SKB

TECHNICAL REPORT 94-22

Evaluation of stationary and non-stationary geostatistical models for inferring hydraulic conductivity values at Äspö

Paul R La Pointe

Golder Associates Inc., Seattle, WA, USA

November 1994

SVENSK KÄRNBRÄNSLEHANTERING AB
SWEDISH NUCLEAR FUEL AND WASTE MANAGEMENT CO

BOX 5864 S-102 40 STOCKHOLM
TEL. 08-665 28 00 TELEX 13108 SKB
TELEFAX 08-661 57 19

EVALUATION OF STATIONARY AND NON-STATIONARY GEOSTATISTICAL MODELS FOR INFERRING HYDRAULIC CONDUCTIVITY VALUES AT ÄSPÖ

Paul R La Pointe

Golder Associates Inc., Seattle, WA, USA

November 1994

EVALUATION OF STATIONARY AND NON-STATIONARY GEOSTATISTICAL MODELS FOR INFERRING HYDRAULIC CONDUCTIVITY VALUES AT ÄSPÖ

November 1994

Paul R. La Pointe

Golder Associates Inc.
Seattle, WA USA

Keywords: geohydrology, hydraulic conductivity, geostatistics, covariance models, semivariograms, stochastic estimation, Äspö

November 9, 1994

IC3-1129

TABLE OF CONTENTS

		Page No.
1	INTRODUCTION	1
2	EXPLORATORY DATA ANALYSIS	9
	2.1 Overview	9
	2.2 Subpopulations	9
	2.3 Trends	23
	2.4 Discussion	23
3	GEOSTATISTICAL ANALYSIS	27
	3.1 Overview	27
	3.2 Anisotropy and Neighborhoods	28
	3.3 Cross-Validation Results of Stationary Geostatistical Models	38
3.3.1	"Delete-1" Cross-Validation Results	38
	3.3.2 "Delete-10%" Cross-Validation Results	41
	3.3.3 "Delete-50%" Cross-Validation Results	45
	3.4 Cross-Validation Results of Non-Stationary Geostatistical Models	48
	3.4.1 "Delete-1" Cross-Validation Results	48
	3.4.2 "Delete-10%" Cross-Validation Results	51
	3.4.3 "Delete-50%" Cross-Validation Results	54
	3.5 Discussion of Results	55
	3.6 Regularization and Upscaling of Packer Tests	60
	3.6.1 Regularization Strategy	60
	3.6.2 Results of Geostatistical Analysis of 30 m Test Data	66
4	CONCLUSIONS AND RECOMMENDATIONS	69
	4.1 Conclusions Concerning the Spatial Statistics of the Packer Test Data	69
	4.2 Recommendations for Stochastic Hydrological Modeling at Äspö	70
5	REFERENCES CITED	75

LIST OF FIGURES

1-1 Location of well data used in this study

1-2 Definition of error statistics

2-1 Comparison of 3 m hydraulic conductivity test results with spinner survey anomalies. a) KAS02 b) KAS03 c) KAS04 d) KAS05 e) KAS06 f) KAS07 g) KAS08

2-2 Histograms and Cumulative Histograms of 3 m packer test data. a) KAS02 b) KAS03 c) KAS04 d) KAS05 e) KAS06 f) KAS07 g)KAS08 h) KLX01

3-1 Alternative methods to compute radial lag pairs. a) conventional method. b) improved box method.

3-2 Directional variograms for 3 m test data. a) east/west b) north/south c) northeast/southwest d) northwest/southeast e) vertical f) horizontal g) isotropic, h) isotropic variogram with number of data pairs.

3-3 Advantages of using a sectored neighborhood to include points from multiple wells. a) single sector neighborhood b) 4 sector neighborhood

3-4 "Delete-1" cross-validation results for alternative neighborhood definitions using the variogram of Figure 3-3.

3-5 Fitted stationary variograms for the "Delete-1" cross-validation tests. a) Nugget + Exponential b) Nugget + Spherical c) Nugget + Spherical (short range) + Exponential (long range) d) Nugget + Spherical (short range) + Spherical (long range)

3-6 Test statistics for "Delete-1" cross-validation of four stationary models

3-7 Test statistics for 10 realizations of the "Delete-10%" cross-validation of the stationary model

3-8 Test statistics for 10 realizations of the "Delete-50%" cross-validation of the stationary model

3-9 Test statistics for 10 realizations of the "Delete-10%" cross-validation of the non-stationary model

3-10 Test statistics for 10 realizations of the "Delete-50%" cross-validation of the non-stationary model

LIST OF FIGURES (Cont.)

3-11 Comparison of cross-validation statistics for the "Delete-10%" cross-validation tests. a) MSE b) ME c) MRE d) MSRE e) R

3-12 Comparison of cross-validation statistics for the "Delete-50%" cross-validation tests. a) MSE b) ME c) MRE d) MSRE e) R

3-13 3 m data regularized to 30 m scale using Norman's (1992a) formula compared to measured 30 m results

3-14 3 m data regularized by arithmetic averaging compared to measured 30 m results.

3-15 Raw variogram and variogram cloud for 30 m test data.

LIST OF TABLES

1.1 Well Data Completeness Summary

3.1 Test statistics for alternative neighborhood models

3.2 Parameters for alternative stationary variograms fitted by Isatis

3.3 Cross validation statistics for "Delete-1" test of alternative stationary models

3.4 Cross-validation statistics for the ten realizations used in the "Delete-10%" test of the stationary model

3.5 Cross-validation statistics for the ten realizations used in the "Delete-50%" test of the stationary model

3.6 Drift analysis results for non stationary models in the "Delete-1" tests.

3.7 Cross-validation statistics for the ten realizations used in the "Delete-1" test of the non-stationary model

3.8 Drift analysis results for non stationary models in the "Delete-10%" tests.

3.9 Cross-validation statistics for the ten realizations used in the "Delete-10%" test of the non-stationary model

3.10 Drift analysis results for non stationary models in the "Delete-50%" tests.

LIST OF TABLES (Cont.)

3.11 Cross-validation statistics for the ten realizations used in the "Delete-50%" test of the non-stationary model

3.12 Cross-validation results for 30 m well tests.

ABSTRACT (Swedish)

Rapporten beskriver en jämförelse mellan stationära och icke-stationära geostatistiska modeller för att beräkna den hydraulika konduktiviteten för bergblock från hydraultester på Äspö. Modellerna jämförs genom att utvärdera korsvalideringsstatistik för tre valideringsfall. Första fallet består av en "Delete-1" test som tidigare har använts på data från Finnsjön. Det andra testet består av "Delete-10%" och det tredje testet är "Delete-50%". Analysen har genomförts med en komersiell programvara Isatis 1.3 (Geovariances, 1994).

Den preliminära analysen visar att data från hydraultester med 3 m och 30 m manschettavstånd kan behandlas som stickprov från samma population vid geostatistiska analyser, möjligen med undantag för data från borrhål KAS02. Analys av 3 m data ger ingen indikation på någon systematisk statistisk förändring med djup, bergart, sprickzon kontra icke-sprickzon eller någon annan kartlagd faktor. Riktningsvariogram är svåra att beräkna och utvärdera på grund av den ringa spridningen hos data (data är grupperade längs borrhålen), men visar ingen tydlig tendens till anisotropi. Kriging med 8 söksektorer och en sökradie på 200 m ger goda resultat jämfört med andra testade sökkriterier och har därför använts i den efterföljande analysen.

Stationär analys antyder att det förekommer en stor rumsligt okorrelerad komponent ("Nugget Effect") för 3 m data, i storleksordningen 60% av den observerade variansen för de olika modellerna. Fyra olika modeller har använts för automatisk anpassning till data:

1) "Nugget Effect + Spherical"
2) "Nugget Effect + Exponential"
3) "Nugget Effect + Spherical(short range) + Exponential(long range)"
4) "Nugget Effect + Spherical(short range) + Spherical(long range)"

Resultaten för alla modellerna vad beträffar korsvalideringsstatistik är väldigt lika för de tre valideringsfallen, men modell 3) ger generellt bättre resultat och har därför används i den efterföljande korsvalideringen.

Icke-stationär analys visar att både graden av "drift" och graden av "intrinsic random function" är låg. För de sökkriterier som använts är "driften" antingen av ordningen noll eller ett och kovariansen av ordningen noll och tre.

Vad gäller korsvalideringsstatistik är de icke-stationära modellerna något bättre än de stationära modellerna för "Mean Error (ME)", men mycket sämre för "Mean Reduced Error (MRE)" och "Mean Squared Reduced Error (MSRE)". "Mean Squared Error (MSE)" är något bättre för det stationära fallet. Dessa resultat antyder starkt att stationära modeller kommer att producera mindre korsvalideringsfel för Äspö data. I motsats till situationen vid Finnsjön verkar data från hydraultester vid Äspö vara mycket stationära utan "trend" och bara med låg grad av "drift", vilket innebär att det

inte är någon fördel att använda "intrinsic random functions" och att stationära modeller möjligen kan beskriva kovariansen bättre.

Uppskalning av konduktivitetsdata från hydraultester i liten skala till större skala (intervall) med Normans formel (1992a) överskattar den hydrauliska konduktiviteten för det uppskalade intervallet. Ett enkelt aritmetiskt medelvärde producerar ett jämförelsevis bättre resultat, men all uppskalning av denna typ introducerar okorrelerat brus i data och tenderar att dölja de sanna rumsliga egenskaperna i konduktivitetsfältet. Uppskalningen med flytande medelvärdesbildning, som för närvarande är implementerad i INFERENS 1.1, kan också introducera andra icke önskvärda effekter. Det rekommenderas därför att geostatistisk analys och skattning utförs på icke uppskalade data i liten skala och att uppskalning till konduktivitetsdata för bergblock i större skala görs efter den geostatistiska skattningen (simuleringen).

Denna studie föreslår också att konventionella korsvalideringsstudier och automatisk variogramanpassning inte nödvändigtvis ger en möjlighet att bedöma hur bra en given modell skattar den hydrauliska konduktiviteten i blockskala. Det föreslås att en serie av stokastiska modellexperiment utförs för att utvärdera hur känsliga resultaten från den hydrologiska modellen är för osäkerheter i den geostatistiska slutledningsmodellen och uppskalningsprocessen.

ABSTRACT (English)

This report describes the comparison of stationary and non-stationary geostatistical models for the purpose of inferring block-scale hydraulic conductivity values from packer tests at Äspö. The comparison between models is made through the evaluation of cross-validation statistics for three experimental designs. The first experiment consisted of a "Delete-1" test previously used at Finnsjon. The second test consisted of "Delete-10%" and the third test was a "Delete-50%" test. The analysis was carried out using the commercial code Isatis 1.3 (Geovariances, 1994).

Preliminary data analysis showed that the 3 m and 30 m packer test data can be treated as a sample from a single population for the purposes of geostatistical analyses, with the possible exception of the data from well KAS02. Analysis of the 3 m data does not indicate that there are any systematic statistical changes with depth, rock type, fracture zone vs non-fracture zone or other mappable factor. Directional variograms are ambiguous to interpret due to the clustered nature of the data, but do not show any obvious anistropy that should be accounted for in geostatistical analysis. A Kriging neighborhood divided into 8 angular sectors with a radius of 200 m provided good results while making computations efficient compared to the other neighborhoods tested. This neighborhood definition was used for the remainder of the analyses.

Stationary analysis suggested that there exists a sizeable spatially uncorrelated component ("Nugget Effect") in the 3 m data, on the order of 60% of the observed variance for the various models fitted. Four different nested models were automatically fit to the data:

1) Nugget Effect + Spherical
2) Nugget Effect + Exponential
3) Nugget Effect + Spherical(short range) + Exponential(long range)
4) Nugget Effect + Spherical(short range) + Spherical(long range)

Results for all models in terms of cross-validation statistics were very similar for the first set of validation tests, but the model consisting of the Nugget Effect, Spherical and Exponential elementary models performed slightly better overall and was chosen for the additional cross-validation tests.

Non-stationary analysis established that both the order of drift and the order of the intrinsic random functions is low. For the neighborhood selected, drift is either zero-th or first order; the elementary covariance models are zero and third order.

In terms of cross-validation statistics, the non-stationary models were slightly better than the stationary models in terms of Mean Error (ME), but much worse in terms of Mean Reduced Error (MRE) and Mean Square Reduced Error (MSRE). Mean Squared Error (MSE) was slightly better for the stationary case. These results strongly suggest that stationary models will produce smaller cross-validation errors at

Äspö. Unlike the situation at Finnsjon, the packer test data at Äspö seem to be very stationary with no trend and only low order drift, so that advantages of using intrinsic random functions are lost and stationary models may more accurately model the covariance.

Regularization of packer test data to larger intervals using Norman's (1992a) formula overpredicts the regularized interval hydraulic conductivity. A simple arithmetic average produces less biased results, but any regularization of this type introduces additional uncorrelated noise into the data, and will tend to obscure the true spatial properties of the hydraulic conductivity field. A moving average regularization process as currently implemented in Inferens 1.1 may also introduce other undesirable effects, and it is recommended that geostatistical analysis and estimation be performed on unregularized data., with regularization to block-scale hydraulic conductivity values performed after geostatistical estimation or simulation.

This study also suggests that conventional cross-validation studies and automatic variogram fitting are not necessarily evaluating how well a model will infer block scale hydraulic conductivity values. It is suggested that a series of stochastic modeling experiments be conducted to examine how sensitive the hydrological modeling results are to uncertainties in the geostatistical inference model and the upscaling process.

1 INTRODUCTION

The goal of a useful stochastic inference model is the ability to make reliable predictions of conductivity properties of a rock mass at large distances away from measured data (Norman, 1992a, pg. 81). Optimizing geostatistical inference models with this goal in mind has been carried out by Norman (1992a) and Geier (1993a) at the Finnsjön site.

Previous cross-validation studies of the Finnsjön rock mass using INFERENS 1.0 led to cross-validation results that were not thought to be sufficiently good (Norman, 1982a, pg 82), in that the proposed geostatistical models with the best jackknife statistics would be rejected at a 95% level of confidence. Examination of the results led to the suggestion that certain improvements to the geostatistical algorithms might improve the cross-validation results. Among the improvements suggested were the addition of nested variograms. These and other changes to Inferens 1.0 are described in detail by Geier (1993a). Subsequent cross validation studies (Geier, 1993a, pg. 64) using the INFERENS 1.1 code, however, did not show much improvement at Finnsjön.

The type of variograms available in both releases of INFERENS presume that the data set analyzed is stationary; such variograms can yield poor results if the data is not stationary. A variety of geostatistical methods exist to interpolate data in certain types of non-stationary situations (Renard, 1989). The purpose of this project is to evaluate the performance of non-stationary geostatistical algorithms for use in interpolating packer test-derived permeability measurements, and to determine whether such algorithms should be incorporated into a new version of INFERENS and HYDRASTAR. The commercial code ISATIS 1.3 (Geovariances, 1994) was used to evaluate the new algorithms. This report also includes a discussion of other stochastic inference methods that might improve hydrological modeling.

