum_i dz} \sum_i K_i = ((<h_1> - <h_2>)/dr) <K> \tag{59}$$

where $<h_k> = \sum_{i=1}^{n} h_k^i/n,\ k = 1,2$ and $<K> = \sum_{i=1}^{n} K_i/n$. Note that this derivation assumes the gradient to be parallel with the radial coordinate. Hence the gradient in the r–direction must be the dominant in order to use this averaging. which is the case in our FSCF10 model.

5 RESULTS

This chapter presents the results for different estimators which are grouped into two categories. In the first, IGLSE–regression is carried out to get the trend parameters. In this procedure one has to estimate the covariance matrix. When the iterations have converged, a covariance function is provided in the same stroke. In this method, referred to as IGLSE/classical in what follows, no special attention is paid to the censored values which are substituted with the measurement limit.

In the second category explored, we use MLE to estimate the trend parameters and the Robinson method to estimate the covariance. In both these methods the censoring is handled in a more adequate way. In the regression adopted here the covariance influence is ignored. This second category is called MLE/Robinson estimation in what follows.

1000 realizations were used in each of the final simulations made to obtain the uncertainties in the performance measures.

5.1 A Simple Estimator Test

In order to see how the estimator chain MLE/Robinson performs we have generated a number of synthetic data sets each consisting of three series of values for $\log K$, using PROPER's random number generator. Every set has an unique collection of properties (such as number of points, covariance structure in the parent distribution etc.) but they are based on the same random seed.

When generating these synthetic data sets we assume the regression parameters β_0 and β_1 to be equal to zero. The number of values is set to either 35 or 100. They appear regularly on data points every 20m with the first point at the depth 100 m . $\log K$ is assumed to be normally distributed with unit variance and an exponential covariance function according to (30). Three different values for the covariance of the first lag are used 0, 0.2 and 0.5. Note that the regression parameters and covariances are used as input for the the data generation process (which is based on the ideas presented in Section 4.1) and should not be confused with the actual outcome of the estimation discussed below. The censoring level is set to –0.43 which causes about 1/3 of the population generated to be censored. Estimation with missing values have been tested. We have also tested Robinson estimation with other values than one on the integers r and s in (40).

The next step is to estimate the regression parameters and the covariance from these three series (here the series are treated as three different processes i.e. the covariance is calculated for each series separately). The results of this simple test is not encouraging for the MLE/Robinson–estimator. When the population of 100 points is used the maximum absolute error for the estimator for the first 4 lags is about 0.25 units i.e. 25% of the variance. The estimation gets even worse when no MLE–regression is performed i.e. the trend with the parameter values $\beta_0 = \beta_1 = 0$ is used when computing the residuals. This indicates that the great deviation is not due to a bias in the MLE–regression. Appendix 2 contains tables for the complete test.

For comparison purposes we tested the IGLSE/classical estimator on one of the sets. Here,the three series contains 100 points. In this comparison test the parent distribution has the unit variance and no correlation are assigned among the points (i.e. the covariance is zero). The use of this estimator gives an maximal error in the covariance estimation of about 7% for the first lags as can be seen in Table 1 below, which shows the result of the comparison for this set of series. On the other hand the variance ($r(\tau = 0)$ c.f the notation in Subsection 3.4.1) is severely underestimated.

	series1		series2		series3		parent
	IGLSE/– classical	MLE/– Robinson	IGLSE/– classical	MLE/– Robinson	IGLSE/– classical	MLE/– Robinson	
β_0	0.3583	0.2165	0.2699	0.0358	0.1548	–0.0249	0
β_1	$-1.451*10^{-4}$	$-1.868*10^{-4}$	$-3.607*10^{-5}$	$-7.123*10^{-5}$	$-4.877*10^{-5}$	$-1.247*10^{-4}$	0
$r(\tau = 0)$	0.515	0.900	0.583	1.184	0.375	0.851	1
$r(\tau = 1)$	0.008	0.014	0.003	0.076	0.023	0.123	0
$r(\tau = 2)$	–0.020	–0.145	0.049	0.233	–0.031	–0.087	0
$r(\tau = 3)$	0.002	–0.121	–0.070	–0.119	–0.069	–0.163	0
$r(\tau = 4)$	0.054	0.024	0.009	0.072	0.051	–0.101	0

Table 1

A more comprehensive statistical analysis of the estimators ought to be done however in order to draw general conclusions.

5.2 Results of the Conductivity Covariance Estimation

Figure 8 and 9 show the result of the covariance estimation when just one regression line (trend) is fitted to all the data (i.e. $N = 1$ in (15)) for the classical and the Robinson method respectively. Figure 8 contains two almost coinciding plots. In one, the regression parameters are determined by *unweighted* least square estimation (MLE) according to (19), in the other the *weighted* regression (20) is used where the covariance matrix V is estimated by iteration. Obviously, the differences are very small for the covariance estimation which also holds for the regression parameters themselves. Figure 9 shows different results of the Robinson method due to the two ways of estimating $\hat{\sigma}$ discussed in Subsection 3.4.3.The cross marks represent the covariances when (38) is used, and the boxes when the $\hat{\sigma}$ estimate is supplied from the MLE. The latter seems to perform best and is used henceforth.

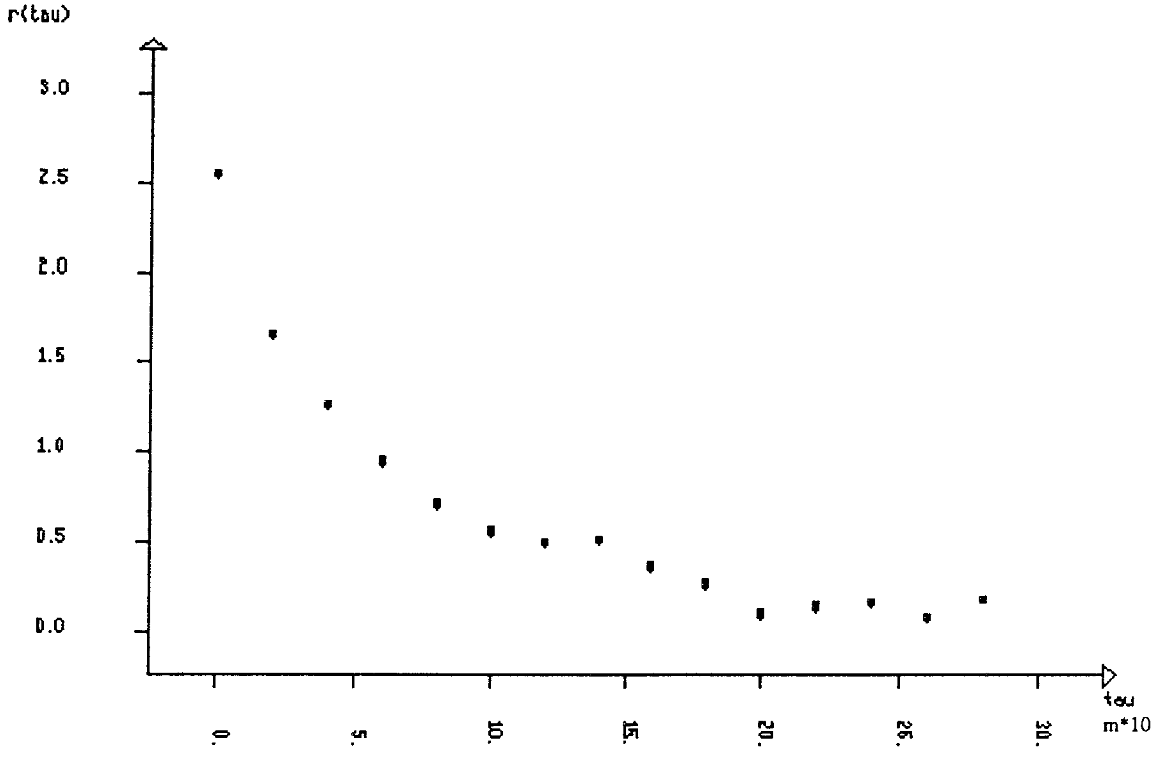


Figure 8, IGLSE/Classical, N=*1*

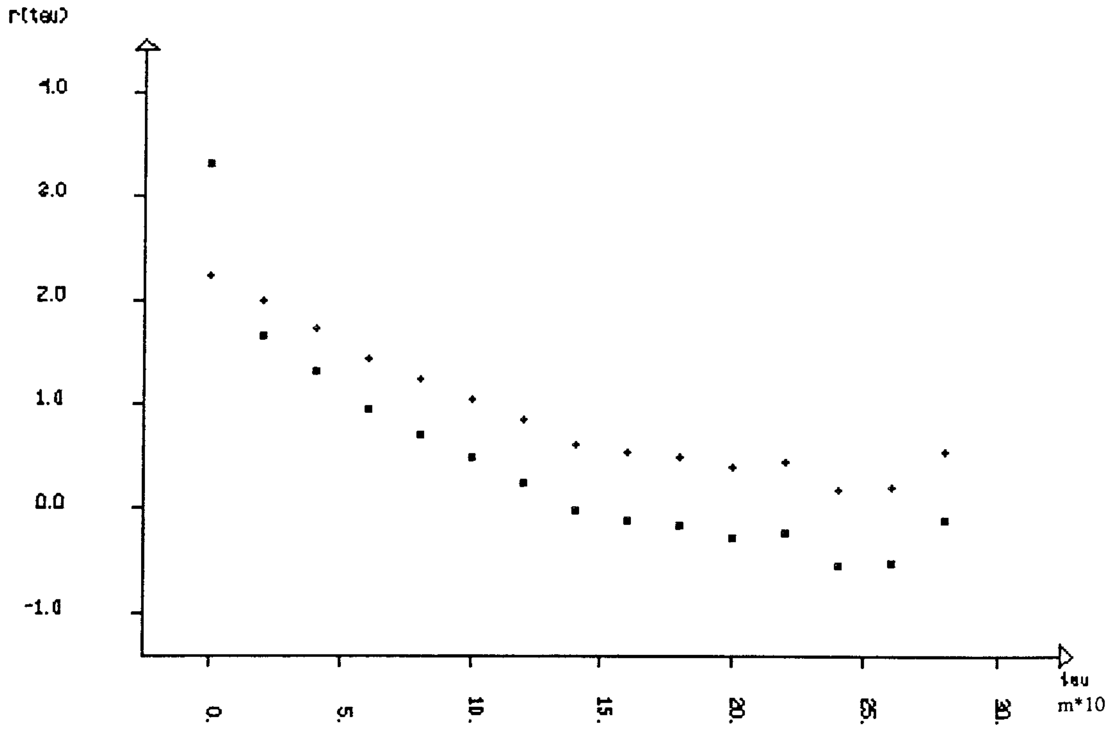


Figure 9 , MLE/Robinson, N=*1*

In Figure 10 and 11 the results of the classical and Robinson estimator are shown respectively when one trend is used per borehole (i.e. $N = 5$). Figure 10 looks very similar to Figure 8 except for the translation in the negative r–direction. The boxes in Figure 11 represent the result obtained if (by contrast to the former Robinson plots) the variance is multiplied by l_i/n_i according to (43). This version is used in the simulation which will be described in the next section. The covariance here, when a separate trend is used for each borehole, is lower than the one in Figure 9 as expected. The covariance function parameters $r(0)$, d_0 and d_{cut} are presented in Table 2 Section 5.3, the covariance radius is approximately 200 m.