The algorithms were applied to data for the Äspö Hard Rock Laboratory described by Nilsson (1989,1990). The data consisted of 3 meter and 30 meter interval tests from wells KAS02, KAS03, KAS04, KAS05, KAS06, KAS07, KAS08 and KLX01. The value of interval hydraulic conductivity that is used in this study were calculated by Nilsson (1989, 1990) using the Jacob interpretation method. As Follin (1992) has shown for stochastic porous media, intervals with higher conductivity values tend to have a larger region of influence or statistical support than do intervals with lower values. This difference in the support among the test intervals could affect the variogram calculations and Kriging cross-validation results. While a similar tendency may also pertain to packer tests in fractured rock, it is not possible to take results from porous media and transfer them to fractured rock. This is because the head or flow in any location in a fracture network connected to the borehole interval tested in a double packer well test is not a simple function of the Euclidian distance from the interval. Rather, it is the geometry and fluid flow characteristics of the fracture pathway from the interval to the location that affects the head or flow. It is not uncommon to find a very small response in nearby fractures not directly connected to the well, and larger responses at a greater distance. Because the actual geometry of the fracture network for each well test at Aspo is unknown, the potential difference in statistical support for the Jacob values was ignored. Table 1.1 summarizes the data used in this study and its completeness. As Table 1.1 shows, many of the 3 m interval tests produced results below the resolution of the instruments. This threshold varied from well to well (Nilsson, 1989, 1990). Approximately half of the 3 m data obtained is above the threshold.

Utilizing data below measurement resolution to compute variograms has drawbacks. One strategy might be to use the data *above* measurement resolution to infer the statistical properties of the data below, and then to generate stochastic realizations of the data below the threshold. However, this imposes a spatial correlation structure on the data below measurement resolution. When it is then combined with the other

data, the effect will be to artificially reduce the variability of the calculated variogram. No new information will actually be added, and moreover, a model has been imposed upon a portion of the data that may actually not be accurate.

Another strategy might be to discretize the well tests into classes and to apply Indicator geostatistical methods. In this scheme, measurements below the threshold would fall into classes. There are problems with this procedure for the Aspo data. The first problems is that the threshold varies from well to well, so it is not possible to put all of the below-threshold data into their own class or classes. Secondly, INFERENS and HYDRASTAR do not support Indicator geostaistical procedures. Finally, discretization of the measurements above threshold into classes reduces the precision of estimation for cross-validation tests and eventually, for creating permeability input for HYDRASTAR or similar code. For these reasons, only the measurements above the threshold were included.

Figure 1-1 shows the spatial location of the 3 meter and 30 meter data. The fact that test data is very unevenly distributed and hierarchically structured (that is, it is acquired as a line sample) makes accurate interpolation quite challenging. However, this is exactly the situation that will be encountered in a typical site hydrological modeling effort, as it was in the Finnsjön study (Norman, 1992a), so the evaluation of the algorithmic performance under these conditions is appropriate.

Table 1.1

PACKER TEST DATA SUMMARY				
Well Name	Total Number of 3 m Tests	Number of 3 m Tests Above Measurement	% 3 m Tests Above Threshold	Total Number of 30 m Tests
KAS02	233	107	45.92%	23
KAS03	150	113	75.33%	21
KAS04	108	68	62.96%	none
KAS05	128	41	32.03%	none
KAS06	162	86	53.09%	none
KAS07	163	52	31.90%	none
KAS08	158	68	43.04%	none
KLX01	197	125	63.45%	19
Total	1299	660	50.81%	63

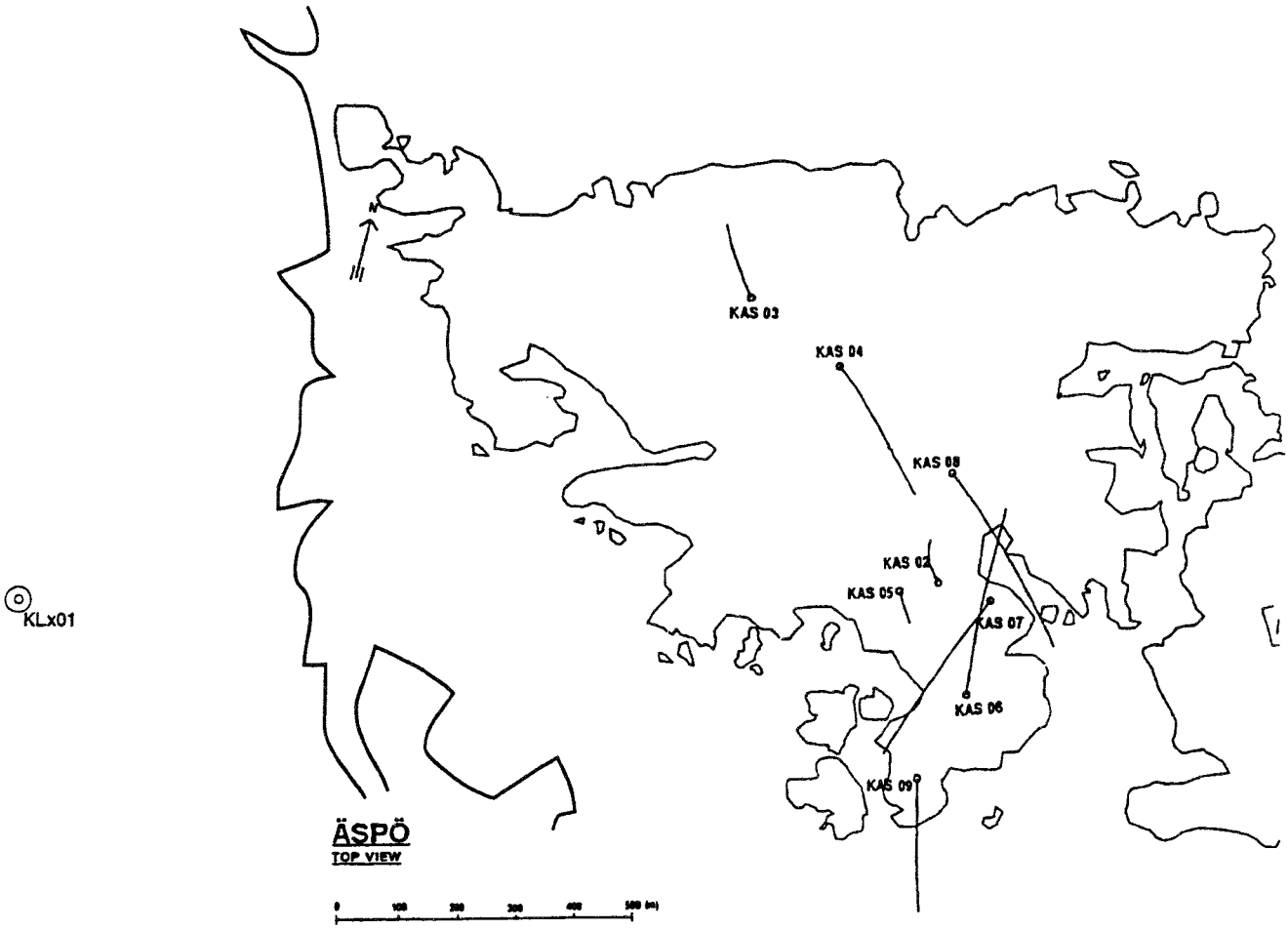


FIGURE **1-1**

LOCATION OF WELL DATA USED IN THIS STUDY

SKB/REPORT/SWEDEN

PROJECT NO. IC3 1129.000 DRAWING NO. 40868 DATE 06/27/94 DRAWN BY TB

Golder Associates

There are many ways to evaluate the performance of a model. The INFERENS codes use a jackknife cross-validation scheme in which each data point used to construct the variogram is deleted from the data set, and is then estimated based upon the remaining data and the variogram model selected. This type of cross-validation could be termed a "Delete-1" algorithm. The disparity between the predicted value and the actual value is quantified in terms of different error statistics. For the cross-validation calculations in this report, the following statistics were used: the Mean Error (ME), the Mean Reduced Error (MRE), the Mean Square Reduced Error (MSRE), the Mean Squared Error (MSE) and Variance Ratio (R). Their definitions and their relation to error statistics used previously at Finnsjon are described in Figure 1-2. In the literature, there are different definitions of the ME and MRE statistics. The definitions for ME and MRE in this report follow Russo and Jury (1987, pg. 1260). MSRE is calculated using the equation presented in Norman (1992a, pg 46). R is a statistic is a ratio of the experimental and theoretical variances as described in Renard (1988, pg 586) and is similar to the MSRE statistic. R, estimated from a single estimation of one data set (or single data point) by another, is biased. This bias can be reduced by randomly dividing the data set into subsets and using each in turn to estimate the other. This procedure is followed in the third series of cross-validation experiments described in this report. Norman (1992a) presents an overview of the utility of error measures in ranking the performance of alternative models. For a perfectly self-consistent model, ME equals 0.0, MRE, MSRE and R equal 1.0 and MSE is small.

Each error statistic shows different information about how well each model performs. In general, those with a theoretical value of 0 show how *unbiased* the model estimates are, while those with a theoretical value of 1 show how well the *size* of the error matches with the expected size distribution given the model and the geometrical arrangement of samples. It is more informative to consider these performance measures separately rather than combining them, as in Norman (1992a). For example, is it worse to have a model that consistently overestimates the permeability, but matches the variance range correctly? Such a model is unbiased, but may not be

very useful. Or is it worse to have a model which is unbiased, but makes much bigger errors than would be expected? Which error is more tolerable or preferable? For design purposes, both types of errors are important, and it is more useful to know which type of error, if any, is predominating. For this reason, it is not very useful to combine the different error statistics into a single score for ranking models.

In addition to the "Delete-1" procedure, this study used two additional experimental designs to evaluate the performance of the algorithms under conditions that are closer to actual site modeling. In such situations, the amount of well test data represents a very small part of the rock mass to be modeled. It is worthwhile to see how well the model could estimate a much larger subset of the well test data, rather than single tests. The second set of cross-validation tests consisted of randomly deleting 10% of the well tests, and estimating it based upon the remaining 90%. These cross-validation tests are termed the "Delete-10%" tests. The final cross validation tests were designed to reduce the effect of outlier data on the cross validation results. In this third set of experiments, the data was randomly divided into two equal sets, and each was used to estimate the other. The statistics for the estimate of set 1 by set 2 and set 2 by set 1 were combined in such a way so as to reduce the effect of outliers, as discussed in greater detail in Section 3.1. Ten independent realizations each for the "Delete-10%" and the "Delete-50%" cross-validation tests were processed.

The remainder of this report is divided into three sections:

Section 2 describes the exploratory data analysis prior to geostatistical calculations. This exploratory analysis was used to confirm that the mathematical requirements of various geostatistical algorithms were or were not fulfilled by the Äspö data, and to determine if there were distinct populations that should not be merged.

Section 3 summarizes the fitting of the stationary and non-stationary models and details the cross-validation results.

Section 4 summarizes the principal conclusions of this study, and describes alternative possibilities for creating stochastic hydrological models at Äspö.

Terminology, this Report	Equivalent Statistic in Norman (1992a) and Geier (1993a)	Abbreviation	Formulation	Theoretical Value
Reduced Error	Reduced Kriging Error	RE	$RE(x_i) = \frac{[u^*(x_i)-u(x_i)]}{Var[u^*(x_i)-u(x_i)]^{1/2}}$	0
Mean Error	Mean Reduced Error	ME	$ME = \frac{1}{n}\sum_{i=1}^{n} RE(x_i)$	0
Mean-Square Reduced Error	$MSRE^{1/2}$	MRE	$MRE = \left[\frac{1}{n}\sum_{i=1}^{n} RE(x_i)^2\right]^{1/2}$	1
Mean Square Error	$MSE^{1/2}$	MSE	$MSE = \left[\frac{1}{n}\sum_{i=1}^{n}[u^*(x_i)-u(x_i)]^2\right]^{1/2}$	0
Mean Square Reduced Error	Not used	MSRE	$MSRE = \frac{1}{n}\sum_{i=1}^{n} RE(x_i)^2$	1
R	Not used	R	$R = \frac{\sum_{i=1}^{n}[u^*(x_i)-u(x_i)]^2}{\sum_{i=1}^{n} Var[u^*(x_i)-u(x_i)]}$	1

FIGURE 1-2
DESCRIPTION OF ERROR STATISTICS
SKB/GEOSTATISTICS/SWEDEN

PROJECT NO. IC3 1129 DRAWING NO. 41074 DATE 11/8/94 DRAWN BY TB

2 EXPLORATORY DATA ANALYSIS

2.1 Overview

It is important to understand the statistical properties of the packer test data and their relation to geology before carrying out geostatistical cross-validation computations. Geostatistical methods make certain requirements for interpolation or simulation, among them, that the data come from a population with certain types of stationarity, and a population frequency distribution that is Gaussian or can be transformed into a Gaussian distribution. These assumptions may be violated if otherwise statistically acceptable data subsets are indiscriminately mixed or lumped together. Moreover, the components of any statistically valid data set can be quite complex, and it is the purpose of preliminary data analysis to sort out some of the components for the appropriate type of analysis. A convenient way to organize the data is into populations, and for each population, to sort out the trend and residual. Trends are systematic or deterministic spatial changes in the statistical properties of the data. Residuals are stochastic. The residual itself may have three components; the drift, the spatially correlated component, and the non-spatially correlated component. Geostatistical analysis focuses on the three components of the residual. Preliminary data analysis usually focuses on the populations and their trend functions.

2.2 Subpopulations

All subsequent geostatistical analysis of the packer data rests upon how the data is initially divided into subpopulations. Since packer test results depend upon the properties of fractures intersecting the well interval, and fracture properties depend upon geology, it is worth testing whether mappable geological units can be correlated with fracture properties or well test results. Two possible geological factors have been considered as a basis for distinguishing subpopulations at Äspö: rock type and fracture zone vs. non-fracture zone.

Liedholm (1991; TN16) analyzed the relation between rock type and hydraulic conductivity measurements for the 3 m tests using stepwise multiple linear regression. The 3 m packer test data constitute the best data set for analysis since it has an order of magnitude more samples, and is spatially more extensive than the 30 m test data. Liedholm (1992) reports that the packer test data only weakly corresponds to mappable petrologic units at Äspö.

This is not entirely unsurprising, as the geometry and intensity of fracturing may be more controlled by mappable *mechanical* boundaries than *geological* boundaries (Stearns and Friedman, 1972). Given the relatively high strength, stiffness, fracture toughness and others mechanical aspects of all of the crystalline rocks types found at Äspö, it is not unexpected that packer tests values do not strongly correspond with the mappable petrologic units.

Another possibility is that packer test values can be split into two populations consisting of tests in identified fracture zones and those outside of fracture zones. If packer tests in fracture zones are a distinct population, then their first or second order statistical moments ought to be such that combining them with non-fracture zone measurements reduces geostatistical estimation/simulation performance. For example, clear evidence that fracture zones should be treated as a separate populations for geostatistical inference is if there are bi-modal distributions of packer tests results, and that each mode consists of either fracture zone and non-fracture zone components.

Subpopulations might also occur for reasons not geologically obvious. Statistical analysis of the data can point out areas of Äspö that may have anomalously high or low packer test measurements, independent of whether there is a geological model to explain the results. If these regions have properties that would violate the assumptions underpinning the geostatistical methods, then it would be inappropriate to lump them together for analysis.

There are ways of qualitatively examining the possibility that fracture zones or other presently unknown geological units constitute distinct populations and that packer results from these zones should not be mixed with the remaining measurements. Several lines of investigation, however, do not indicate that there are distinct subpopulations in the 3 m Äspö data.