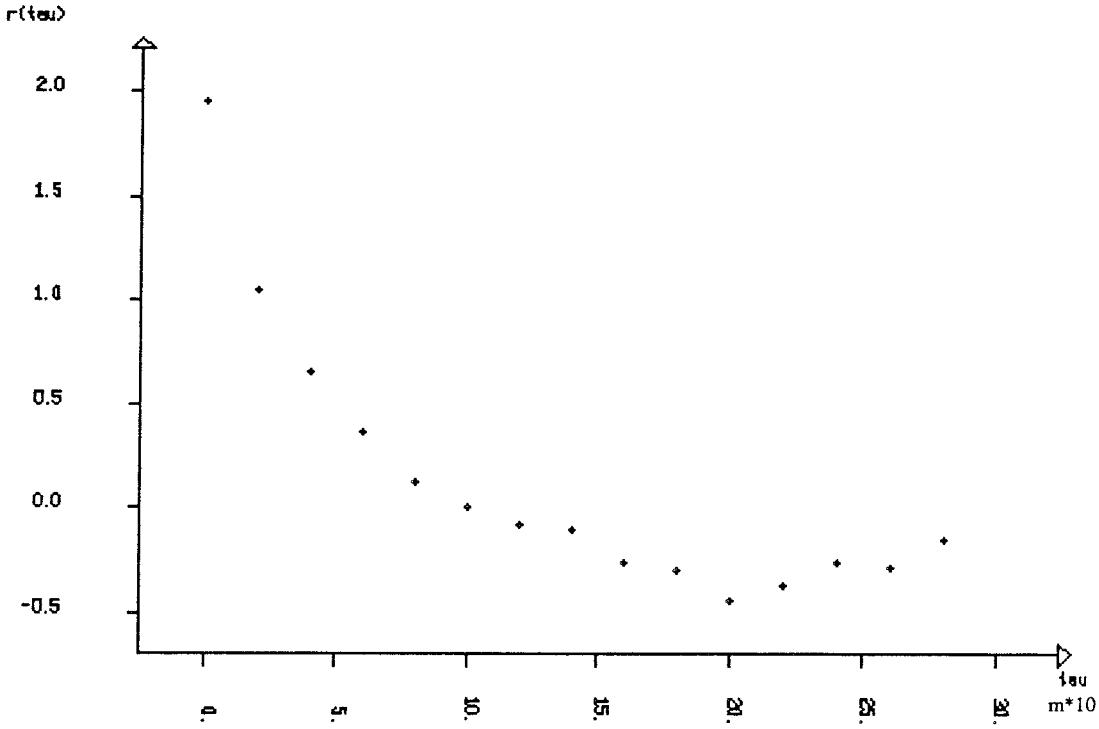


Figure 10, IGLSE/Classical N=5

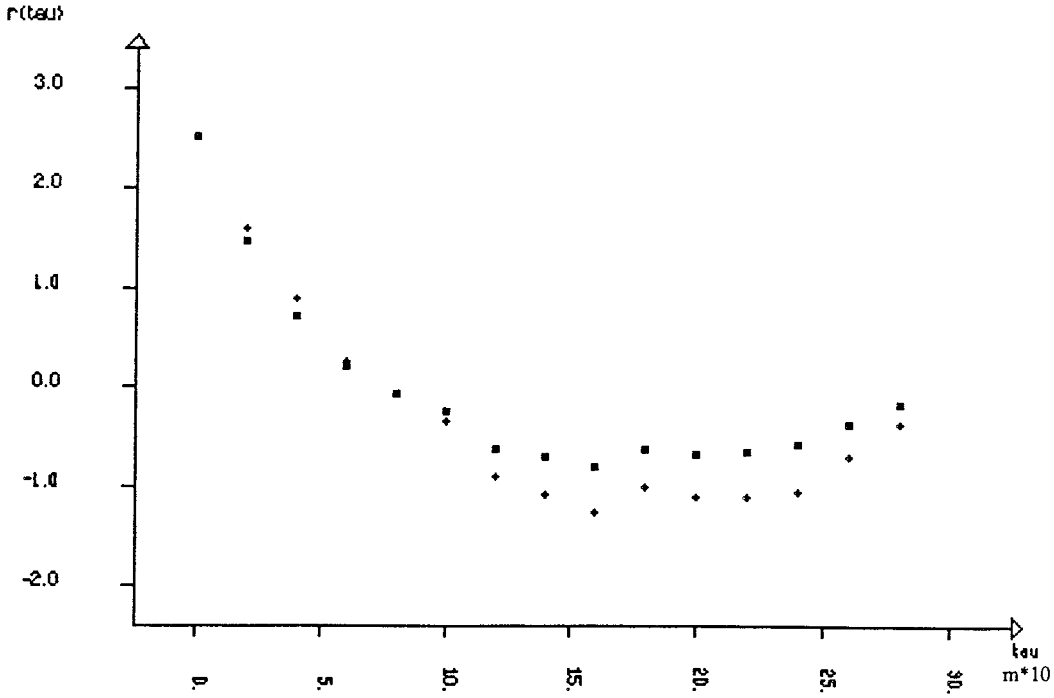


Figure 11, MLE/Robinson N=5

Figure 12 and 13 show the classical and the Robinson estimator with 9 trends respectively. Regarding the borehole plots in Section 3.1 one may assign borehole 1 and 5 two subtrends and borehole 3 three subtrends, as shown in Figure 14–16. (the trends illustrated in these figures are fitted by eye so they may not coincide with the actual outcome of the regression calculations). Here the variance has decreased further, as expected.

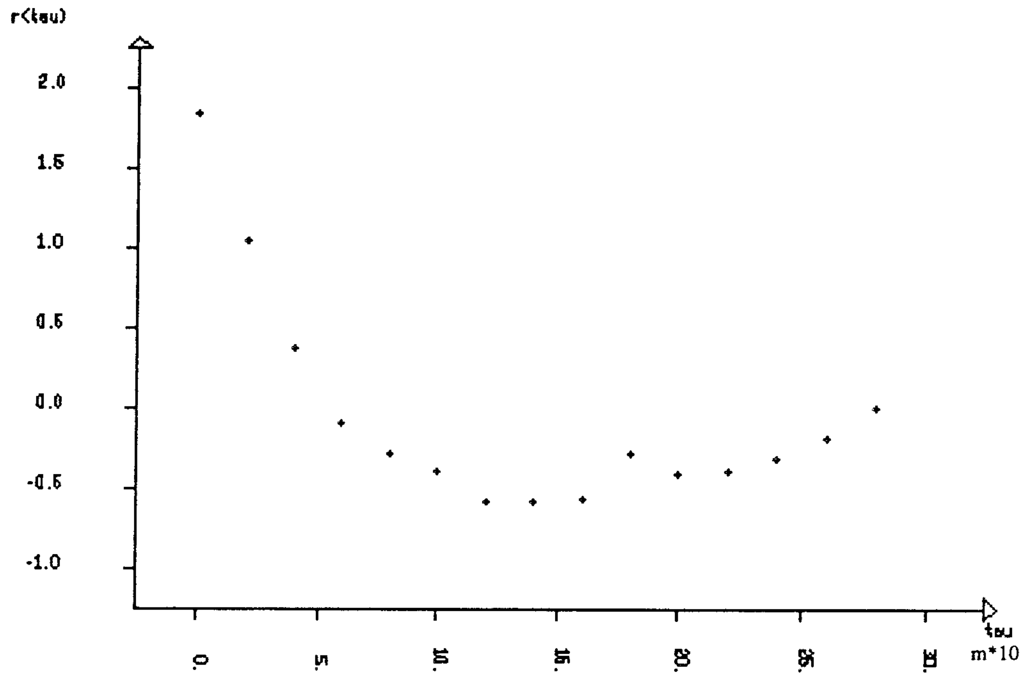


Figure 12, IGLSE/Classical N=9

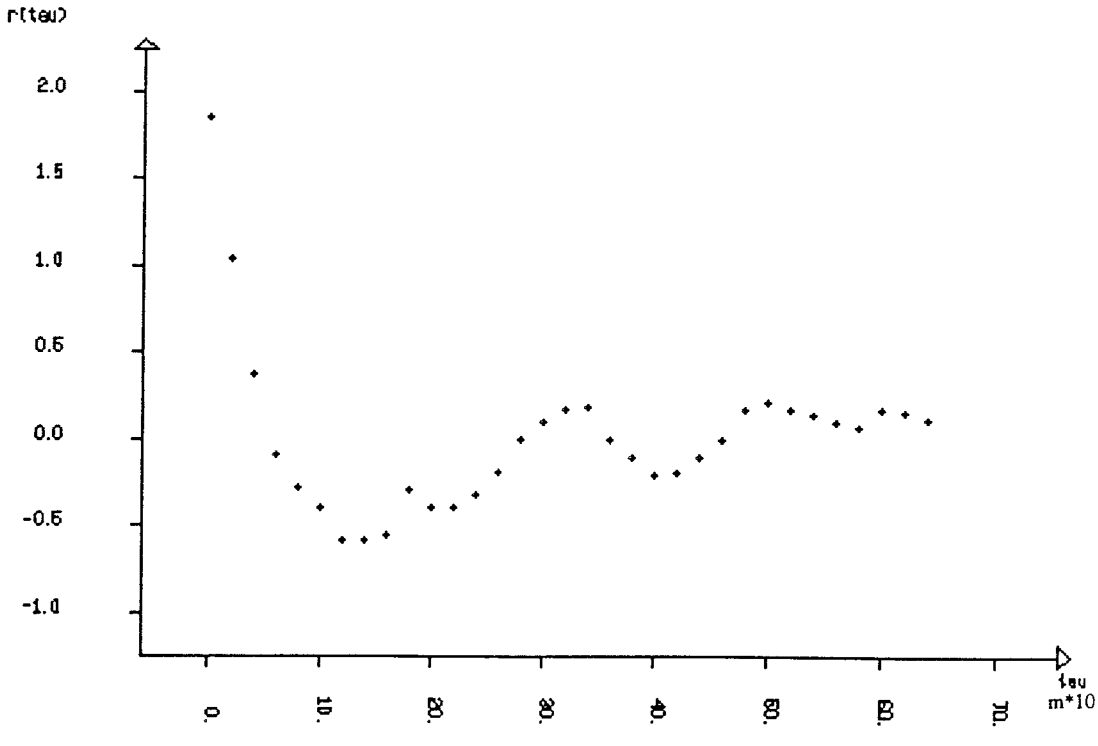


Figure 13, MLE/Robinson N=9

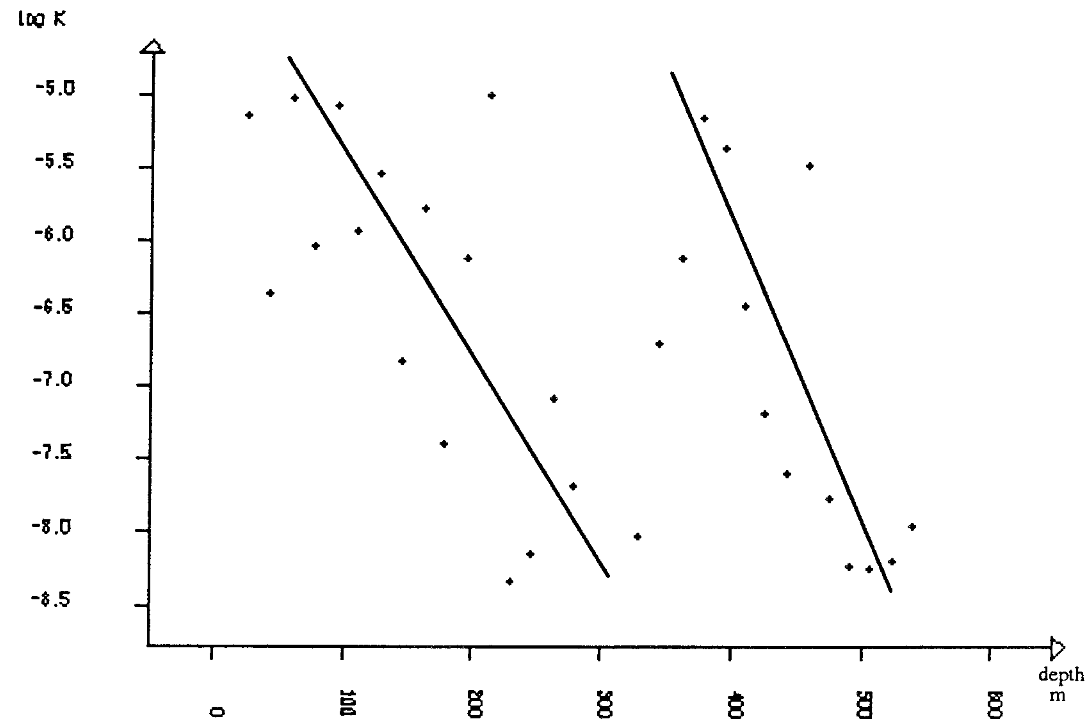


Figure 14, Borehole1 two trends fitted.