Figures 2-1a-g show the hydraulic conductivity values of the 3 m packer tests for wells KAS02 through KAS08 as solid connected lines. The black squares represent the location of significant conductive features as identified by the spinner survey logs. The value of each black square is the inferred transmissivity of each conductive feature. In many cases, these spinner log anomalies correspond to packer tests with high values. However, there are also many packer tests with high values that do not correspond to spinner anomalies.

Histograms of the 3 m packer data (Figures 2-2a through h) show no bi-modality, so that even if fracture zone tests are a different population, they seem to conform to a single uni-modal population that is suited to geostatistical methods.

Cumulative histograms for each well (Figure 2-3) show that, with the possible exception of KAS02, all of the other wells have nearly the same cumulative distribution. This is consistent with the assumption that each set of packer measurements from a particular well can be viewed as a realization of a single underlying population that has reasonably stable (stationary) first and second-order moments over the region sampled by the wells.

Geostatistical analysis described in Sec. 3.2 generates results consistent with a single population. As Geier (1994) mentions in describing the geological causes of nested variograms, a strongly nested variogram can occur when the fractures and well tests within a fracture zone are more similar over short distances than fractures or tests outside the zone. Another type a variogram that occurs more frequently (La Pointe, 1980) is a pseudo-periodicity imposed on the variogram. This leads to variograms of

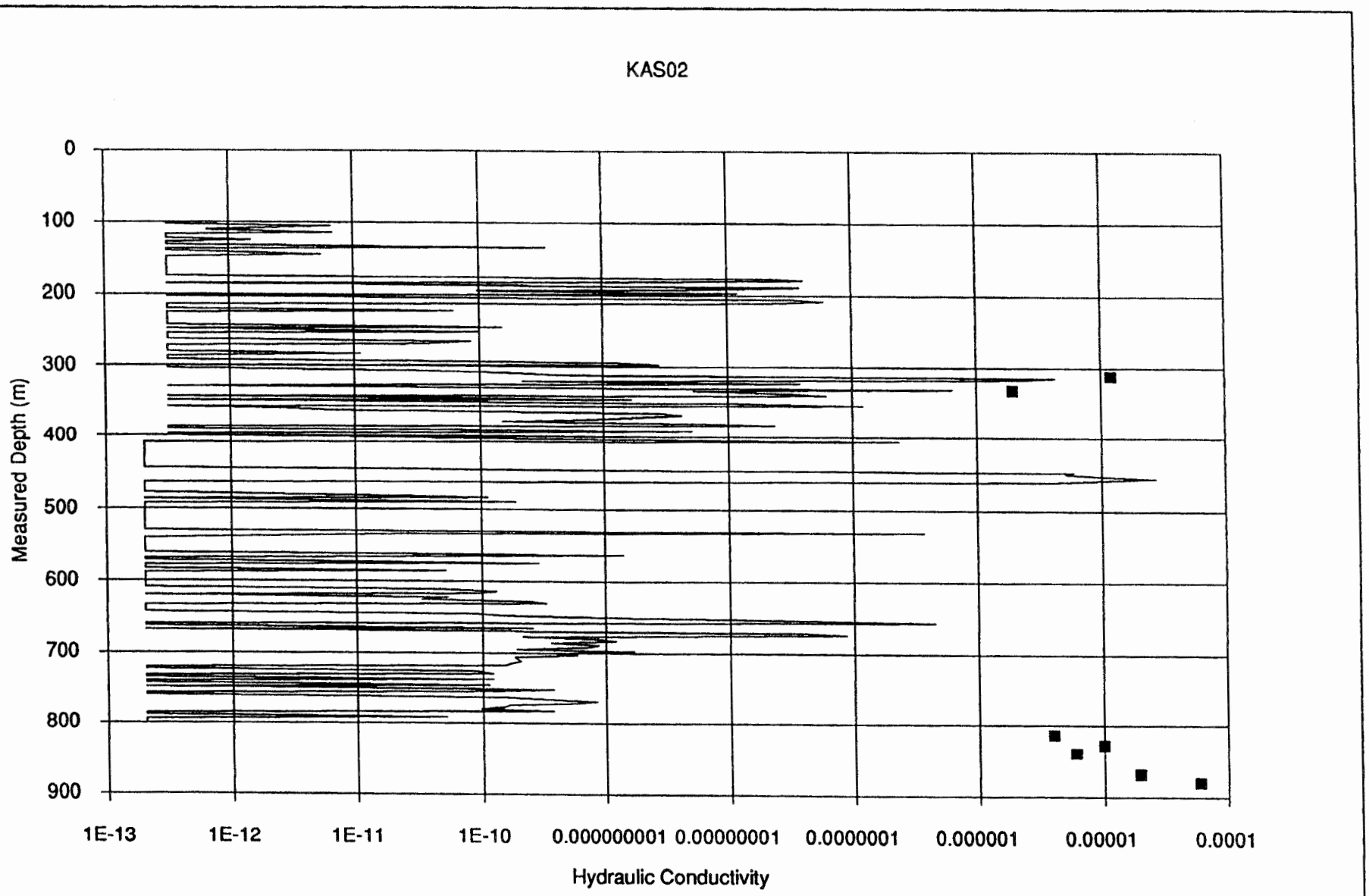


FIGURE **2-1a**

COMPARISON OF 3 M HYDRAULIC CONDUCTIVITY TEST RESULTS WITH SPINNER SURVEY ANOMALIES

SKB/GEOSTATISTICS/SWEDEN

PROJECT NO. IC2 1129.000 DRAWING NO. [illegible] DATE [illegible]

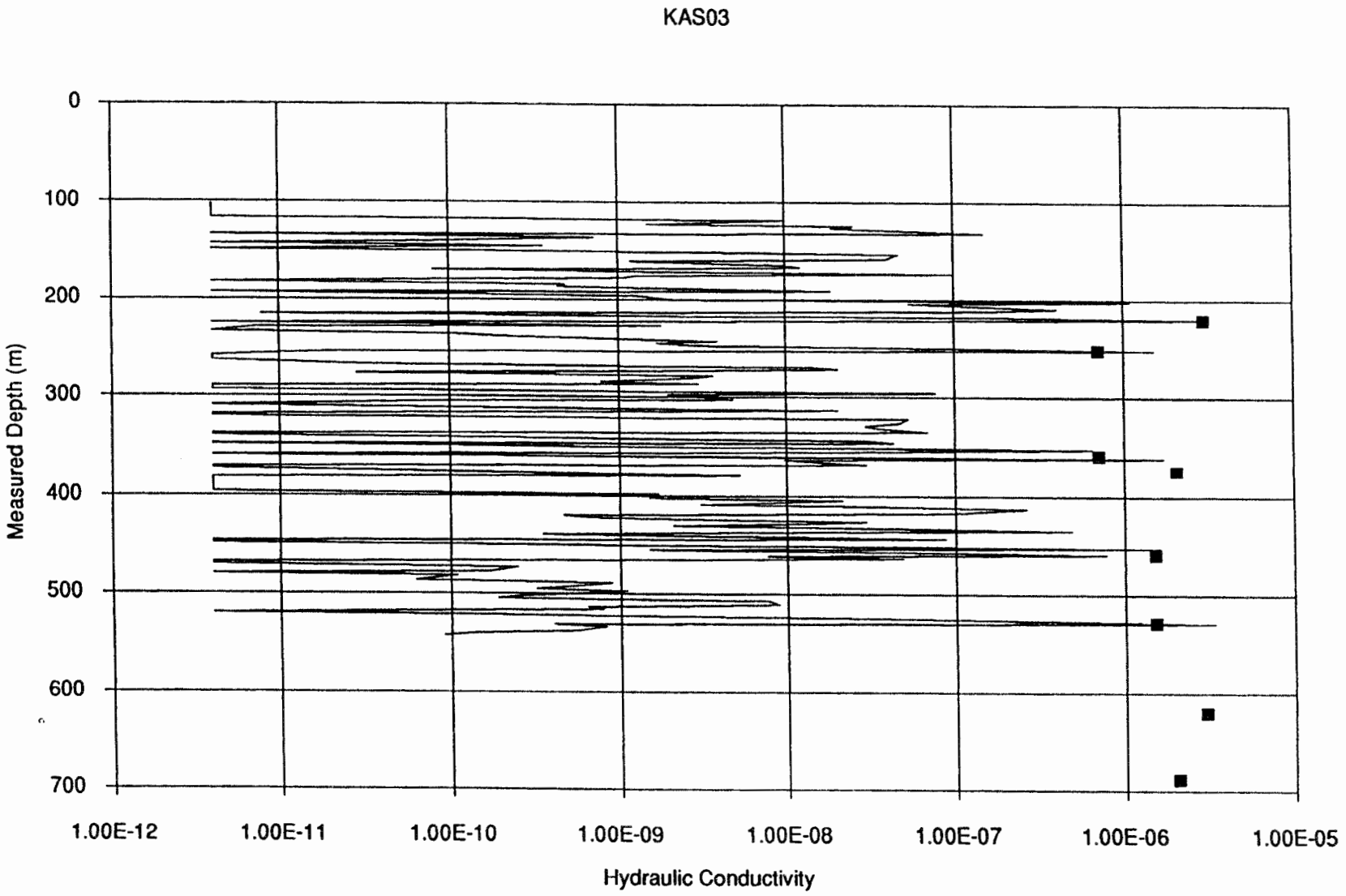


FIGURE **2-1b**

COMPARISON OF 3 M HYDRAULIC CONDUCTIVITY TEST RESULTS WITH SPINNER SURVEY ANOMALIES

SKB/GEOSTATISTICS/SWEDEN

PROJECT NO. IC3 1129.000 DRAWING NO. 49851 DATE 6/29/94 DRAWN BY CB

Golder Associates

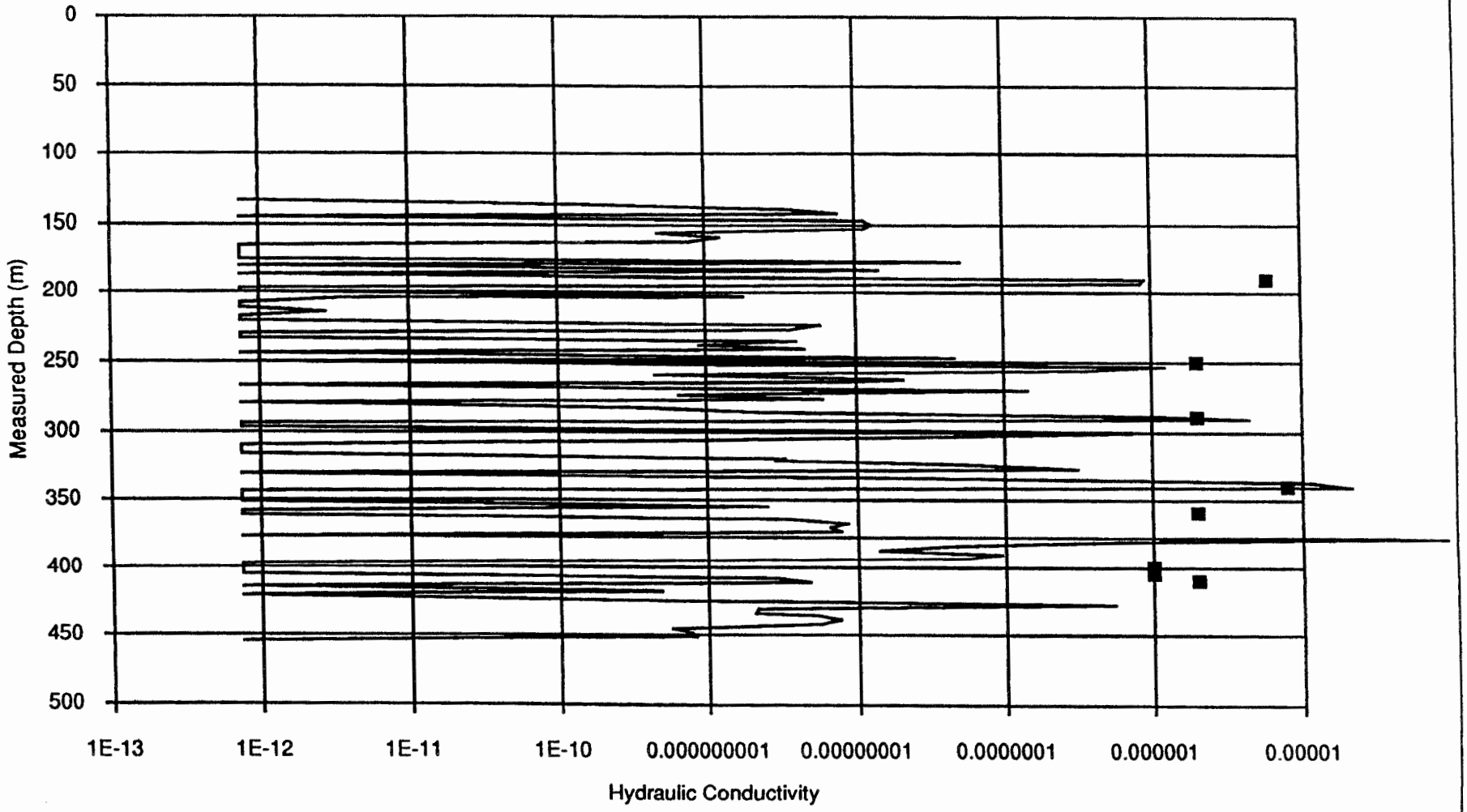


FIGURE **2-1c**

COMPARISON OF 3 M HYDRAULIC CONDUCTIVITY TEST RESULTS WITH SPINNER SURVEY ANOMALIES

SKB/GEOSTATISTICS/SWEDEN

PROJECT NO. IC3 1129.000 DRAWING NO. 49852 DATE 6/29/94 DRAWN BY CP

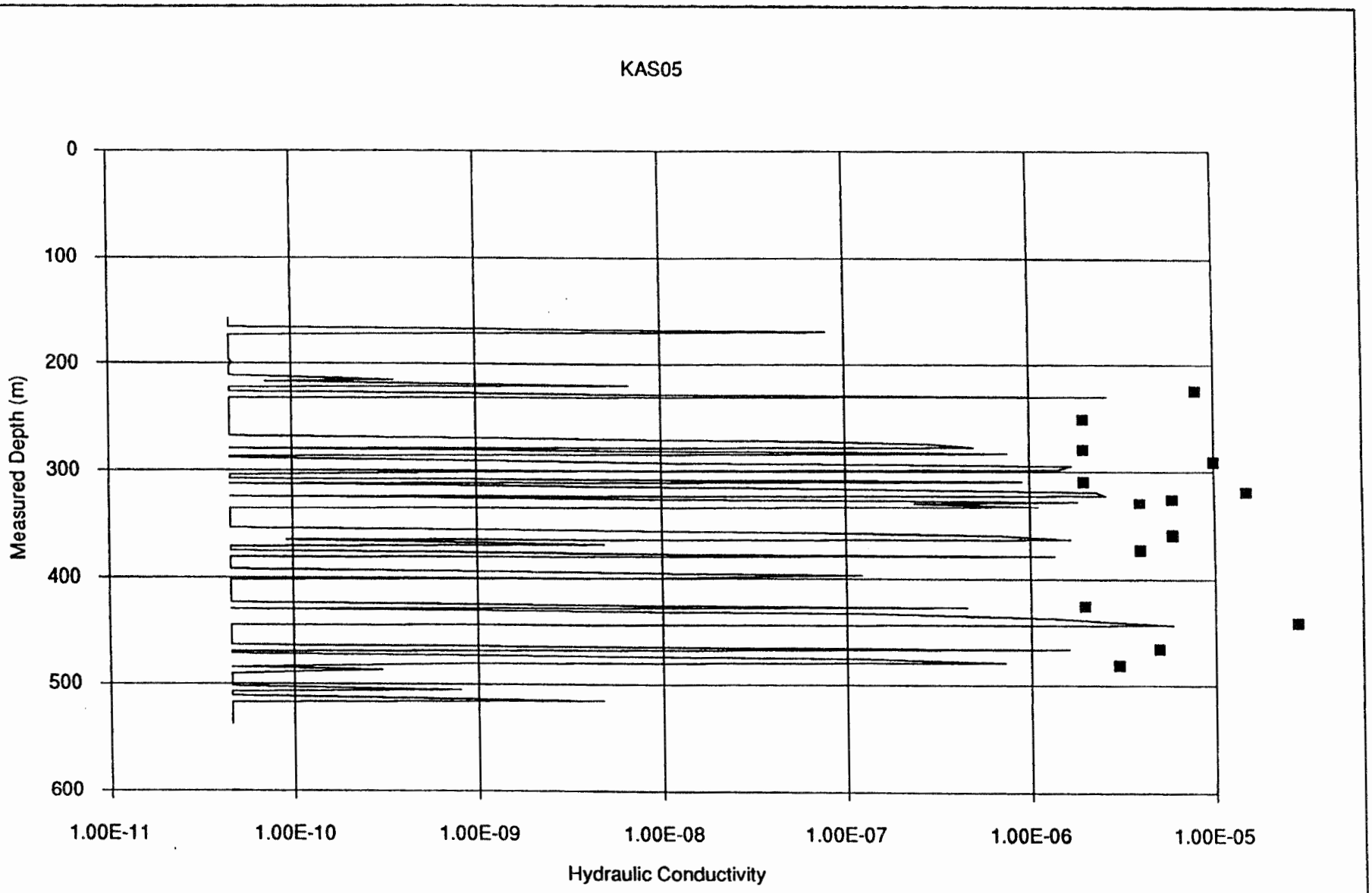


FIGURE **2-1d**

COMPARISON OF 3 M HYDRAULIC CONDUCTIVITY TEST RESULTS WITH SPINNER SURVEY ANOMALIES

SKB/GEOSTATISTICS/SWEDEN

PROJECT NO. IC3 1129.000 DRAWING NO. 49853 DATE 8/29/94 DRAWN BY CB

Golder Associates

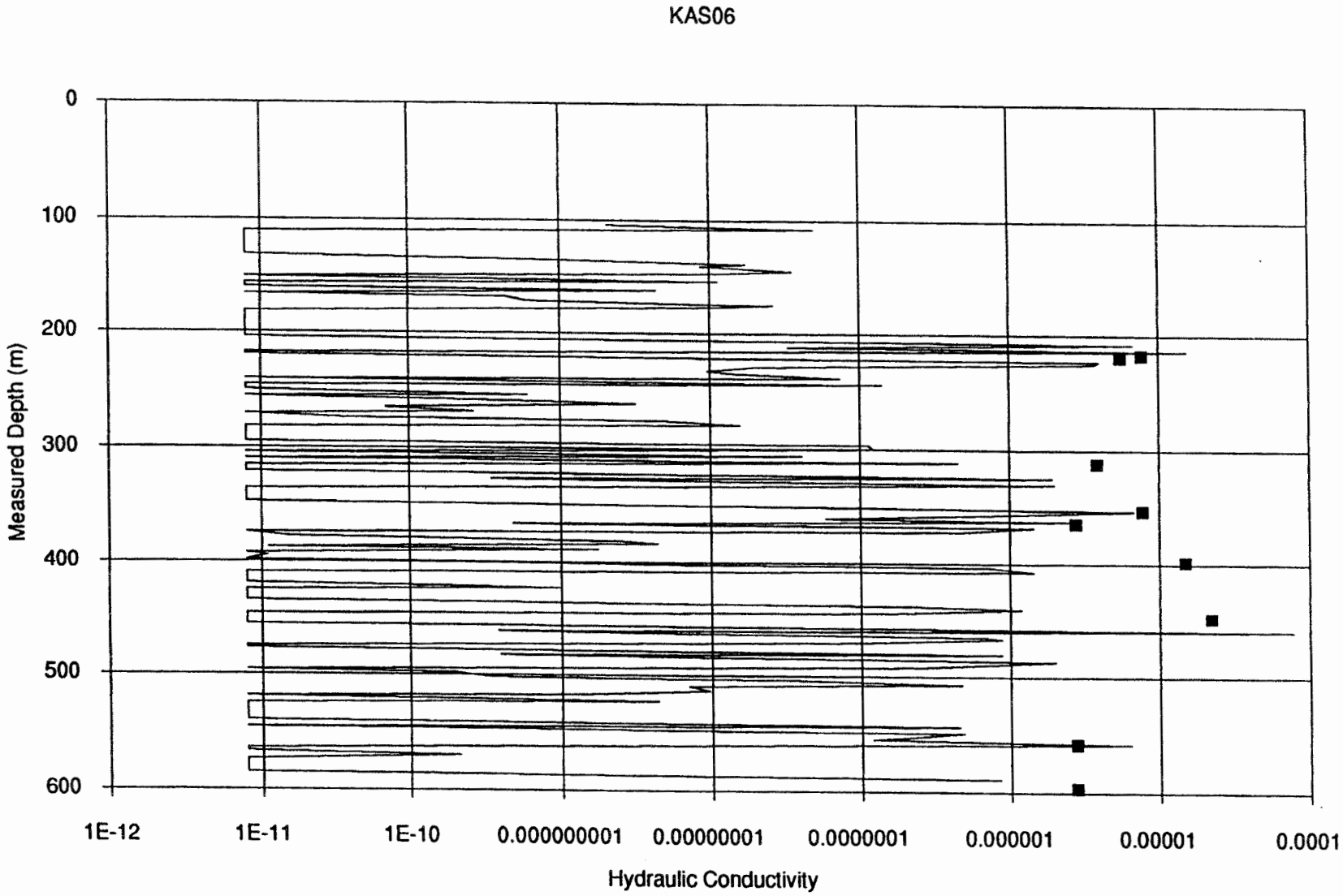


FIGURE **2-1e**

COMPARISON OF 3 M HYDRAULIC CONDUCTIVITY TEST RESULTS WITH SPINNER SURVEY ANOMALIES

SKB/GEOSTATISTICS/SWEDEN

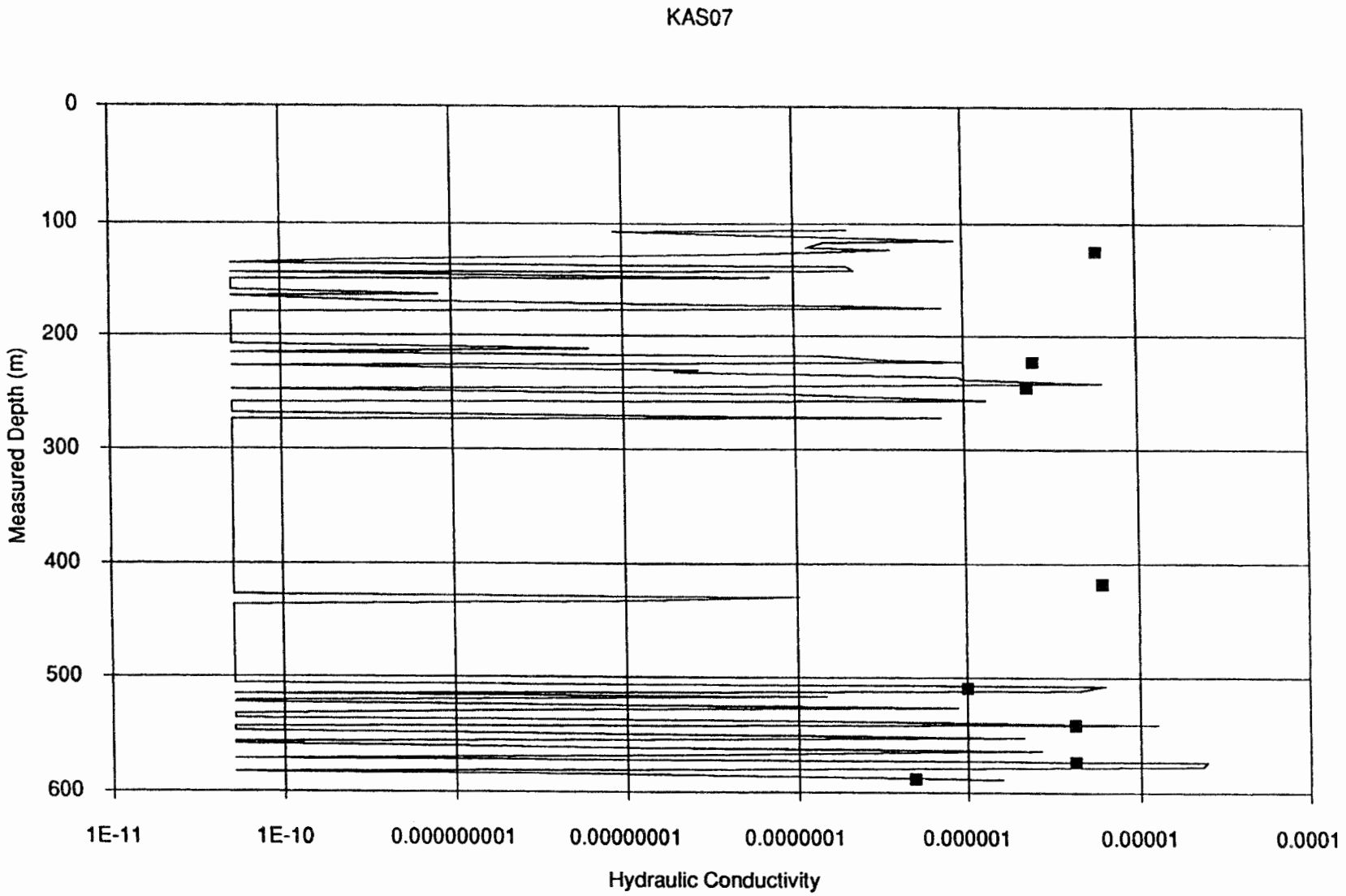


FIGURE **2-1f**

COMPARISON OF 3 M HYDRAULIC CONDUCTIVITY TEST RESULTS WITH SPINNER SURVEY ANOMALIES

SKB/GEOSTATISTICS/SWEDEN

PROJECT NO. IC3 1129.000 DRAWING NO. 49855 DATE 6/29/94 DRAWN BY CB

Golder Associates

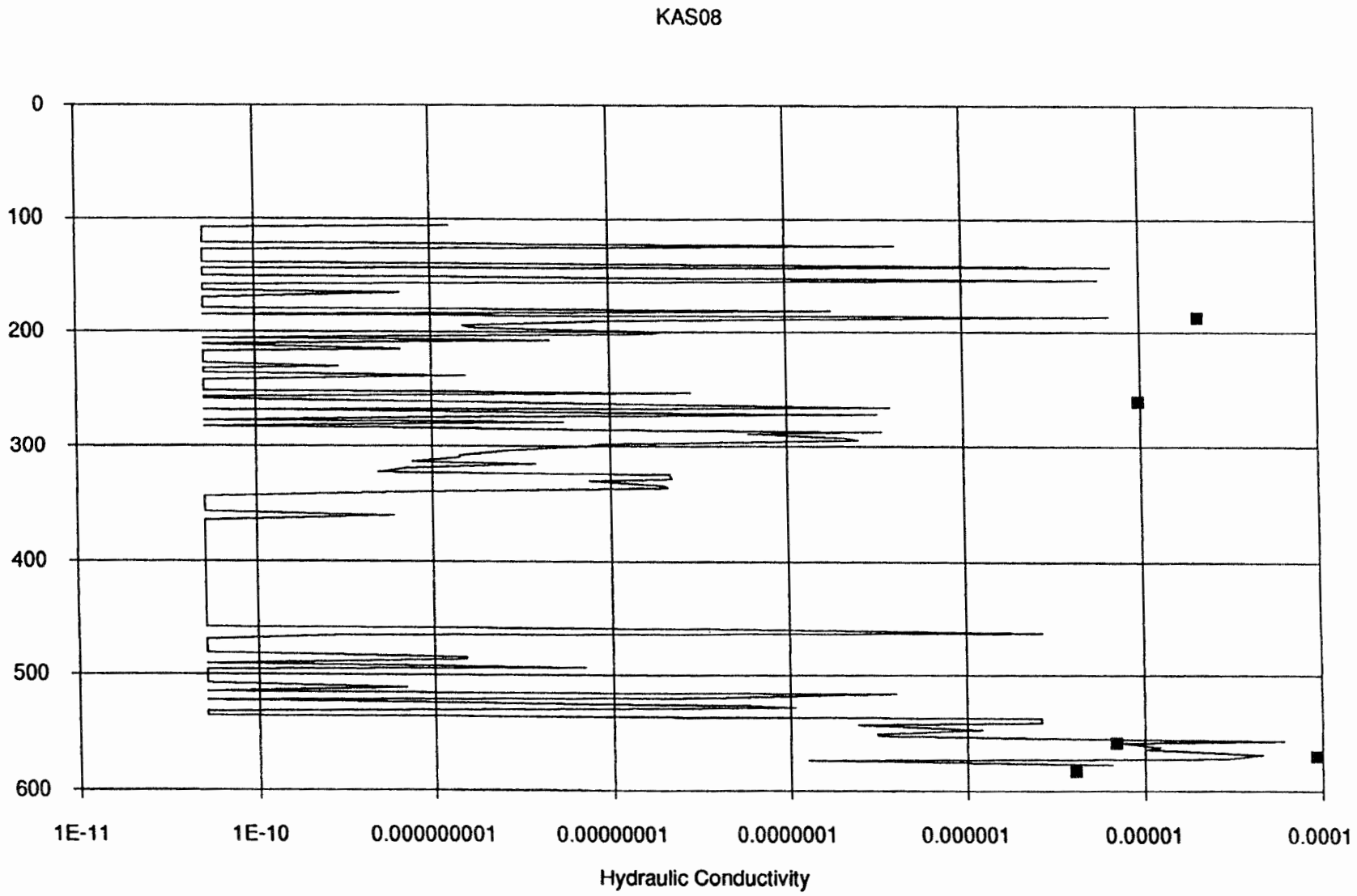
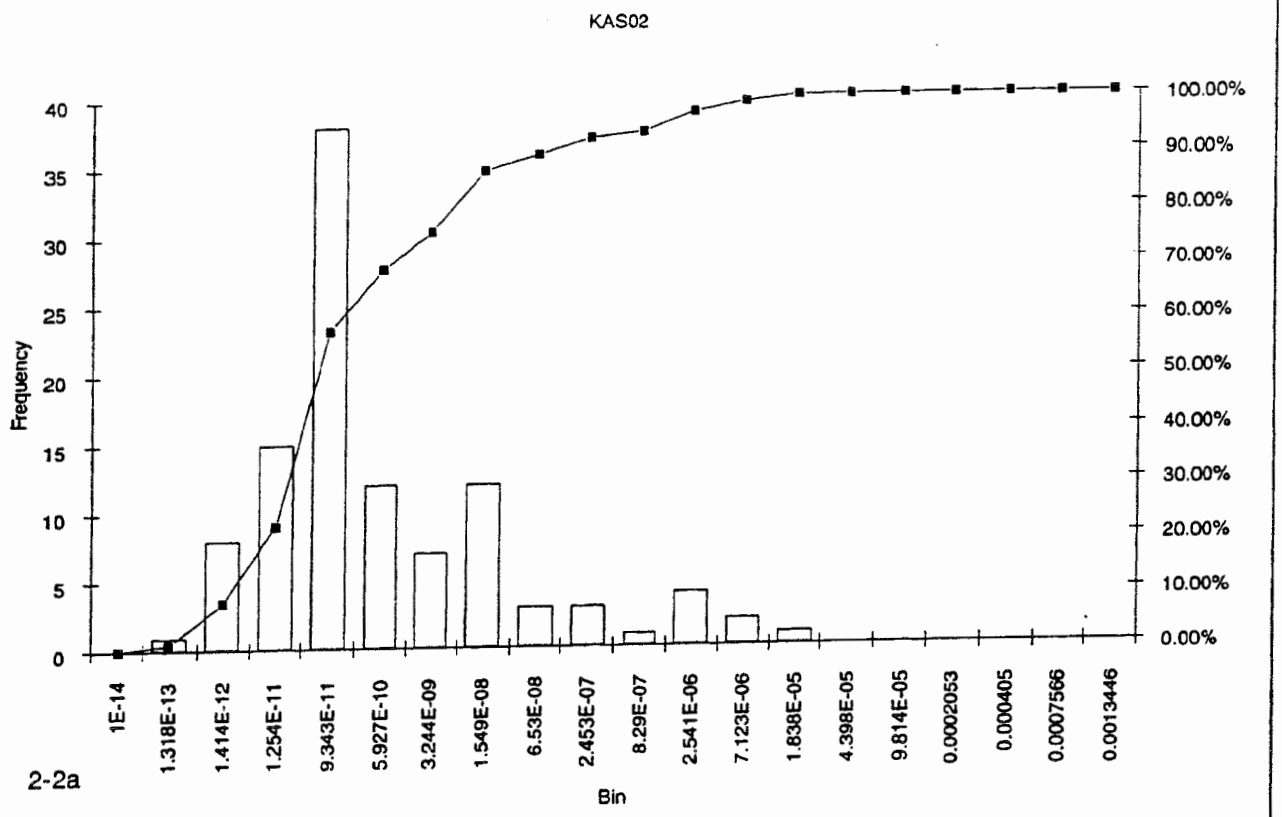