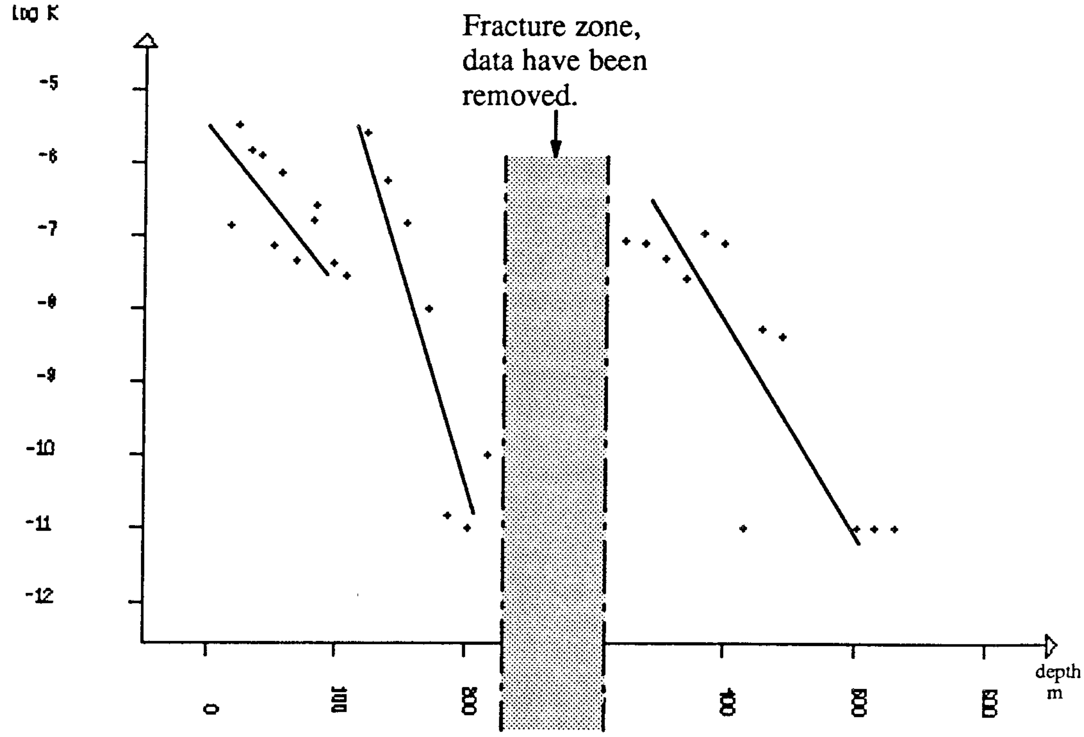


Figure 15, Borehole3 three trends fitted.

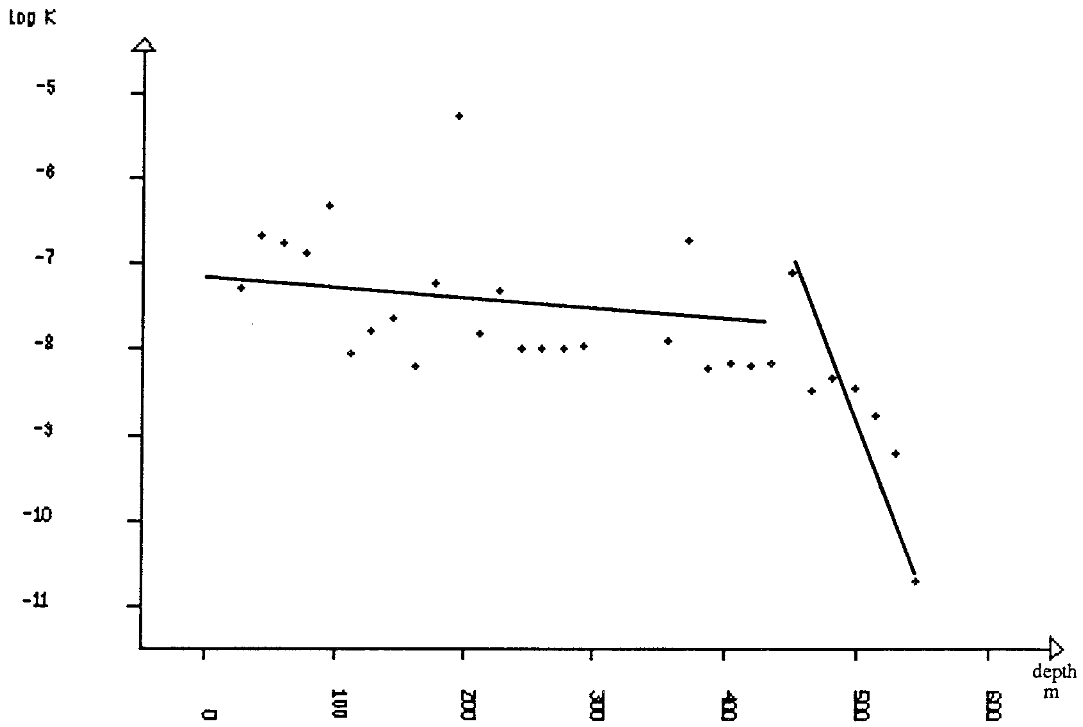


Figure 16, Borehole5 two trends fitted.

Most of the estimators have a "dip" for lags around 200m indicating an illusory negative correlation. This is probably due the error in modelling the trend of $\log K$ as a linear function of z . In a test including the term $\beta_2^i \cdot (z^i)^2$ in the regression expression for $i = 1$ the use of the MLE/Robinson estimator (Figure 17) gives no "dip". The term β_2^i becomes positive and when extrapolated to the depth of 530m $\log K$ starts to increase with increasing depth which is not reasonable. This example highlights a problem. In the FSCF10 model we are modelling a rock mass that extends 3000m below the ground surface based on conductivity data down to about 650m. The lack of information forces us to extrapolations which may be a poor model of the nature of the phenomenon.

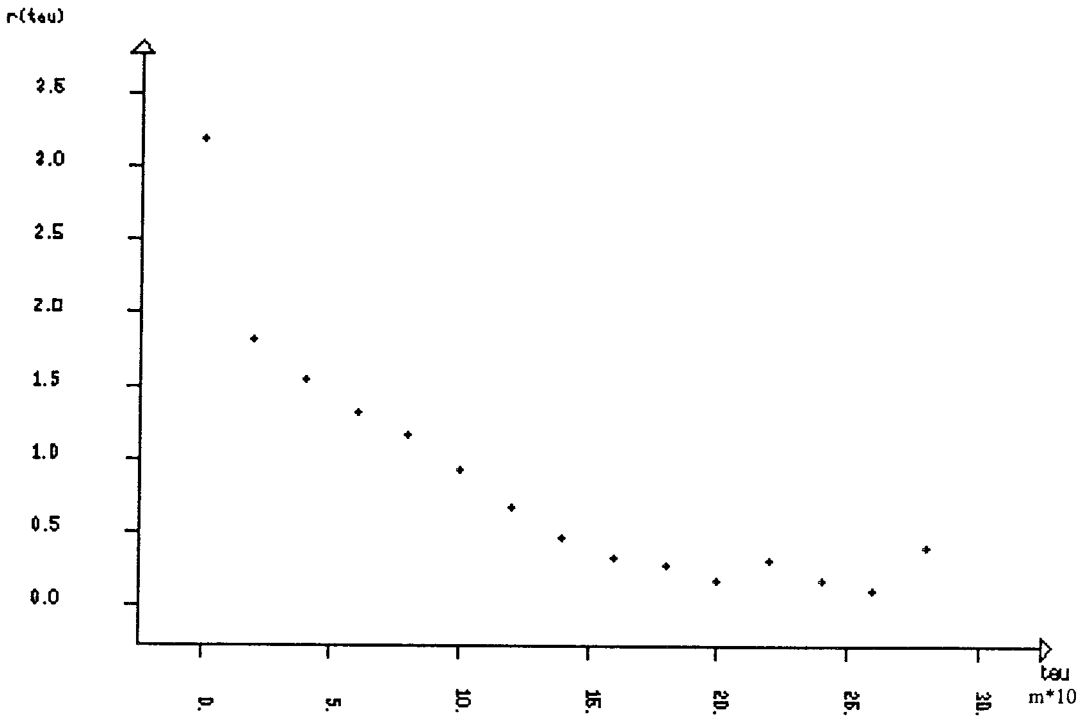


Figure 17

5.3 Drilling in the Model

In this section we present some sample boreholes "drilled" in the FSCF10–model. The conductivity fields are generated according to the procedure described above using the results of the two estimator concepts. A single realization is used. The size of the FSCF10 model is smaller than the average distance between the boreholes at the Klipperås study site (Ref.4) where the data were collected. What we are modelling (in generic terms) is therefore an area corresponding to one hole rather than a region containing all the holes. Hence, we use the estimators based on one trend per borehole (Figure 10 and 11 above). The parameter d_0 is then estimated using (30). The same vector of independent normal variates was used to generate the conductivity field in both cases. Table 2 shows the parameters used in either case. To get typical values of β_0 and β_1, we use the estimates from the two regression techniques MLE and IGLSE assuming one common trend only.

	IGLSE/– classical	MLE/– Robinson
β_0	–6.568	–6.470
β_1	$5.018*10^{-3}$	$5.925*10^{-3}$
$r(0)$	2.010	2.504
d_0	76.88	69.85
d_{cut}	900	900

Table 2

Figure 18 and 19 show the resulting values of $\log K$ for the two element columns far to the left in the model based on both the IGLSE/Classical and MLE/Robinson estimators. What we see is the conductivity calculated in the centre of the of the mesh element (or subelements for large elements c.f Subsection 4.2.2.) Figure 20 and 21 similarly show the values for column 3 and 6 (from the left of the model). Comparing with the source data (Figures 4–6) these simulations look quite alike.