FIGURE **2-1g**

COMPARISON OF 3 M HYDRAULIC CONDUCTIVITY TEST RESULTS WITH SPINNER SURVEY ANOMALIES

SKB/GEOSTATISTICS/SWEDEN

PROJECT NO. IC2 1139 000 DRAWING NO. [illegible] DATE



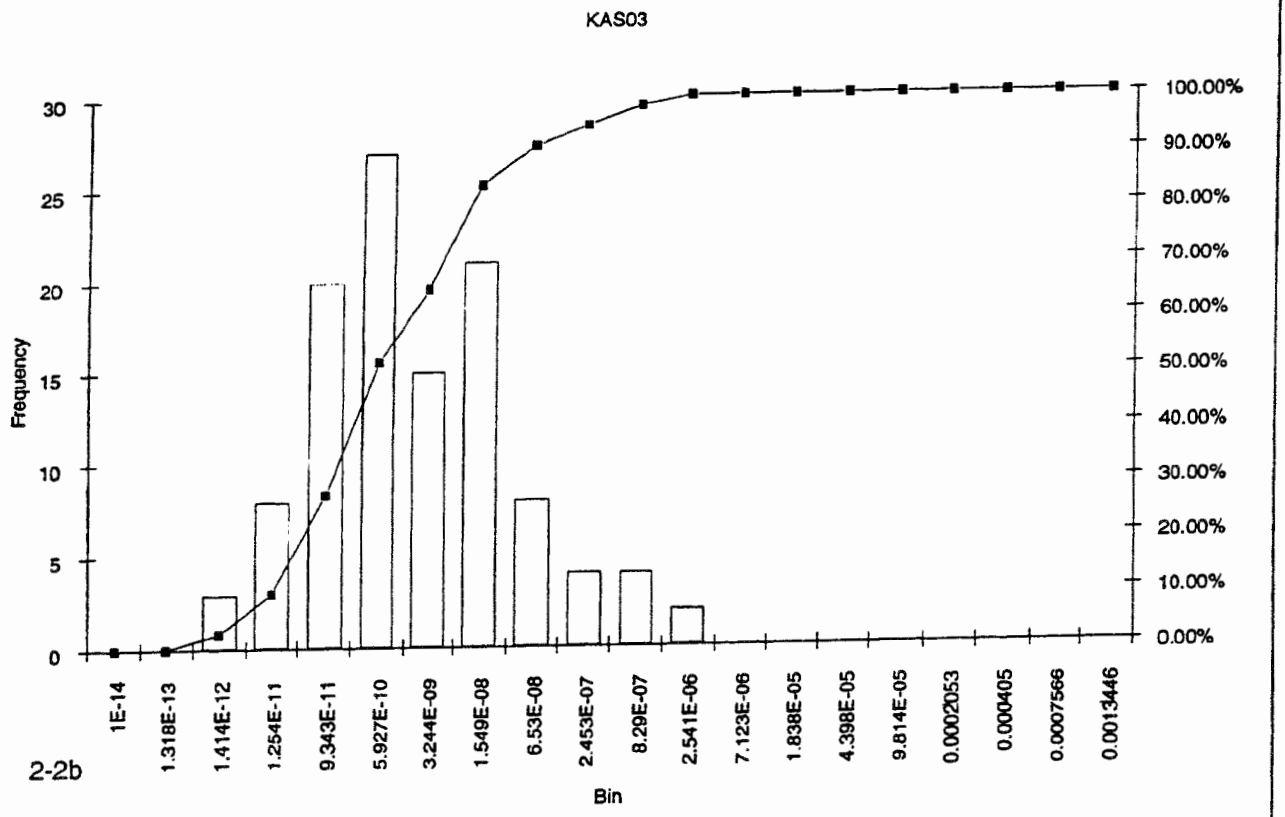
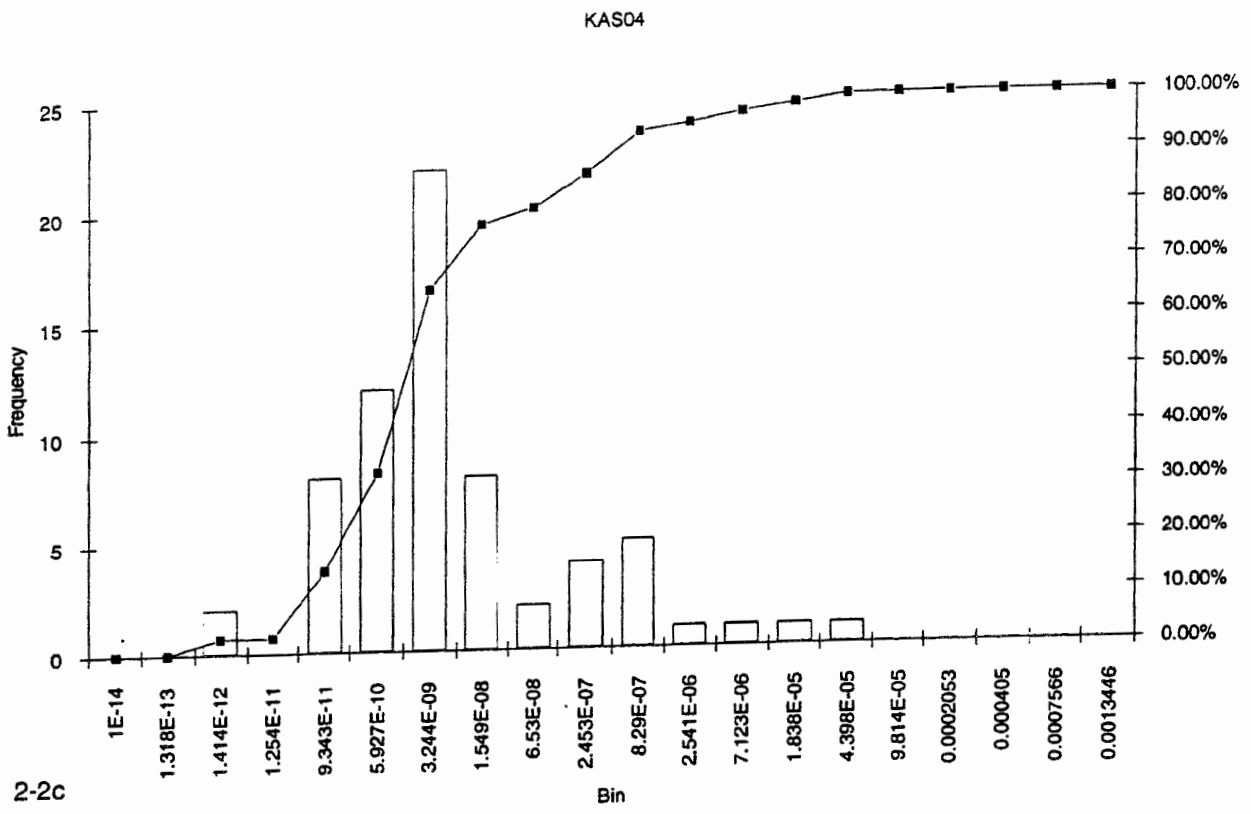


FIGURE **2-2**

HISTOGRAMS AND CUMULATIVE HISTOGRAMS OF 3 M PACKER TEST DATA

SKB/GEOSTATISTICS/SWEDEN

PROJECT NO. ID2 1102 DRAWING NO. 41065 DATE 11/7/94 DRAWN BY TB



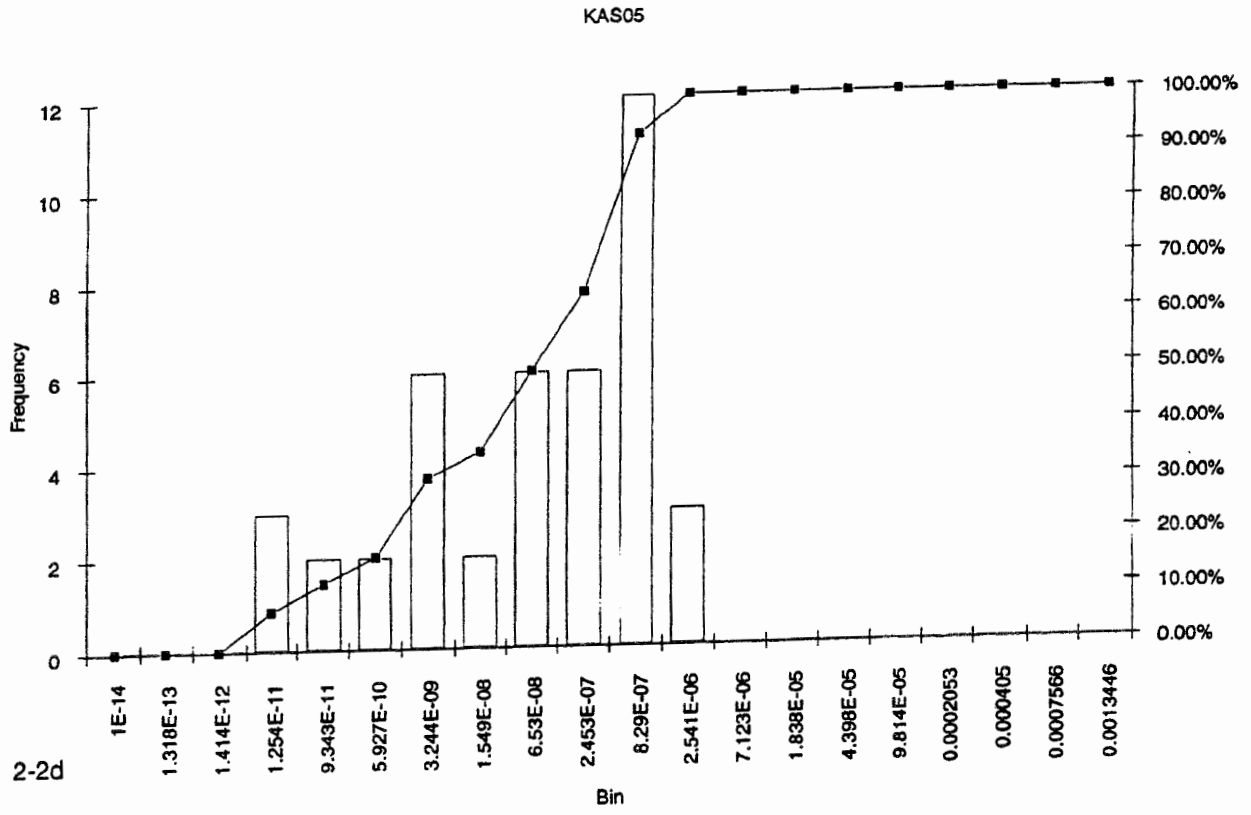
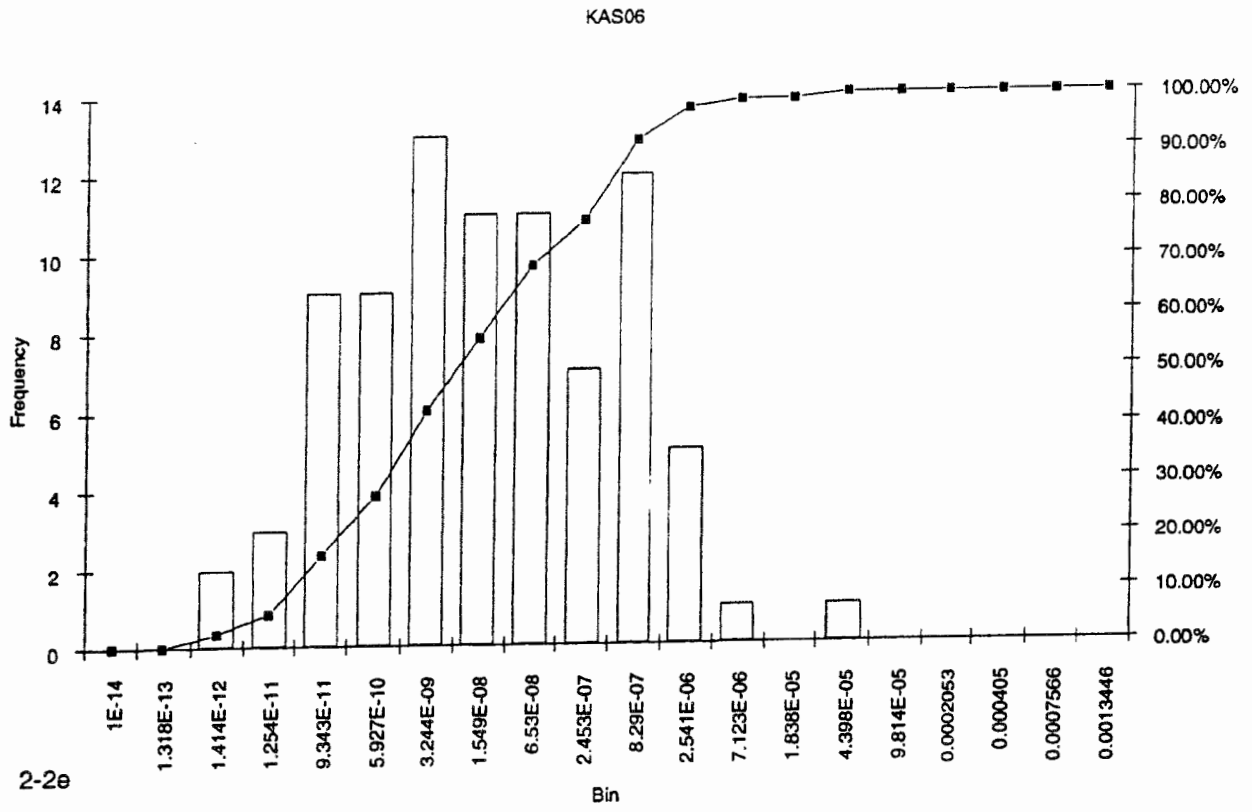