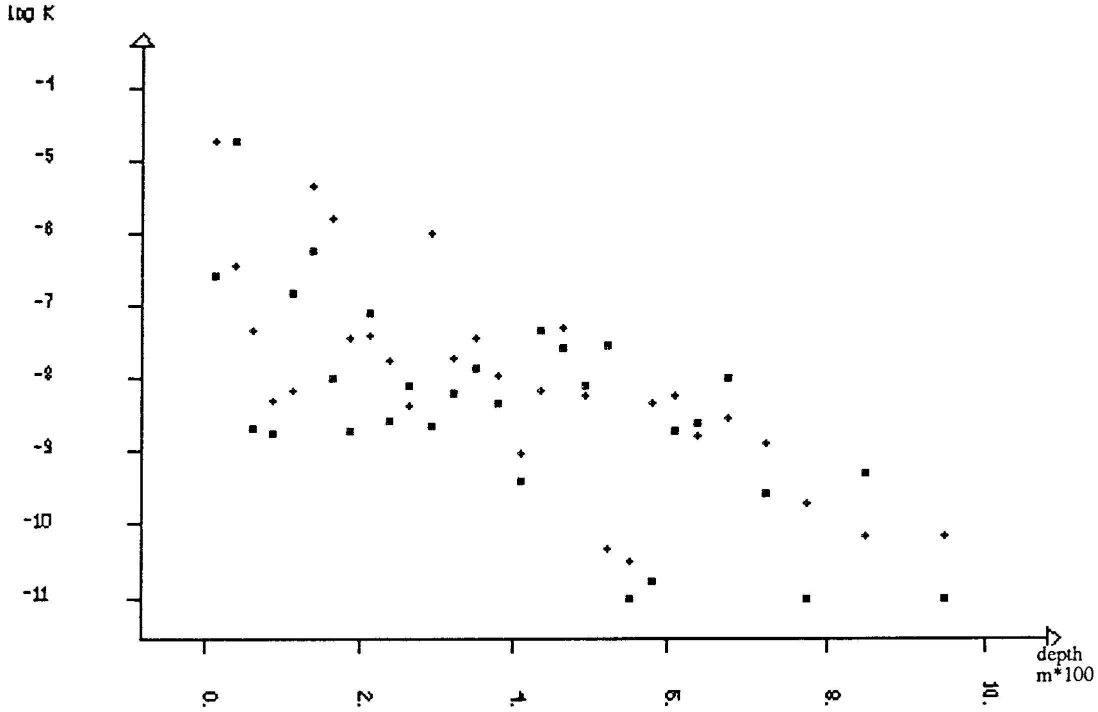


Figure 18, col. 1&2 IGLSE/Classical

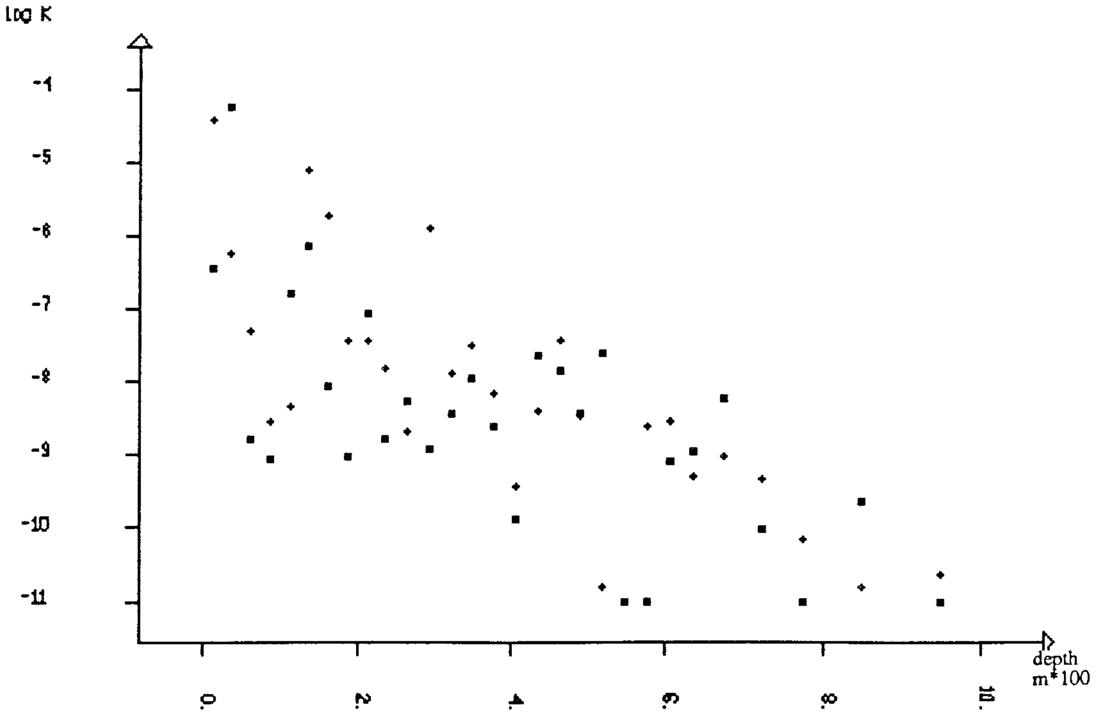


Figure 19, col. 1 & 2, MLE/Robinson

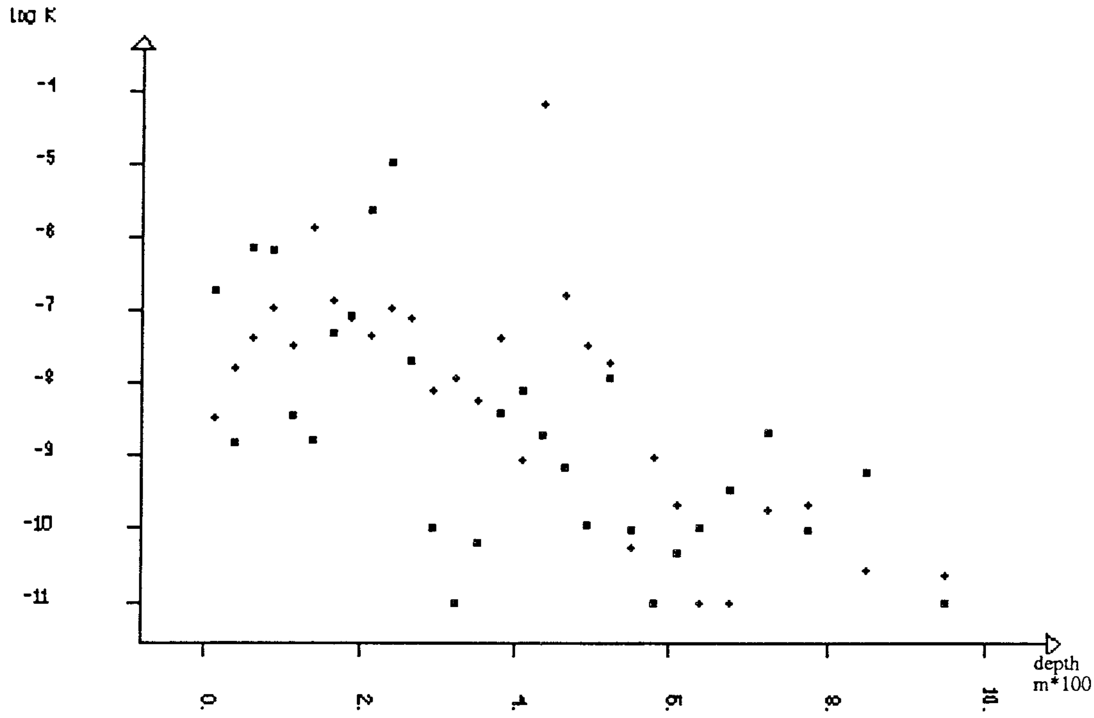


Figure 20, col. 3 & 6, IGLSE/Classical

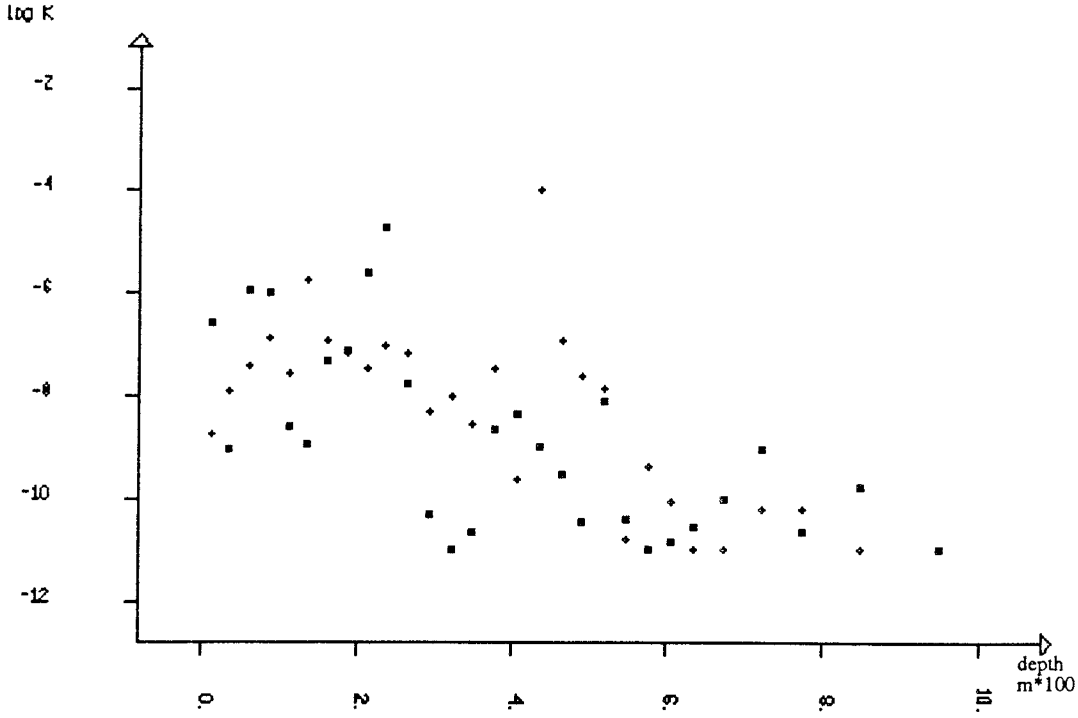


Figure 21, col. 3 & 6 MLE/Robinson

5.4 The Performance Measures

Figure 22–25 show the distributions of the performance measure discussed in Section 2.4 given the uncertainty induced by the spatial variability of the hydraulic conductivity. The dotted and plain lines represent results obtained using the MLE/Robinson and the IGLSE/Classical estimator respectively. The simulation based on the MLE/Robinson estimation shows a greater variability which also is expected due the greater value on the estimated variance (c.f Figure 10 and 9). The variability in the performance measures associated with the farfield is large (Figure 24 and 25) compared with the ditto obtained in a preliminary study where the conductivity field was modelled by $\log K = AFX + BCF \log z$ and values of the parameters AFX and BCF were generated in each realization and used throughout the model which gives the field an uniform behaviour. In this present model there is no such overall structure in the field. Besides the short ranged covariance coupling the conductivities are generated individually in each FSCF10 element.

Another, at a first glance, surprising result is that no significant correlation is detected between the annual recharge and any of the performance measures, when a sensitivity analysis is performed in the PROPER postprocessor POSTREG (c.f. also the POSTMON/GPLOT scatterplots in Figure 26). In the preliminary study a strong dependency was found. The lack of uniform behaviour in this model combined with the exponential nature by which K is related to z (see (15) and Section 4.1) give a negligible contribution to the recharge from elements deep down where the repository is located. Thus, the flow velocity deep down in the model does not have to be strongly related to the recharge.

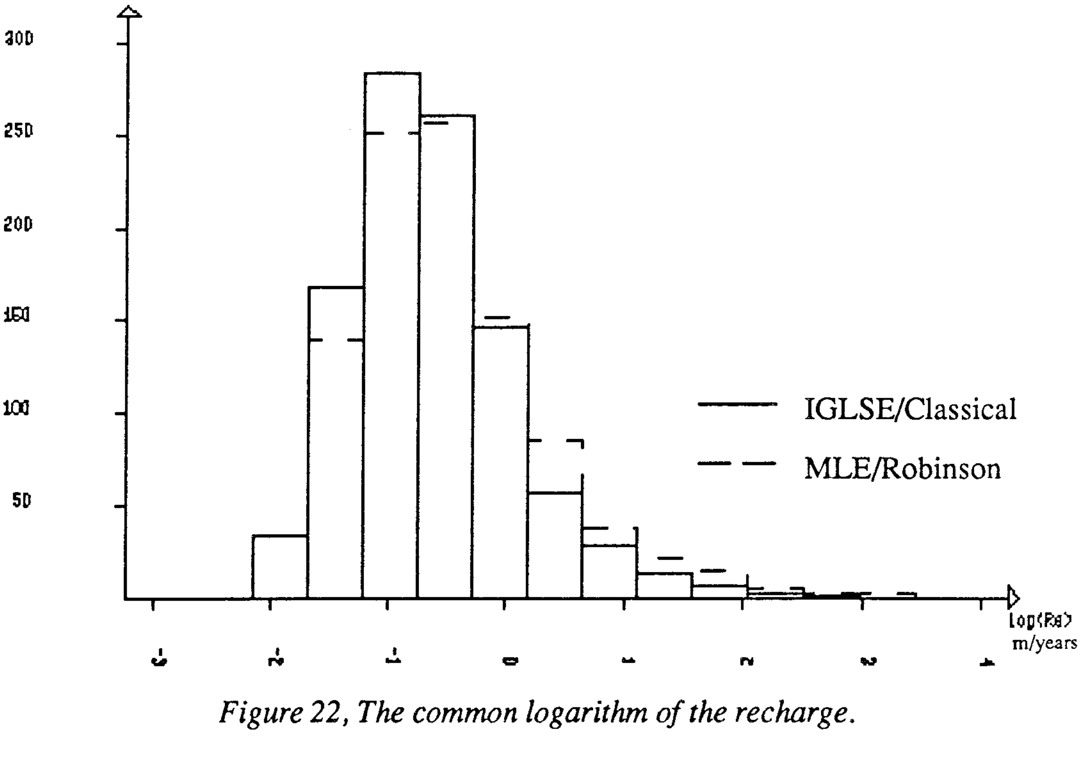


Figure 22, The common logarithm of the recharge.

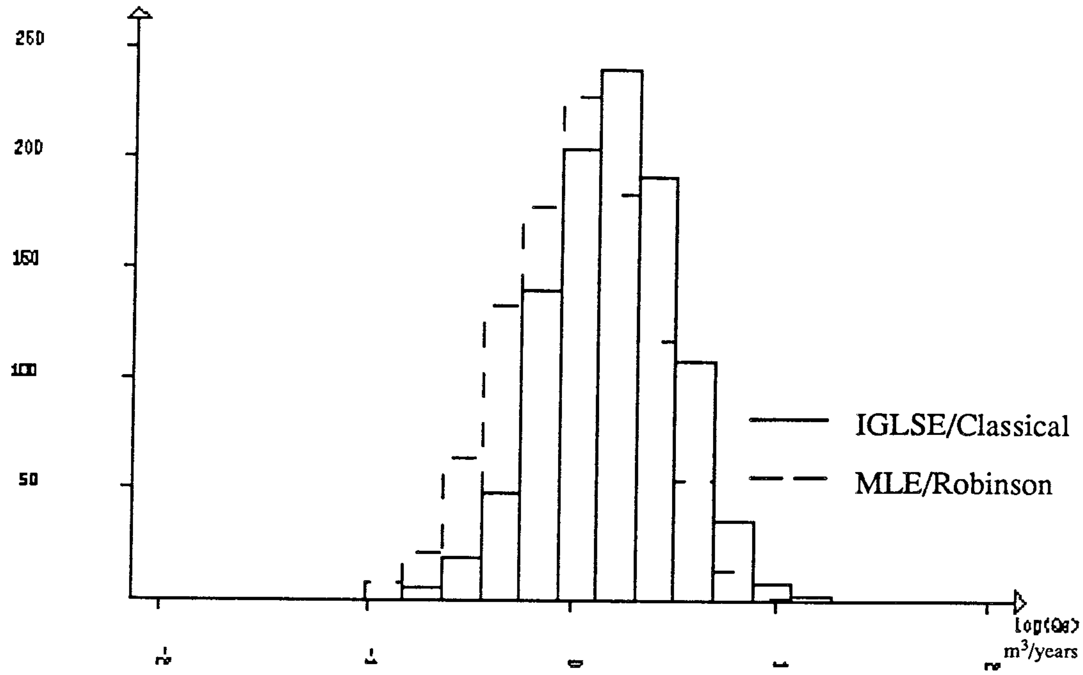


Figure 23, The common logarithm of Qeq.

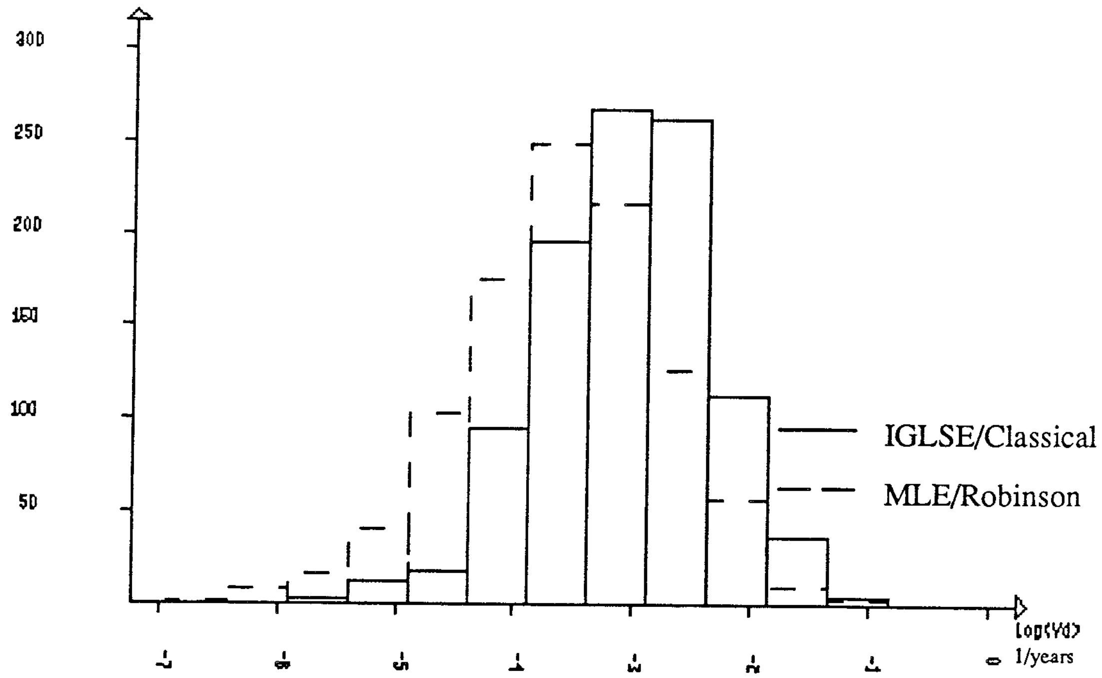


Figure 24, The common logarithm of Vd

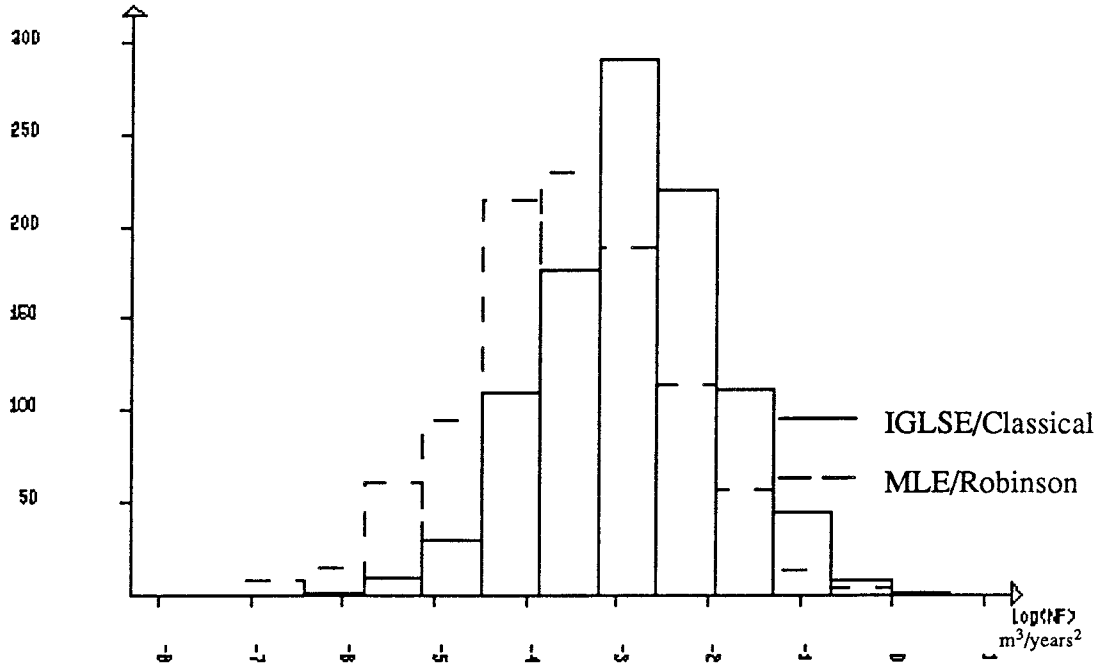


Figure 25, The common logarithm of NFFF

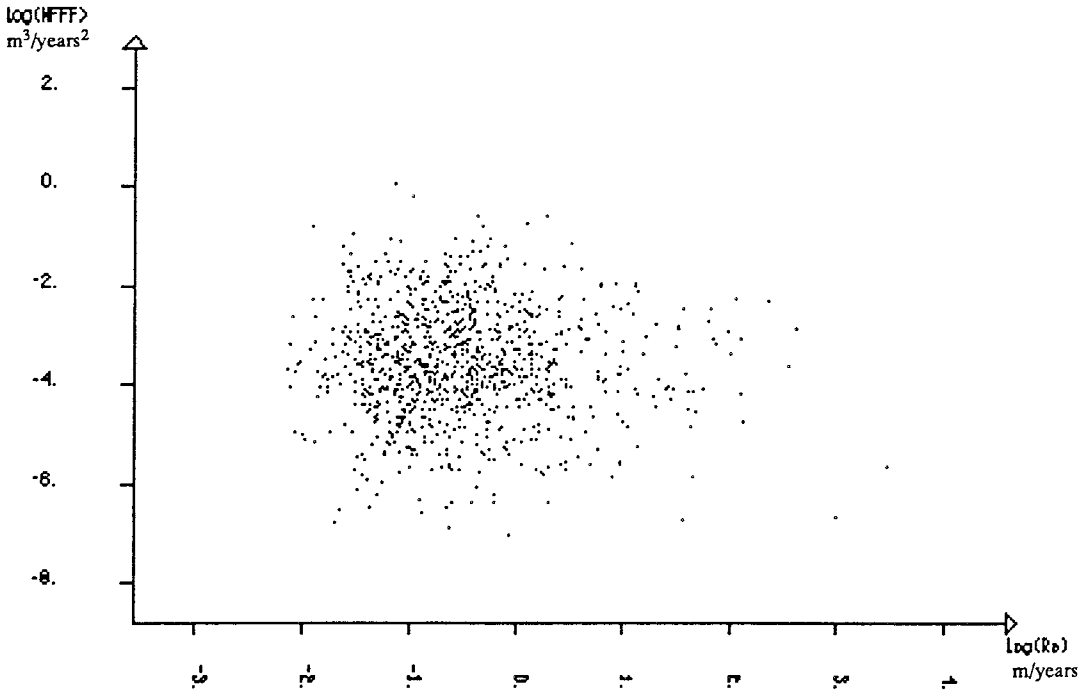


Figure 26 log(NFFF) v.s. log(recharge)

6 CONCLUSIONS AND CAVEATS

The present study certainly has its shortcomings, such as not dealing strictly with the 3–D problem, restrictions to continuous media etc. but it could never the less be concluded that spatial variability in the hydraulic conductivity is an important source of uncertainty in repository performance predictions.

The validity of this conclusion has to do with the fact that the scale of the variability is comparable to the distance between the outermost part of the repository and "the accessible environment", the latter represented by vertical fracture zones, assuming the delay in a zone is short. This is reflected in the large uncertainties in the performance measures involving the farfield. The actual extent of the repository as such is sufficient to make the variabilities average out, reflected in total Q_{eq}, were it not for the proximity of the vertical zones.

Some additional conclusions can be drawn from the study:

- hydrologic stochastic simulation is compatible with the PSAC approach to uncertainty analysis,
- it would be desirable to go to strict three–dimensionality since the variability is 3–D in nature,
- refinement and statistical analysis of the estimation procedures are desirable if the approach of the study is to be pursued,
- the influence of covariances between the residuals seem to have a small effect on the trend estimation.

7 REFERENCES

1 NERETNIEKS, I
Transport of Oxidants and Radionuclides through a Clay Barrier.
KBS Technical Report 79, Project Kärnbränslesäkerhet,
February 1978.

2 ANDERSSON, G, RASMUSON, A, and NERETNIEKS, I,
Migration Model for the Near Field: Final Report.
KBS Technical Report 82–24, Swedish Nuclear Fuel Supply Co,
November 1982.

3 HIGBIE, R,
The Rate of Absorption of a Pure Gas into a Still Liquid during Short Periods of Exposure. Trans AIChE 31, 365–389, 1935.

4 GENTZSCHEIN, B
Hydrogeological Investigations at the Klipperås Study Site.
SKB Technical Report 86–08, Swedish Nuclear Fuel and Waste Management Co, June 1986.

5 DRAPER, N and SMITH, H,
Applied Regression Analysis,Second Edition. John Wiley & Sons, 1978.

6 NEUMAN, S P and JACOBSON, E A,
Analysis of Nonintrinsic Spatial Variability by Residual Kriging with Application to Regional Groundwater Levels. Mathematical Geology, Vol. 16, No 5, 1984

7 LAWLESS, J F
Statistical Models and Methods for Lifetime Data. Wiley's Series on Probability Analysis, 1982.

8 HELSEL, D R and COHN, T A
Estimation of Descriptive Statistics for Multiply Censored Water Quality Data. Water Resources Research, vol. 24, No 12, 1997–2004, December 1988.