FIGURE **2-2**

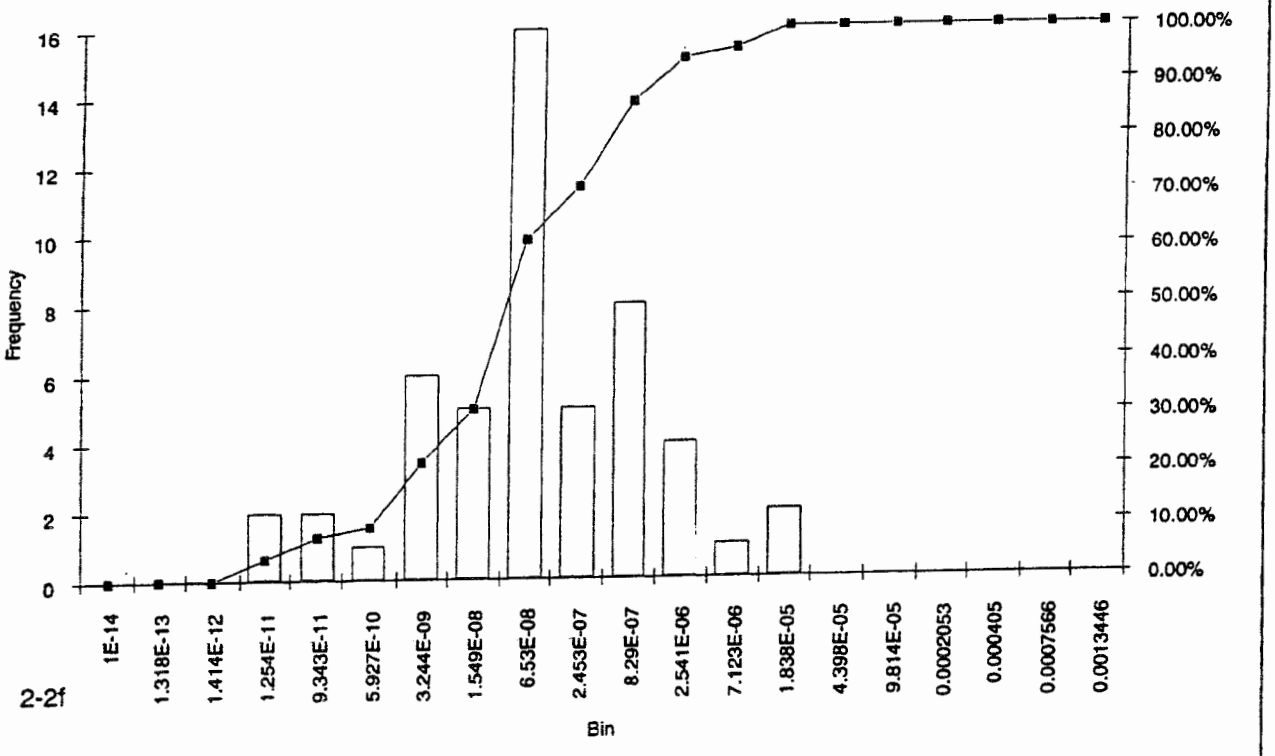
HISTOGRAMS AND CUMULATIVE HISTOGRAMS OF 3 M PACKER TEST DATA

SKB/GEOSTATISTICS/SWEDEN

PROJECT NO. IP3-1129 DRAWING NO. 41066 DATE 11/7/94 DRAWN BY TR



2-2e



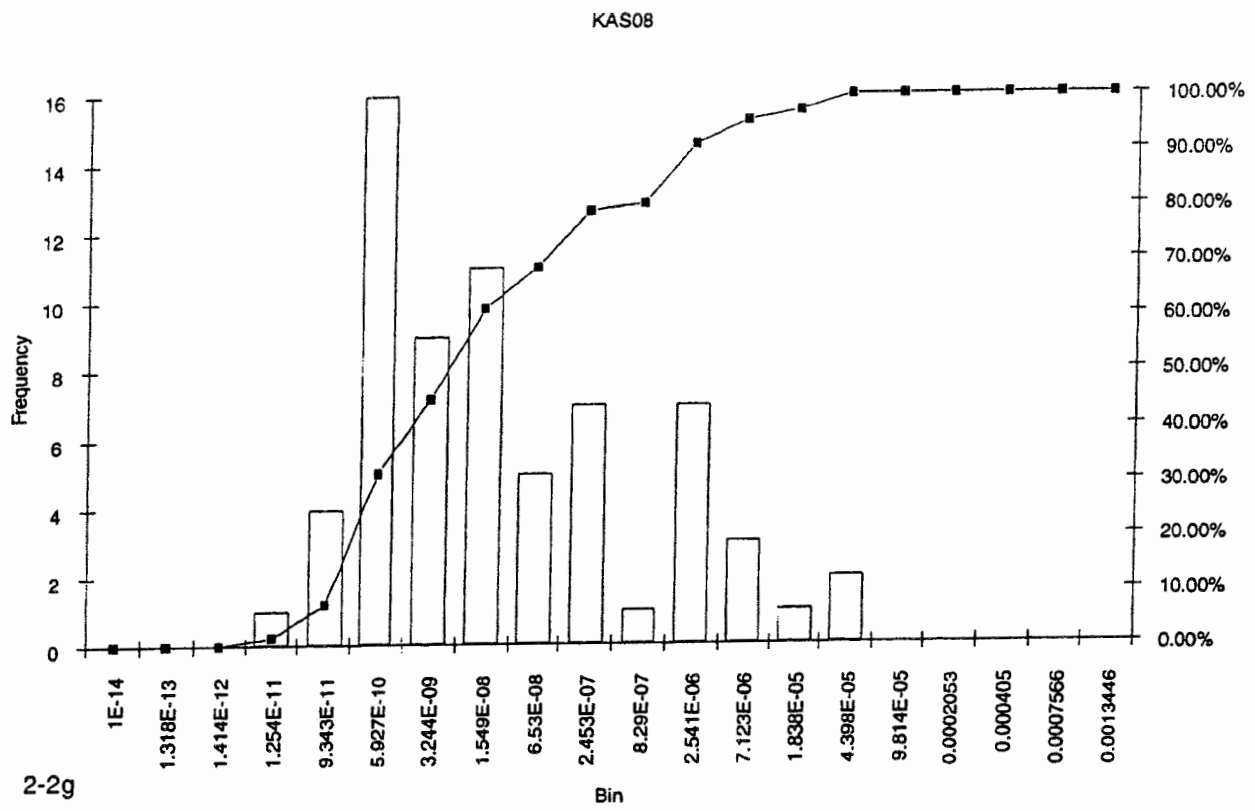
2-2f

FIGURE **2-2**

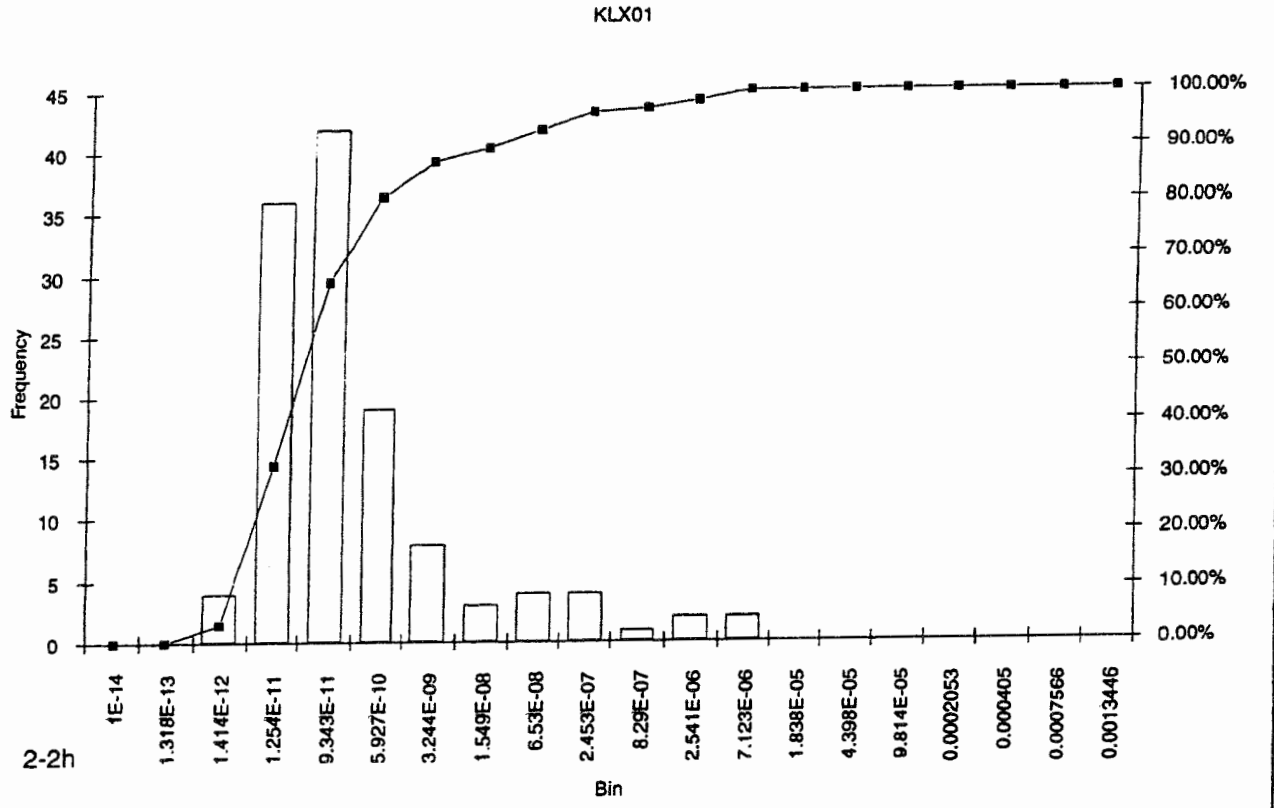
HISTOGRAMS AND CUMULATIVE HISTOGRAMS OF 3 M PACKER TEST DATA

SKB/GEOSTATISTICS/SWEDEN

PROJECT NO. IP3-1120 DRAWING NO. 41067 DATE 11/7/94 DRAWN BY TP



2-2g



2-2h

FIGURE **2-2**

HISTOGRAMS AND CUMULATIVE HISTOGRAMS OF 3 M PACKER TEST DATA

SKB/GEOSTATISTICS/SWEDEN

PROJECT NO. IP3.1139 DRAWING NO. 41068 DATE 11/1/94 DRAWN BY TR

the "Hole Effect" variety, such as the Cardinal Sinus model or the Bessel models. The raw variogram data discussed in Section 3.2 does not indicate that the variograms are strongly nested or well-approximated by a Cardinal Sinus or Bessel model.

2.3 *Trends*

Next, it is necessary to consider trends. Trends are described by global functions that relate the expected value of packer test data to their spatial position. Trends constitute a systematic or deterministic change in the statistics of the data to some underlying geological or spatial component. In folded sedimentary rocks, there may be systematic increases in fracture intensity and conductivity according to structural position, bedding thickness or curvature. At Äspö, the most obvious trend to consider is whether there is any regular change with depth.

Liedholm, (1991; TN8) examined the 3 m packer data to determine whether there were changes in the magnitudes of conductivity measurements as a function of depth. He concluded that the wells showed no systematic change or weak increases or decreases with depth. In light of these results, it does not seem necessary to try to fit a global depth-related trend function to the Äspö data, and perform the geostatistical analyses on the residuals.

The similarity in the cumulative histograms for each well (Figure 2-3) also suggests that there are no systematic areal changes in the sample data.

2.4 *Discussion*

An explanation for the reason why the Äspö packer data may constitute a single population is that fracturing at the Äspö site is a mechanical response of the rock to strain events that were more or less homogeneous at the scale of the site. Although there are mappable geological units, the differences in mechanical properties among

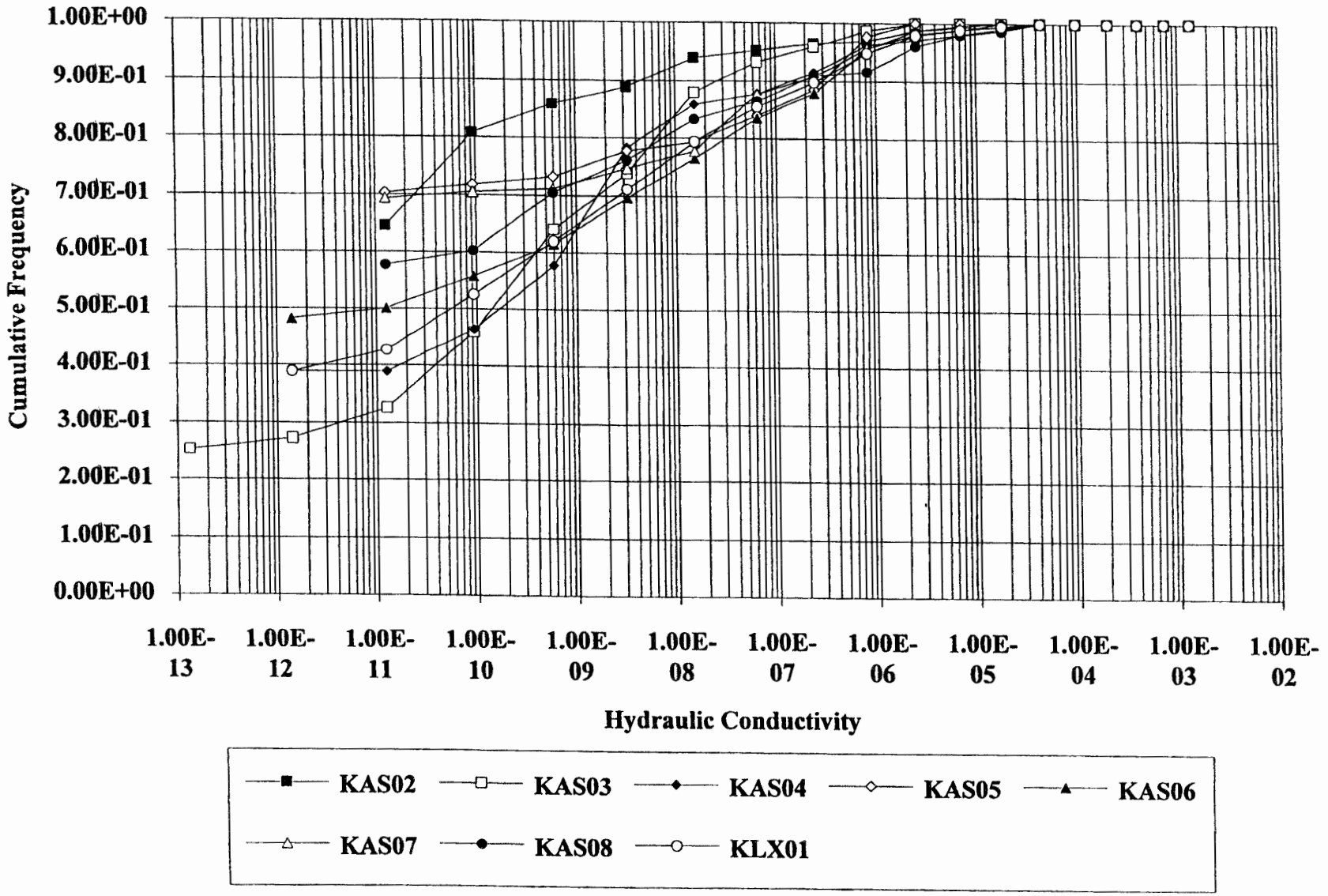


FIGURE **2-3**

NORMALIZED CUMULATIVE FREQUENCY DISTRIBUTION FOR 3 m DATA

them is not sufficient to produce different fracture patterns. The fracture zones are part of the extremities of the natural fracture distribution. These fractures form an interconnected network, in many ways similar to a three-dimensional percolating lattice.