9 GLEIT, A
Estimation from Small Normal Data Sets with Detection Limits.
Environ. Sci. Technol., 19, 1201–1206, 1985.

10 SHENTON, L R and BOWMAN, K O
Maximum Likelihood Estimation in Small Samples. Griffin, London 1977.

11 PARZEN, E
On Spectral Analysis with Missing Observations and Amplitude Modulation. Sankhyã, Series A 25, 383–392, 1963.

12 JENKINS, G M and WATTS, D G,
Spectral Analysis and it's Applications. Holden–Day,
San Francisco 1969.

13 DUNSMUIR, W
Large Sample Properties of Estimation in Time Series Observed at Unequally Spaced Times. Ed. BRILLINGER, D, FIENBER, S, GANI, J, HERTIGAN, J, and KRICKENBERG, K, Time Series Analysis of Irregularly Observed Data, Springer–Verlag 1984.

14 ROBINSON, P M
Estimation and Forecasting for Time Series Containing Censored or Missing Observations. Ed. ANDERSSON, O D, Time Series, North–Holland Publishing Company 1980.

15 BRATLEY, P, FOX, B L and SCHRAGE, L E,
A Guide to Simulation, Second Edition, Springer–Verlag.

16 BRAESTER, C, and THUNVIK, R,
Numerical Simulation of Double Packer Tests, Calculation of Rock Permeability. KBS Technical Report 82–06, Swedish Nuclear Fuel Supply Co, June 1982.

17 GENTZSCHEIN, B
Description of Hydrogeological Data in SKB's Data Base GEOTAB. SKB Technical Report 86–22, Swedish Nuclear and Fuel Management Co, December 1986.

APPENDIX 1

Extracting conductivity data from geotab

The Hydrogeological Database GEOTAB

SKB and SGAB have over the years collected huge amounts of data from different study sites all over Sweden. These data has been stored in SKB's database GEOTAB (Ref.17). GEOTAB is a relational database and it is based on a program from Mimer Information System. Among the data in GEOTAB it is the conductivities we are interested in.

Using Tables in GEOTAB

Data on the conductivity in the bedrock are found in the table SHTINJCD. In this study data from the Klipperås study site have been used. Conductivity measures are available for the following boreholes: KKL01, KKL02, KKL06, KKL09, KKL12, KKL13 and KKL14, but data from the two first in this list do not participate in our calculations. In KKL01 the conductivity is measured in sections of 25 m while a section length of 20 m is used for the others, so the use of KKL01 would disrupt the 20 m regularity that is utilized in the covariance estimations. KKL02 data are excluded because this hole has an angle of inclination that deviates much from those of the others. Furthermore most of the conductivity values in this hole are found to lie below the measurement limit (indicated as $K = -99$ m/s) making the estimates less reliable. Information on this limit is stored in the table SHTINJF2. The remaining five boreholes KKL06–KKL14 are referred to as borehole 1–5 in this report.

In the study we model rock mass, hence data corresponding to major fracture zones are removed and are looked upon as missing values (c.f Chapter 3). Those zones are located by other means that conductivity measurements (e.g. surface geophysical measurements see Ref.4).

The depth corresponding to some conductivity value is obtained by interpolation between the tables SHTINJCD and BHCOORD where the conductivity and coordinates respectively are stored as a functions of length along the hole. Finally the depth coordinate is interpolated to the section centres. In GEOTAB the coordinates are given for the upper end of the sections, but here it was assumed that the measured conductivity values correspond to the centres of the sections.

APPENDIX 2

The estimator test

In order to see how the estimator chain MLE/Robinson performs we have generated a number of synthetic data sets each consisting of three series of values for $\log K$, using PROPER's random number generator given the distribution function. Every set has a unique collection of properties (such as number of points, covariance structure in the parent distribution etc.) but they are based on the same random seed.

When generating these synthetic data sets we assume the regression parameters β_0 and β_1 to be equal to zero. The number of values is set to either 35 or 100. They appear regularly on data points every 20m with the first point at the depth 100 m . $\log K$ is assumed to be normally distributed with unit variance and an exponential covariance function according to (30). Three different values for the covariance of the first lag are used 0, 0.2 and 0.5. Note that the regression parameters and covariances are used as input for the the data generation process (which is based on the ideas presented in Section 4.1) and should not be confused with the actual outcome of the estimation tabled. The censoring level is set to –0.43 which causes about 1/3 of the population generated to be censored. Estimation with missing values have been tested. We have also tested Robinson estimators with other values than one on the integers r and s in (40), "`delta10`" means the estimator with r=1 and s=0.

The Robinson covariance is estimated both separately based on ($\beta_0 = \beta_1 = 0, \sigma = 1$) and after that the trend obtained from a MLE–regression is removed. In the list that follows the label "`ROBINSON VARIANCE:`" refers to the result obtained from the variance estimator (38). The value of the variance used in the covariance estimation is that supplied by the MLE which corresponds to "`COVARIANCE`" for "`DISTANCE`" = 0 metres.

```
 Regression parameter 1 {B(0)}:
0
 Regression parameter 2 {B(1)}:
0
 Number of boreholes :
3
 Number of desired points/borehole :
100
 The variance :
1
 The covariance for the first lag (as a fraction of the variance):
0
 The censoring level (in log(conductivity))  ;
-.43

** without regression **

BOREHOLE =  1
REGRESSIONS PARAMETERS:   0.0000       0.0000
NUMBER OF CENSORED:  30
NUMBER OF POINTS: 100
ROBINSON VARIANCE:     1.025

 DISTANCE(m) COVARIANCE
  0  1.000000000000
 20  0.1536191000000E-02
 40 -0.2265261000000
 60 -0.1987131000000
 80 -0.9012343000000E-02

BOREHOLE =  2
REGRESSIONS PARAMETERS:   0.0000       0.0000
NUMBER OF CENSORED:  36
NUMBER OF POINTS: 100
ROBINSON VARIANCE:     1.178

 DISTANCE(m) COVARIANCE
  0  1.000000000000
 20  0.1084925000000
 40  0.2691611000000
 60 -0.5474831000000E-01
 80  0.9913125000000E-01

BOREHOLE =  3
REGRESSIONS PARAMETERS:   0.0000       0.0000
NUMBER OF CENSORED:  40
NUMBER OF POINTS: 100
ROBINSON VARIANCE:    0.6589

 DISTANCE(m) COVARIANCE
  0  1.000000000000
 20 -0.2156211000000
 40 -0.4241783000000
 60 -0.4750810000000
 80 -0.4326651000000

** with regression **

BOREHOLE =  1
REGRESSIONS PARAMETERS:   0.2165      -0.1868E-03
NUMBER OF CENSORED:  30
NUMBER OF POINTS: 100
ROBINSON VARIANCE:     1.035

 DISTANCE(m) COVARIANCE
  0  0.8994565000000
 20  0.1386692000000E-01
 40 -0.1452411000000
 60 -0.1211467000000
 80  0.2482342000000E-01
```

```
 BOREHOLE =  2
 REGRESSIONS PARAMETERS:   0.3580E-01  -0.7123E-04
 NUMBER OF CENSORED:  36
 NUMBER OF POINTS: 100
 ROBINSON VARIANCE:     1.254

  DISTANCE(m) COVARIANCE
   0  1.183830000000
  20  0.7598563000000E-01
  40  0.2329044000000
  60 -0.1194265000000
  80  0.7209144000000E-01

 BOREHOLE =  3
 REGRESSIONS PARAMETERS:  -0.2468E-01  -0.1247E-03
 NUMBER OF CENSORED:  40
 NUMBER OF POINTS: 100
 ROBINSON VARIANCE:    0.7913

  DISTANCE(m) COVARIANCE
   0  0.8510776000000
  20  0.1225850000000
  40 -0.8780709000000E-01
  60 -0.1627259000000
  80 -0.1012216000000

 ***************************************

  Regression parameter 1 {B(0)}:
 0
  Regression parameter 2 {B(1)}:
 0
  Number of boreholes :
 3
  Number of desired points/borehole :
 100
  The variance :
 1
  The covariance for the first lag (as a fraction of the variance):
 .2
  The censoring level (in log(conductivity))  ;
 -.43

 ** without regression **

 BOREHOLE =  1
 REGRESSIONS PARAMETERS:   0.0000       0.0000
 NUMBER OF CENSORED:  29
 NUMBER OF POINTS: 100
 ROBINSON VARIANCE:    0.9374

  DISTANCE(m) COVARIANCE
   0  1.000000000000
  20  0.2153775000000
  40 -0.9652350000000E-01
  60 -0.1135945000000
  80  0.2437166000000E-01

 BOREHOLE =  2
 REGRESSIONS PARAMETERS:   0.0000       0.0000
 NUMBER OF CENSORED:  35
 NUMBER OF POINTS: 100
 ROBINSON VARIANCE:     1.184

  DISTANCE(m) COVARIANCE
   0  1.000000000000
  20  0.2504759000000
  40  0.2133265000000
  60 -0.6936812000000E-01
  80  0.3522971000000E-01

 BOREHOLE =  3
```

```
REGRESSIONS PARAMETERS:   0.0000       0.0000
NUMBER OF CENSORED:  38
NUMBER OF POINTS: 100
ROBINSON VARIANCE:    0.7345

 DISTANCE(m) COVARIANCE
  0  1.000000000000
 20 -0.1077582000000
 40 -0.4480524000000
 60 -0.5696497000000
 80 -0.4583662000000

** with regression **

BOREHOLE =  1
REGRESSIONS PARAMETERS:   0.3437      -0.2868E-03
NUMBER OF CENSORED:  29
NUMBER OF POINTS: 100
ROBINSON VARIANCE:    0.9588

 DISTANCE(m) COVARIANCE
  0  0.8503359000000
 20  0.2044345000000
 40 -0.1205027000000E-01
 60 -0.1710557000000E-01
 80  0.6451920000000E-01

BOREHOLE =  2
REGRESSIONS PARAMETERS:   0.8443E-01  -0.1141E-03
NUMBER OF CENSORED:  35
NUMBER OF POINTS: 100
ROBINSON VARIANCE:     1.207

 DISTANCE(m) COVARIANCE
  0  1.162153000000
 20  0.2569603000000
 40  0.1845136000000
 60 -0.1215457000000
 80  0.1403328000000E-01

BOREHOLE =  3
REGRESSIONS PARAMETERS:   0.9048E-02  -0.1437E-03
NUMBER OF CENSORED:  38
NUMBER OF POINTS: 100
ROBINSON VARIANCE:    0.7982

 DISTANCE(m) COVARIANCE
  0  0.7993429000000
 20  0.2344402000000
 40 -0.1160718000000
 60 -0.2328710000000
 80 -0.1256772000000

***********************************

 Regression parameter 1 {B(0)}:
0
 Regression parameter 2 {B(1)}:
0
 Number of boreholes :
3
 Number of desired points/borehole :
100
 The variance :
1
 The covariance for the first lag (as a fraction of the variance):
.5
 The censoring level (in log(conductivity))  ;
-.43

** without regression **
```

```
BOREHOLE =  1
REGRESSIONS PARAMETERS:   0.0000       0.0000
NUMBER OF CENSORED:  29
NUMBER OF POINTS: 100
ROBINSON VARIANCE:     1.089

 DISTANCE(m) COVARIANCE
  0  1.000000000000
 20  0.6263550000000
 40  0.2810170000000
 60  0.1699154000000
 80  0.1767798000000

BOREHOLE =  2
REGRESSIONS PARAMETERS:   0.0000       0.0000
NUMBER OF CENSORED:  39
NUMBER OF POINTS: 100
ROBINSON VARIANCE:     1.077

 DISTANCE(m) COVARIANCE
  0  1.000000000000
 20  0.3855616000000
 40  0.1495973000000
 60 -0.6146874000000E-01
 80 -0.6718557000000E-01

BOREHOLE =  3
REGRESSIONS PARAMETERS:   0.0000       0.0000
NUMBER OF CENSORED:  38
NUMBER OF POINTS: 100
ROBINSON VARIANCE:    0.6896

 DISTANCE(m) COVARIANCE
  0  1.000000000000
 20 -0.2954870000000E-01
 40 -0.4429040000000
 60 -0.6207560000000
 80 -0.5190855000000


** with regression **

BOREHOLE =  1
REGRESSIONS PARAMETERS:   0.5391      -0.4592E-03
NUMBER OF CENSORED:  29
NUMBER OF POINTS: 100
ROBINSON VARIANCE:    0.9438

 DISTANCE(m) COVARIANCE
  0  0.8423051000000
 20  0.5503365000000
 40  0.3225990000000
 60  0.2517350000000
 80  0.1990486000000

BOREHOLE =  2
REGRESSIONS PARAMETERS:   0.1128      -0.1905E-03
NUMBER OF CENSORED:  39
NUMBER OF POINTS: 100
ROBINSON VARIANCE:     1.128

 DISTANCE(m) COVARIANCE
  0  1.211107000000
 20  0.4831440000000
 40  0.1757441000000
 60 -0.5289333000000E-01
 80 -0.3623273000000E-01

BOREHOLE =  3
REGRESSIONS PARAMETERS:  -0.8835E-01  -0.1001E-03
NUMBER OF CENSORED:  38
NUMBER OF POINTS: 100
```

```
ROBINSON VARIANCE:    0.8333

 DISTANCE(m) COVARIANCE
  0  0.7705071000000
 20  0.3722377000000
 40 -0.3640557000000E-01
 60 -0.2183599000000
 80 -0.9604086000000E-01

**********************************

 Regression parameter 1 {B(0)}:
0
 Regression parameter 2 {B(1)}:
0
 Number of boreholes :
3
 Number of desired points/borehole :
35
 The variance :
1
 The covariance for the first lag (as a fraction of the variance):
0
 The censoring level (in log(conductivity))  ;
-.43

** without regression **

BOREHOLE =  1
REGRESSIONS PARAMETERS:   0.0000       0.0000
NUMBER OF CENSORED:  10
NUMBER OF POINTS:  35
ROBINSON VARIANCE:     1.067

 DISTANCE(m) COVARIANCE
  0  1.000000000000
 20  0.2226459000000
 40 -0.4959803000000
 60 -0.4126567000000
 80 -0.1987596000000

BOREHOLE =  2
REGRESSIONS PARAMETERS:   0.0000       0.0000
NUMBER OF CENSORED:  10
NUMBER OF POINTS:  35
ROBINSON VARIANCE:     1.426

 DISTANCE(m) COVARIANCE
  0  1.000000000000
 20  0.2848449000000
 40  0.1913951000000
 60  0.2424496000000
 80  0.5499847000000

BOREHOLE =  3
REGRESSIONS PARAMETERS:   0.0000       0.0000
NUMBER OF CENSORED:  13
NUMBER OF POINTS:  35
ROBINSON VARIANCE:    0.3910

 DISTANCE(m) COVARIANCE
  0  1.000000000000
 20 -0.8487454000000
 40 -0.6058108000000
 60 -0.7112620000000
 80 -0.6069178000000

** with regression **

BOREHOLE =  1
REGRESSIONS PARAMETERS:   0.4340      -0.8059E-03
NUMBER OF CENSORED:  10
```

```
 NUMBER OF POINTS:  35
 ROBINSON VARIANCE:     0.9932

  DISTANCE(m) COVARIANCE
   0  0.9501093000000
  20  0.5851704000000E-01
  40 -0.5431473000000
  60 -0.4130934000000
  80 -0.3504275000000

 BOREHOLE =  2
 REGRESSIONS PARAMETERS:  -0.4588       0.1328E-02
 NUMBER OF CENSORED:  10
 NUMBER OF POINTS:  35
 ROBINSON VARIANCE:     1.082

  DISTANCE(m) COVARIANCE
   0  1.048457000000
  20 -0.2473393000000
  40 -0.2741064000000
  60 -0.2988296000000
  80  0.5762685000000E-01

 BOREHOLE =  3
 REGRESSIONS PARAMETERS:   0.3679E-01  -0.6349E-03
 NUMBER OF CENSORED:  11
 NUMBER OF POINTS:  32
 ROBINSON VARIANCE:     0.6089

  DISTANCE(m) COVARIANCE
   0  0.4557592000000
  20 -0.9161615000000E-01
  40  0.8431831000000E-01
  60  0.7608604000000E-01
  80  0.4701186000000E-01

 *******************************

  Regression parameter 1 {B(0)}:
 0
  Regression parameter 2 {B(1)}:
 0
  Number of boreholes :
 3
  Number of desired points/borehole :
 35
  The variance :
 1
  The covariance for the first lag (as a fraction of the variance):
 .2
  The censoring level (in log(conductivity))  ;
 -.43

 ** without regression **

 BOREHOLE =  1
 REGRESSIONS PARAMETERS:   0.0000       0.0000
 NUMBER OF CENSORED:   9
 NUMBER OF POINTS:  35
 ROBINSON VARIANCE:     1.073

  DISTANCE(m) COVARIANCE
   0  1.000000000000
  20  0.4234822000000
  40 -0.3344702000000
  60 -0.4199496000000
  80 -0.1662586000000


 BOREHOLE =  2
 REGRESSIONS PARAMETERS:   0.0000       0.0000
 NUMBER OF CENSORED:   8
```