In this model, flow through the Äspö fracture system occurs in the subset of fractures known as the "backbone" of the network. The term backbone is used in percolation theory (e.g. Feder, 1988) to describe the set consisting of all self-avoiding walks from point A to point B. In a network, the backbone is a small subset of the connected lattice, and the majority of fluid flow takes place through this subset. In this situation, one would expect to see the same types of correspondences and lack of correspondences between the spinner logs and the interval tests that one sees at Äspö. A packer test straddling the backbone would have both high conductivity and a spinner anomaly because the fractures are part of the regional well-connected conductive portion of the network. A test of the lattice outside of the backbone could intersect high conductivity fractures that were less well integrated.

These and previous investigations suggest that the 3 m Äspö packer test hydraulic conductivity data can be treated as a single population for geostatistical inference, with the possible exception of the data from KAS02. There appear to be no systematic trends in the data, either with depth or in areal direction. This indicates that the data variability is mostly described by the three stochastic components of local drift, the spatially correlated component and the non-spatially correlated component. Variogram estimation and geostatistical interpolation or simulation are more accurate when the empirical data more closely match the model used to characterize the data. This self-evident statement implies that the greater the proportion of spatially correlated variability is to non-spatially correlated noise, the more effective will geostatistical methods be for estimation. Mixing data from different populations will reduce this proportion, and consequently, the estimation accuracy. Although the 3 m packer test data from KAS02 may not necessarily come from a different population, it has been excluded from all of the analyses reported in

Section 3 in order to avoid the risk of mis-estimating the spatial correlation structure of the data.

3 GEOSTATISTICAL ANALYSIS

3.1 *Overview*

The process of fitting stationary models and non-stationary models follows different paths. Stationary models require estimating parameters for a theoretical model composed of a nugget effect and one or more nested structures, along with anisotropy. Non-stationary models do not consist of estimating parameter values. Rather, it focuses on estimating the *order* of the drift and the determination of which elementary generalized covariance models should be used. Both processes depend upon the type and definition of the Kriging neighborhood.

Section 3.2 describes the evaluation of anisotropy and different neighborhoods, leading to the selection of a single neighborhood and anisotropy model for subsequent cross-validation.

Section 3.3 describes the results of the cross-validation calculations for the stationary models.

Section 3.4 describes the results of the cross-validation calculations for the non-stationary intrinsic models.

Section 3.5 is a general discussion of the comparison between the cross-validation results for the stationary and non-stationary models.

Section 3.6 discusses upscaling of 3 meter tests to 30 m scale tests and their geostatistical properties.

3.2 *Anisotropy and Neighborhoods*

A key problem produced by this typical spatial arrangement of the packer test data is that only the short scale vertical correlation structure and the long scale horizontal correlation structure can be studied directly. This problem is often termed the "data clustering" problem. The maximum vertical or sub-vertical scale that can be robustly studied is on the order of half the distance of the length of a well, which is on the order of 200-400 meters. The horizontal scale is dictated by the spacing of the wells, and for the Äspö data, is on the order of hundreds of meters to kilometers and with many gaps in data over this range. The problem for interpolating well data is not so much that the long distance correlation structure cannot be directly calculated for the vertical direction, but rather that the shorter scale variogram cannot be calculated for the horizontal directions. In Kriging, the neighboring data values play a much more significant role than the distant data values. The way in which data is selected for Kriging, that is, the definition of the Kriging neighborhoods, can improve the clustering problem.

It is possible to test whether some simple assumptions appear valid in order to estimate a reasonable short-scale horizontal variogram structure. The most straightforward way is to determine whether the correlation is isotropic. If this can be shown to be the case, then the short scale vertical variogram should be a reasonable surrogate for the short scale horizontal variogram. This is not a straightforward problem if the data come from non-coaxial wells in which the data have been regularized with overlapping support, see Sec. 3.6. The Äspö wells are definitely not co-axial. For this reason, the anisotropy was studied through analysis of the raw (unregularized) 3 m packer tests.

ISATIS 1.3 contains some very flexible ways to estimate directional variograms that are not found in INFERENS or most other geostatistical software. A common way to estimate directional variograms (for example, Geier, 1993b, p. 15) is to specify and angular and radial tolerance about the desired variogram calculation direction. While

this works well for nearby points, data located hundreds of meters away will lie in sectors of considerable size. For example, a 10 degree horizontal sector 1 km from a data point will include points that extend nearly 100 meters above or below the horizontal (Fig. 3-1a). This can have the effect of including in the directional variogram calculation many data points that are significantly out of the desired direction, which obscures directional correlation properties if they exist. ISATIS 1.3 has a way of mitigating this problem. It allows the user to circumscribe the angular sectors by a three-dimensional box (Fig 3-1b). Points lying outside of the box are never used in the calculation, regardless of whether they belong to the angular sector class under consideration. In other words, a box with a vertical dimension of 10 m would prevent any points more than 5 m above or 5 m below the elevation of the origin from entering into the calculation.

Figures 3-2a-g show directional variograms for the 3 m test data. For these variograms, only the most robust data was used. As described in section 2.2, the most robust data consists of the data above the measurement resolution. Moreover, because well KAS02 may not be part of the same population, all data from this well was excluded from the directional variogram analysis.

The directional variograms shown in Figs. 3-2a-e used a box of dimension 10 m in the two directions perpendicular to the direction of estimation. The horizontal variogram (Fig. 3-2f) constrained the z-dimension to a 10 m window. There was no box for the isotropic variogram shown in Fig. 3-2g. Figure 3-2h shows the number of lag pairs as a function of distance superimposed on the isotropic variogram.