NUMBER OF POINTS: 35
ROBINSON VARIANCE: 1.252

DISTANCE(m)	COVARIANCE
0	1.000000000000
20	0.6215946000000
40	0.4725727000000
60	0.4965321000000
80	0.6808305000000

BOREHOLE = 3
REGRESSIONS PARAMETERS: 0.0000 0.0000
NUMBER OF CENSORED: 16
NUMBER OF POINTS: 35
ROBINSON VARIANCE: 0.1895

DISTANCE(m)	COVARIANCE
0	1.000000000000
20	-0.8023465000000
40	-0.7407620000000
60	-0.8122340000000
80	-0.7866788000000

** with regression **

BOREHOLE = 1
REGRESSIONS PARAMETERS: 0.5837 -0.1032E-02
NUMBER OF CENSORED: 9
NUMBER OF POINTS: 35
ROBINSON VARIANCE: 1.032

DISTANCE(m)	COVARIANCE
0	0.9084604000000
20	0.1657536000000
40	-0.4637544000000
60	-0.4630260000000
80	-0.4001214000000

BOREHOLE = 2
REGRESSIONS PARAMETERS: -0.4264 0.1502E-02
NUMBER OF CENSORED: 8
NUMBER OF POINTS: 35
ROBINSON VARIANCE: 1.007

DISTANCE(m)	COVARIANCE
0	0.8655828000000
20	-0.2603658000000E-01
40	-0.2901833000000
60	-0.2378764000000
80	0.6795360000000E-02

BOREHOLE = 3
REGRESSIONS PARAMETERS: 0.5223E-01 -0.8616E-03
NUMBER OF CENSORED: 9
NUMBER OF POINTS: 23
ROBINSON VARIANCE: 0.3003

DISTANCE(m)	COVARIANCE
0	0.4163807000000
20	-0.1559293000000
40	-0.2168923000000
60	-0.4967748000000E-01
80	-0.2598356000000

****

Regression parameter 1 {B(0)}:

```
0
 Regression parameter 2 (B(1)):
0
 Number of boreholes :
3
 Number of desired points/borehole :
35
 The variance :
1
 The covariance for the first lag (as a fraction of the variance):
.5
 The censoring level (in log(conductivity))  ;
-.43

** without regression **

BOREHOLE =  1
REGRESSIONS PARAMETERS:   0.0000       0.0000
NUMBER OF CENSORED:   7
NUMBER OF POINTS:  35
ROBINSON VARIANCE:     1.029

 DISTANCE(m) COVARIANCE
  0  1.000000000000
 20  0.6724530000000
 40 -0.6316369000000E-01
 60 -0.2912286000000
 80 -0.4679500000000E-01


BOREHOLE =  2
REGRESSIONS PARAMETERS:   0.0000       0.0000
NUMBER OF CENSORED:   9
NUMBER OF POINTS:  35
ROBINSON VARIANCE:     1.460

 DISTANCE(m) COVARIANCE
  0  1.000000000000
 20  0.9979726000000
 40  0.9979726000000
 60  0.9979726000000
 80  0.9486658000000


BOREHOLE =  3
REGRESSIONS PARAMETERS:   0.0000       0.0000
NUMBER OF CENSORED:  17
NUMBER OF POINTS:  35
ROBINSON VARIANCE:    0.1767

 DISTANCE(m) COVARIANCE
  0  1.000000000000
 20 -0.7761088000000
 40 -0.8620119000000
 60 -0.8920774000000
 80 -0.9020616000000



** with regression **

BOREHOLE =  1
REGRESSIONS PARAMETERS:   0.8458      -0.1341E-02
NUMBER OF CENSORED:   7
NUMBER OF POINTS:  35
ROBINSON VARIANCE:    0.9622

 DISTANCE(m) COVARIANCE
  0  0.7108958000000
 20  0.2988527000000
 40 -0.2599414000000
 60 -0.4296202000000
```

```
 80 -0.4086976000000


BOREHOLE =  2
REGRESSIONS PARAMETERS:  -0.7471       0.2352E-02
NUMBER OF CENSORED:   8
NUMBER OF POINTS:  33
ROBINSON VARIANCE:    0.7844

 DISTANCE(m) COVARIANCE
  0  0.8708097000000
 20  0.3489486000000
 40 -0.5149958000000E-01
 60 -0.1154720000000
 80 -0.7892986000000E-01


BOREHOLE =  3
REGRESSIONS PARAMETERS:   0.1170      -0.1146E-02
NUMBER OF CENSORED:   8
NUMBER OF POINTS:  19
ROBINSON VARIANCE:    0.1309

 DISTANCE(m) COVARIANCE
  0  0.2467309000000
 20 -0.2285989000000E-01
 40 -0.9868227000000E-01
 60 -0.1078231000000
 80 -0.1199522000000


****************************

** 300-400m & 600-720m skipped **

 Regression parameter 1 {B(0)}:
0
 Regression parameter 2 {B(1)}:
0
 Number of boreholes :
3
 Number of desired points/borehole :
50
 The variance :
1
 The covariance for the first lag (as a fraction of the variance):
0
 The censoring level (in log(conductivity))  ;
-.43

** without regression **

BOREHOLE =  1
REGRESSIONS PARAMETERS:   0.0000       0.0000
NUMBER OF CENSORED:  10
NUMBER OF POINTS:  37
ROBINSON VARIANCE:     1.251

 DISTANCE(m) COVARIANCE
  0  1.000000000000
 20  0.2193590000000
 40 -0.2366432000000
 60 -0.5337340000000
 80 -0.2244877000000


BOREHOLE =  2
REGRESSIONS PARAMETERS:   0.0000       0.0000
NUMBER OF CENSORED:  13
NUMBER OF POINTS:  37
ROBINSON VARIANCE:    0.4790
```

```
 DISTANCE(m) COVARIANCE
  0  1.000000000000
 20 -0.6382105000000
 40 -0.4022835000000
 60 -0.4895308000000
 80 -0.4426647000000


BOREHOLE =  3
REGRESSIONS PARAMETERS:   0.0000       0.0000
NUMBER OF CENSORED:  13
NUMBER OF POINTS:  37
ROBINSON VARIANCE:     1.665

 DISTANCE(m) COVARIANCE
  0  1.000000000000
 20 -0.7381669000000E-01
 40  0.6776657000000
 60  0.1662018000000
 80  0.9507925000000E-01


** with regression **

BOREHOLE =  1
REGRESSIONS PARAMETERS:   0.2849      -0.3086E-03
NUMBER OF CENSORED:  10
NUMBER OF POINTS:  37
ROBINSON VARIANCE:     1.154

 DISTANCE(m) COVARIANCE
  0  0.9657983000000
 20  0.3222058000000E-01
 40 -0.3235887000000
 60 -0.6991059000000
 80 -0.4445513000000


BOREHOLE =  2
REGRESSIONS PARAMETERS:  -0.6404E-01  -0.1147E-03
NUMBER OF CENSORED:  13
NUMBER OF POINTS:  37
ROBINSON VARIANCE:    0.5852

 DISTANCE(m) COVARIANCE
  0  0.5792913000000
 20 -0.1439270000000
 40  0.6012533000000E-01
 60  0.3970416000000E-01
 80  0.2738474000000E-01


BOREHOLE =  3
REGRESSIONS PARAMETERS:  -0.5130E-01  -0.1765E-04
NUMBER OF CENSORED:  13
NUMBER OF POINTS:  37
ROBINSON VARIANCE:     1.693

 DISTANCE(m) COVARIANCE
  0  1.381483000000
 20 -0.2253912000000
 40  0.6259455000000
 60  0.4913032000000E-01
 80 -0.5695138000000E-01


************************

** 300-400m & 600-720m skipped **

 Regression parameter 1 {B(0)}:
0
```

```
 Regression parameter 2 {B(1)}:
0
 Number of boreholes :
3
 Number of desired points/borehole :
50
 The variance :
1
 The covariance for the first lag (as a fraction of the variance):
.2
 The censoring level (in log(conductivity))  ;
-.43

** without regression **

BOREHOLE =  1
REGRESSIONS PARAMETERS:   0.0000       0.0000
NUMBER OF CENSORED:   8
NUMBER OF POINTS:  37
ROBINSON VARIANCE:     1.080

 DISTANCE(m) COVARIANCE
  0  1.000000000000
 20  0.4237328000000
 40 -0.1152675000000
 60 -0.4346467000000
 80 -0.1482476000000


BOREHOLE =  2
REGRESSIONS PARAMETERS:   0.0000       0.0000
NUMBER OF CENSORED:  14
NUMBER OF POINTS:  37
ROBINSON VARIANCE:    0.3594

 DISTANCE(m) COVARIANCE
  0  1.000000000000
 20 -0.5751540000000
 40 -0.5215539000000
 60 -0.5829236000000
 80 -0.6624223000000


BOREHOLE =  3
REGRESSIONS PARAMETERS:   0.0000       0.0000
NUMBER OF CENSORED:  13
NUMBER OF POINTS:  37
ROBINSON VARIANCE:     1.485

 DISTANCE(m) COVARIANCE
  0  1.000000000000
 20  0.3160341000000E-01
 40  0.5887223000000
 60  0.1213113000000
 80  0.1457949000000E-01


** with regression **

BOREHOLE =  1
REGRESSIONS PARAMETERS:   0.3975      -0.3626E-03
NUMBER OF CENSORED:   8
NUMBER OF POINTS:  37
ROBINSON VARIANCE:     1.090

 DISTANCE(m) COVARIANCE
  0  0.8370599000000
 20  0.1628277000000
 40 -0.2505899000000
 60 -0.5982199000000
 80 -0.4027219000000
```

```
 BOREHOLE =  2
 REGRESSIONS PARAMETERS:   0.5731E-01  -0.3613E-03
 NUMBER OF CENSORED:  14
 NUMBER OF POINTS:  37
 ROBINSON VARIANCE:    0.4646

  DISTANCE(m) COVARIANCE
   0  0.5064536000000
  20 -0.1083253000000E-01
  40  0.5456385000000E-01
  60  0.4296564000000E-01
  80 -0.3699069000000E-01


 BOREHOLE =  3
 REGRESSIONS PARAMETERS:  -0.2362E-01  -0.1029E-03
 NUMBER OF CENSORED:  13
 NUMBER OF POINTS:  37
 ROBINSON VARIANCE:     1.492

  DISTANCE(m) COVARIANCE
   0  1.413765000000
  20 -0.5564978000000E-01
  40  0.5677003000000
  60  0.6475119000000E-01
  80 -0.1115966000000


********************************

** 300-400m & 600-720m skipped **

 Regression parameter 1 {B(0)}:
0
 Regression parameter 2 {B(1)}:
0
 Number of boreholes :
3
 Number of desired points/borehole :
50
 The variance :
1
 The covariance for the first lag (as a fraction of the variance):
.5
 The censoring level (in log(conductivity))  ;
-.43

** without regression **

BOREHOLE =  1
REGRESSIONS PARAMETERS:   0.0000      0.0000
NUMBER OF CENSORED:   9
NUMBER OF POINTS:  37
ROBINSON VARIANCE:    0.9911

 DISTANCE(m) COVARIANCE
  0  1.000000000000
 20  0.7504585000000
 40  0.1493966000000
 60 -0.1728832000000
 80  0.1605256000000E-01


BOREHOLE =  2
REGRESSIONS PARAMETERS:   0.0000      0.0000
NUMBER OF CENSORED:  15
NUMBER OF POINTS:  37
ROBINSON VARIANCE:    0.4527

 DISTANCE(m) COVARIANCE
  0  1.000000000000
```

```
  20 -0.2647521000000
  40 -0.4662836000000
  60 -0.7262391000000
  80 -0.7672804000000


 BOREHOLE =  3
 REGRESSIONS PARAMETERS:   0.0000        0.0000
 NUMBER OF CENSORED:  12
 NUMBER OF POINTS:  37
 ROBINSON VARIANCE:     1.095

  DISTANCE(m) COVARIANCE
   0  1.000000000000
  20  0.3056069000000
  40  0.5536355000000
  60  0.3575814000000
  80  0.4371487000000E-01


 ** with regression **

 BOREHOLE =  1
 REGRESSIONS PARAMETERS:   0.6211      -0.6718E-03
 NUMBER OF CENSORED:   9
 NUMBER OF POINTS:  37
 ROBINSON VARIANCE:    0.8582

  DISTANCE(m) COVARIANCE
   0  0.7906615000000
  20  0.4298934000000
  40 -0.1416041000000E-01
  60 -0.3769861000000
  80 -0.3860365000000


 BOREHOLE =  2
 REGRESSIONS PARAMETERS:   0.2470      -0.7918E-03
 NUMBER OF CENSORED:   9
 NUMBER OF POINTS:  25
 ROBINSON VARIANCE:    0.6193

  DISTANCE(m) COVARIANCE
   0  0.5156996000000
  20  0.3189856000000
  40  0.1133570000000
  60 -0.9053656000000E-01
  80 -0.3994449000000E-01


 BOREHOLE =  3
 REGRESSIONS PARAMETERS:  -0.4458E-01   0.1742E-03
 NUMBER OF CENSORED:  12
 NUMBER OF POINTS:  37
 ROBINSON VARIANCE:    0.9746

  DISTANCE(m) COVARIANCE
   0  1.075045000000
  20  0.7936543000000E-01
  40  0.3788300000000
  60  0.1147357000000
  80 -0.1449667000000


***********************************

The estimator delta21:

 Regression parameter 1 {B(0)}:
0
 Regression parameter 2 {B(1)}:
0
```

```
 Number of boreholes :
3
 Number of desired points/borehole :
100
 The variance :
1
 The covariance for the first lag (as a fraction of the variance):
0
 The censoring level (in log(conductivity))  ;
-.43

** without regression **

BOREHOLE =  1
REGRESSIONS PARAMETERS:   0.0000      0.0000
NUMBER OF CENSORED:  30
NUMBER OF POINTS: 100
ROBINSON VARIANCE:     1.025

 DISTANCE(m) COVARIANCE
  0  1.000000000000
 20 -0.1056786000000
 40 -0.2688800000000
 60 -0.2485747000000
 80 -0.1130831000000


BOREHOLE =  2
REGRESSIONS PARAMETERS:   0.0000      0.0000
NUMBER OF CENSORED:  36
NUMBER OF POINTS: 100
ROBINSON VARIANCE:     1.178

 DISTANCE(m) COVARIANCE
  0  1.000000000000
 20 -0.3120687000000E-01
 40  0.7903141000000E-01
 60 -0.1453283000000
 80 -0.3768471000000E-01


BOREHOLE =  3
REGRESSIONS PARAMETERS:   0.0000      0.0000
NUMBER OF CENSORED:  40
NUMBER OF POINTS: 100
ROBINSON VARIANCE:    0.6589

 DISTANCE(m) COVARIANCE
  0  1.000000000000
 20 -0.2609031000000
 40 -0.4176118000000
 60 -0.4574852000000
 80 -0.4242072000000


** with regression **

BOREHOLE =  1
REGRESSIONS PARAMETERS:   0.2165     -0.1868E-03
NUMBER OF CENSORED:  30
NUMBER OF POINTS: 100
ROBINSON VARIANCE:     1.035

 DISTANCE(m) COVARIANCE
  0  0.8994565000000
 20 -0.1075637000000
 40 -0.2176110000000
 60 -0.2007178000000
 80 -0.1001014000000


BOREHOLE =  2
```

```
 REGRESSIONS PARAMETERS:   0.3580E-01  -0.7123E-04
 NUMBER OF CENSORED:  36
 NUMBER OF POINTS: 100
 ROBINSON VARIANCE:     1.254

  DISTANCE(m) COVARIANCE
   0  1.183830000000
  20 -0.2546172000000E-01
  40  0.8748903000000E-01
  60 -0.1689185000000
  80 -0.2828683000000E-01


 BOREHOLE =  3
 REGRESSIONS PARAMETERS:  -0.2468E-01  -0.1247E-03
 NUMBER OF CENSORED:  40
 NUMBER OF POINTS: 100
 ROBINSON VARIANCE:    0.7913

  DISTANCE(m) COVARIANCE
   0  0.8510776000000
  20 -0.3908275000000E-01
  40 -0.1805547000000
  60 -0.2323587000000
  80 -0.1897641000000


 **********************************

 The estimator delta10:

  Regression parameter 1 {B(0)}:
 0
  Regression parameter 2 {B(1)}:
 0
  Number of boreholes :
 3
  Number of desired points/borehole :
 100
  The variance :
 1
  The covariance for the first lag (as a fraction of the variance):
 0
  The censoring level (in log(conductivity))  ;
 -.43

 ** without regression **

 BOREHOLE =  1
 REGRESSIONS PARAMETERS:   0.0000       0.0000
 NUMBER OF CENSORED:  30
 NUMBER OF POINTS: 100
 ROBINSON VARIANCE:     1.025

  DISTANCE(m) COVARIANCE
   0  1.000000000000
  20 -0.2001892000000
  40 -0.4654635000000
  60 -0.4353600000000
  80 -0.2133784000000


 BOREHOLE =  2
 REGRESSIONS PARAMETERS:   0.0000       0.0000
 NUMBER OF CENSORED:  36
 NUMBER OF POINTS: 100
 ROBINSON VARIANCE:     1.178

  DISTANCE(m) COVARIANCE
   0  1.000000000000
  20 -0.6143987000000E-01
  40  0.1643088000000
```

```
 60 -0.2695362000000
 80 -0.7394928000000E-01


BOREHOLE =  3
REGRESSIONS PARAMETERS:   0.0000       0.0000
NUMBER OF CENSORED:  40
NUMBER OF POINTS: 100
ROBINSON VARIANCE:    0.6589

 DISTANCE(m) COVARIANCE
  0  1.000000000000
 20 -0.4537357000000
 40 -0.6608240000000
 60 -0.7056777000000
 80 -0.6684626000000


** with regression **

BOREHOLE =  1
REGRESSIONS PARAMETERS:   0.2165      -0.1868E-03
NUMBER OF CENSORED:  30
NUMBER OF POINTS: 100
ROBINSON VARIANCE:     1.035

 DISTANCE(m) COVARIANCE
  0  0.8994565000000
 20 -0.2022641000000
 40 -0.3825740000000
 60 -0.3566445000000
 80 -0.1890624000000


BOREHOLE =  2
REGRESSIONS PARAMETERS:   0.3580E-01  -0.7123E-04
NUMBER OF CENSORED:  36
NUMBER OF POINTS: 100
ROBINSON VARIANCE:     1.254

 DISTANCE(m) COVARIANCE
  0  1.183830000000
 20 -0.5037581000000E-01
 40  0.1814438000000
 60 -0.3137344000000
 80 -0.5589769000000E-01


BOREHOLE =  3
REGRESSIONS PARAMETERS:  -0.2468E-01  -0.1247E-03
NUMBER OF CENSORED:  40
NUMBER OF POINTS: 100
ROBINSON VARIANCE:    0.7913

 DISTANCE(m) COVARIANCE
  0  0.8510776000000
 20 -0.7637076000000E-01
 40 -0.3228051000000
 60 -0.4012794000000
 80 -0.3372166000000


****************************
```

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

Technical Reports

List of SKB Technical Reports 1990

TR 90-01
FARF31 –
A far field radionuclide migration code for use with the PROPER package
Sven Norman[1], Nils Kjellbert[2]
[1] Starprog AB
[2] SKB AB
January 1990

TR 90-02
Source terms, isolation and radiological consequences of carbon-14 waste in the Swedish SFR repository
Rolf Hesböl, Ignasi Puigdomenech, Sverker Evans
Studsvik Nuclear
January 1990