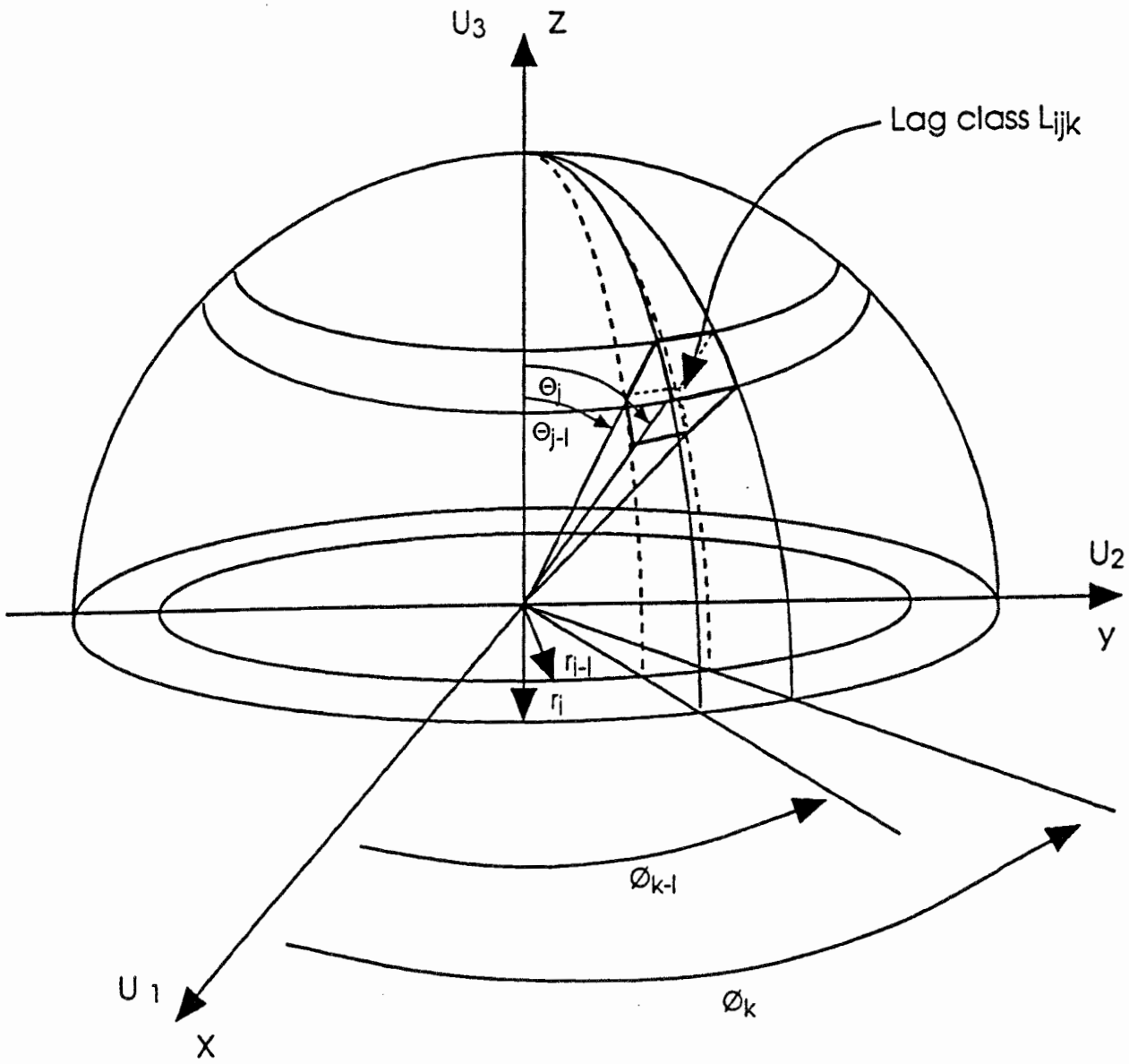


FIGURE **3-1a**
ALTERNATIVE METHODS TO COMPUTE RADIAL LAG PAIRS CONVENTIONAL METHOD
SKB/GEOSTATISTICS/SWEDEN

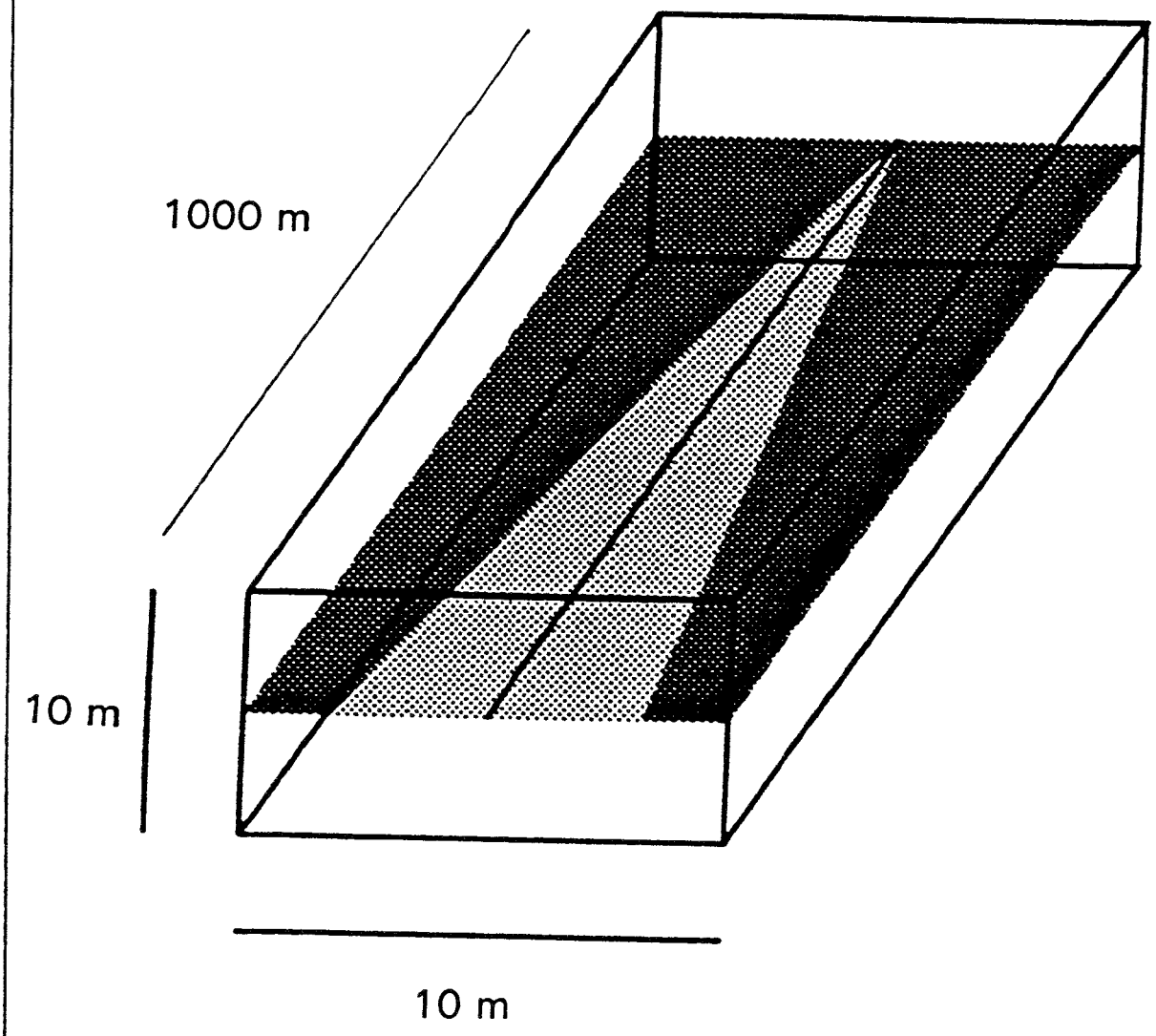


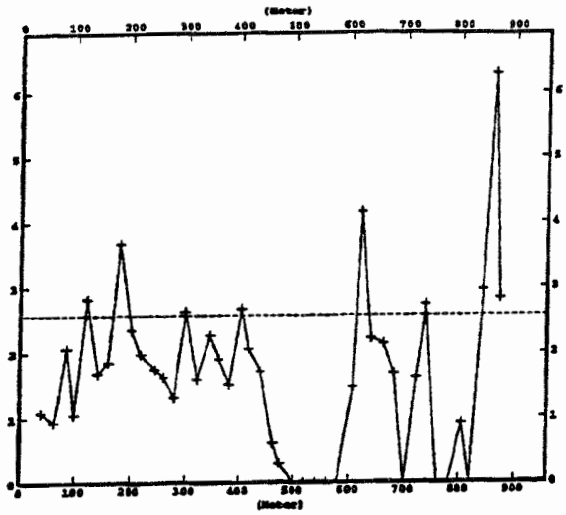
FIGURE **3-1b**

ALTERNATIVE METHODS TO COMPUTE RADIAL LAG PAIRS IMPROVED BOX METHOD

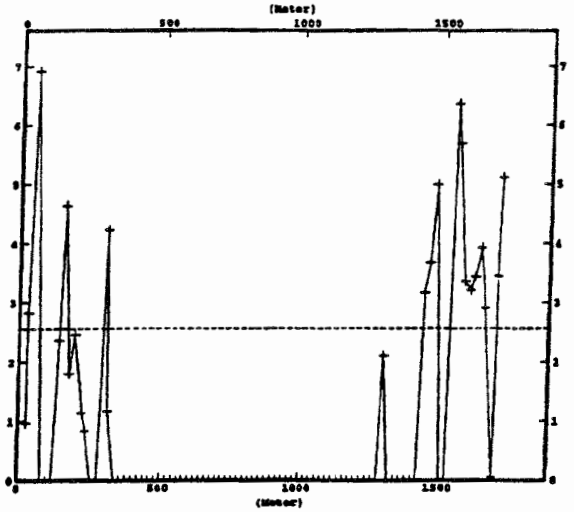
SKB/GEOSTATISTICS/SWEDEN

PROJECT NO. IC3 1129.000 DRAWING NO. 49866 DATE 6/29/94 DRAWN BY CB

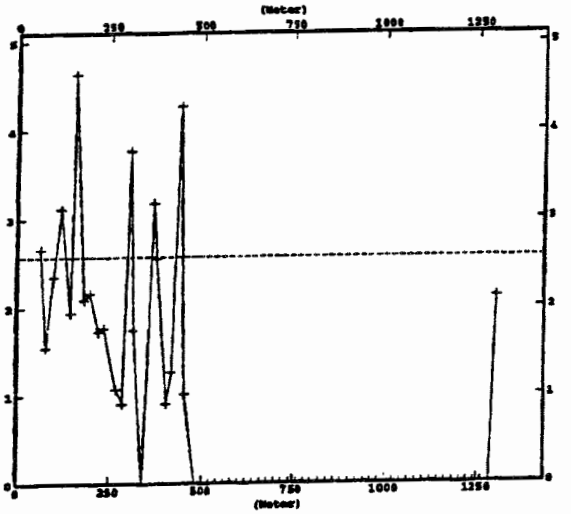
Golder Associates



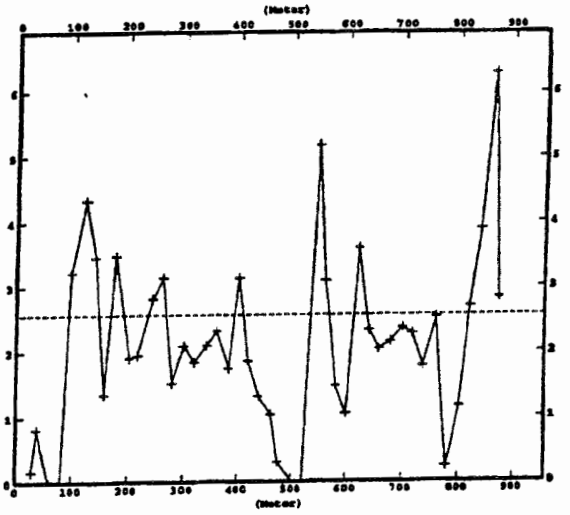
3-2a East/West



3-2b North/South



3-2c Northeast/Southwest



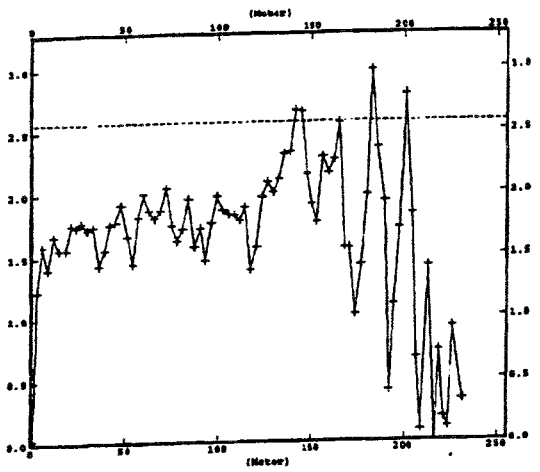
3-2d Northwest/Southeast

FIGURE **3-2**

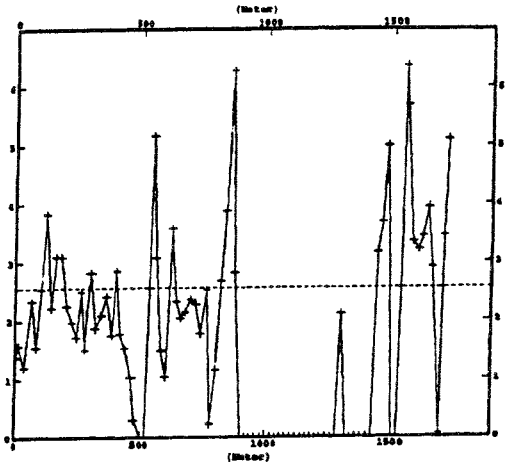
DIRECTIONAL VARIOGRAMS FOR 3 M TEST DATA

SKB/GEOSTATISTICS/SWEDEN

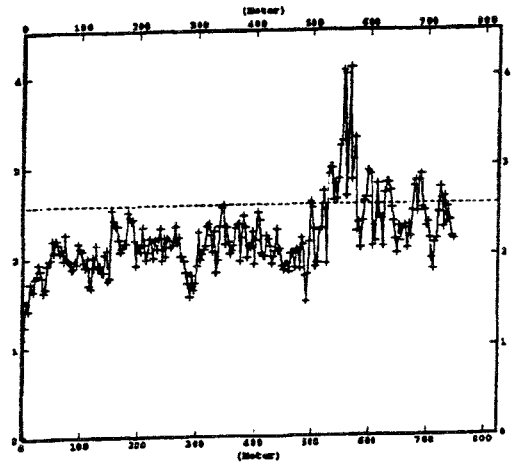
PROJECT NO. IP3 1129 DRAWING NO. 41072 DATE 11/7/94 DRAWN BY TB



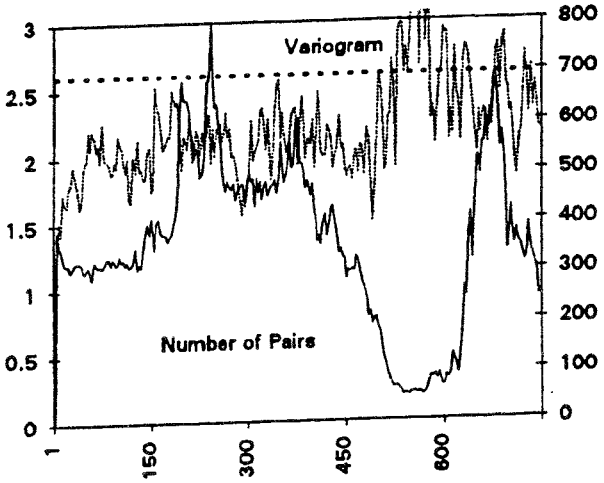
3-2e. Vertical



3-2f. Horizontal



3-2g. All Directions



3-2g. Number of Sample Pairs for All Directions

FIGURE **3-2**
DIRECTIONAL VARIOGRAMS FOR 3M TEST DATA
SKB/GEOSTATISTICS/SWEDEN

PROJECT NO. IC3 1129 DRAWING NO. 41071 DATE 11/7/94 DRAWN BY TB

The vertical variogram (Fig. 3-2e) shows a reasonably well-defined variogram. For distances up to about 125 m, the variogram shows a Nugget Effect (N) around 1.4 and an increasing variogram similar to the Spherical model in that its slope decreases with increasing distance and it appears to reach a sill (C) at a finite distance. The sill is somewhere on the order of 20% of N. This implies that geostatistical methods will not be as advantageous as they might were the proportion of the spatially correlated component higher. The fluctuation in the variogram after 150 m and its decrease is mostly due to the significant decrease in the amount of lag pairs that are used to estimate the variogram. The slight upturn of the variogram around 150 m may reflect underlying correlation structure or the decrease of lag pair data. Based upon the results from horizontal variogram calculations, it is likely that it is in part due to larger scale correlation structure. Without testing any models rigorously, it seems reasonable to fit the raw vertical variogram with a model consisting of a Nugget Effect and a Spherical variogram, which would ignore the upturn around 150 m, or a Nugget Effect and a Gaussian variogram with a long range that captured the upturn better. Another possibility is to fit a nested model with a shorter range Spherical variogram and some appropriate longer range model. However, the lack of significant spatial correlation at distance over 100 m suggests that either model would produce nearly the same results provided that they well-approximate the shorter scale data.

The four horizontal variograms (Figs. 3-2a-d) were for the +X/-X, +Y/-Y, +X+Y/-X-Y and +X-Y/-X+Y directions, referred to subsequently as the East/West, North/South, Northeast/Southwest and Northwest/Southeast directions, respectively. The purpose of calculating these four variograms was to assess any horizontal anisotropy. The problems caused by the horizontal spacing of the wells is quite evident. All four variograms show significant gaps in data and much noise. The North/South and Northwest/Southeast variograms show very little correlation at the scales for which there was data. The Northeast/Southwest variogram shows some correlation, which may be more a result of noise than true structure, since the number of data pairs used to calculate the small distance lag classes in this variogram are

small. The East/West variogram shows the most structure, with a Nugget Effect somewhere on the order of 1.0 to 1.5 and a spherical-like correlation structure with a range of about 200 meters. A global horizontal variogram (Fig. 3-2g) shows a similar structure - a Nugget Effect around 1.3 to 1.7, and a Spherical correlation structure with a range around 200 m. The reason that this horizontal variogram shows the spherical structure is that the short scale lag pairs overwhelmingly come from the East/West variogram. The lack of correlation structure at short scales in the other directional horizontal variograms is quite possibly due to the very small number of data pairs that were used to calculate the variogram, which is graphically demonstrated by the high variability of the raw variograms even at small lag distances.

If this is true, then the one directional horizontal variogram where there are a greater number of data pairs shows very similar variogram structure to the vertical variogram. Both have Nugget Effect somewhere around 1.5, and both rise to sills of around 0.5 reminiscent of a spherical variogram model with a range 200 meters. The variance (σ^2), represented by the dashed line on each variogram is not reached for at least 600 m in the horizontal direction. It is unknown whether the vertical variogram would have reached its final sill at this point or not, since there is no data at these distances. All that can be deduced from the vertical variogram is that the ultimate sill was not reached for distances less than 200m.

The fact that the horizontal and vertical variograms are similar suggests that the correlation is reasonably isotropic. Moreover, there are no obvious geological reasons why there should be correlational anisotropy, as might be expected were Äspö situated in a layered sedimentary rock mass with strong bedding anisotropy. For these reasons, it appears defensible to assume that the correlation structure is isotropic.

Figure 3-2g is the isotropic variogram in which only the distance, not the direction, was used to select lag pairs for the raw variogram. This raw variogram indicates that

there is a sizable Nugget Effect of around 1.4, and a Spherical correlation structure with a Sill about 0.6 and a range on the order of 200 m. At 500 m, there is an increase to the regional Sill (σ^2). The isotropic variogram (Fig. 3-2 g) does not reach the value of σ^2 until about 500 m, although it is no more than 15% below this value for lags pairs greater than 200 m. This suggests that well test data within 500 m is of some limited use in predicting the value of other 3 m well tests because most of the isotropic variogram for the lag pairs between 200 m and 500 m consist of pairs of data from different wells. This contrasts with the variogram lag pairs for distance less than 200 m, which are increasingly dominated by data pairs from the same well. This correlation suggests that there may be some hydrologic features at Aspo that are on the scale of half a kilometer in their horizontal extent. While it might be possible to study the tests forming the lag pairs in this range to investigate this further, they will play a very minor role in spatial interpolation.

Neighborhood definition also plays an important role in stationary and non-stationary geostatistics. There are several considerations and strategies for determining efficient and useful neighborhoods. Use of all of the samples in estimation is called a "Unique Neighborhood". Neighborhoods that do not use all of the available samples, so that different subsets of the data are used for each point or region estimated change, are termed "Moving Neighborhoods". The size of the neighborhood controls how much data will be used to estimate any sample or block. While using more data generally produces greater accuracy, there are pitfalls and pragmatic considerations to limit the amount of data. First, in realistically-sized problems such as the 3 m data at Äspö, there are over 500 well tests that are above the detection limits. If all of these points are used, very large matrices requiring long computer runs are required to perform cross-validation tests, with perhaps, not a great increase in accuracy, since samples more than a few tens of meters from the point to be estimated have negligible weights. Samples at distances greater than the range have no correlation to the point or block being estimated anyway, so the gain in estimation accuracy is small. For these reasons, Unique neighborhoods were not used for the Äspö analyses.

Although it may seem sensible to use the 10 or 20 closest samples for estimation in a Moving neighborhood, this can be a problem for data collected hierarchically (as line samples from wells). Selecting the N nearest samples can lead to using samples only from a single well for cross-validation calculations. If the region to be estimated lies between wells, the estimation becomes more an exercise in extrapolation than interpolation, with all of the inherent pitfalls. It should generally be more accurate to estimate data from samples obtained in several nearby wells, although the selected samples may not be the closest. This problem was evaluated by Norman (1992a) for the Finnsjön data. Possible strategies to deal with this problem include the use of nested models, prohibiting using more than K samples from any one well, requiring that the search neighborhood be divided into angular sectors, or computing cross-validation statistics for more sizable deletion subsets.

There is insufficient data in the Äspö data set to prohibit using more than K samples from a well, so this option was not tested. Thus, it was decided to compare six alternative neighborhoods. One set consisted of a spherical neighborhood divided into 8 sectors, while the other set consisted of a single sector. The neighborhood radius took on three different values: 25 m, 80 m, and 200 m. The comparison of neighborhoods consisted of visually evaluating the locations and weights of samples in these neighborhoods for typical situations in the Äspö rock mass and cross-validation statistics. An accurate and efficient neighborhood is one which contains very few low-weight samples and tends to select points from more than one well when there several wells at approximately the same distance from the estimation point. Kriging weights are a function of the variogram, so it is not possible to optimize the Kriging weights without specifying a variogram model. The weights and subsequent "Delete-1" cross-validation statistics were derived for the optimized non-stationary intrinsic random functions. Figure 3-3 shows the de-clustering effect of using a sectored neighborhood. The simple "Delete-1" cross-validation statistics are summarized in Table 3-1 and Figure 3-4. The error statistics reported in this and following tables follow the definitions shown in Figure 1-2. The results show that the larger the

neighborhood and the more sectors, the better the statistics. The best overall statistics were achieved with a sectored neighborhood and 8 sectors.

Table 3.1

Neighborhood Model	MSE	ME	MRE	MSRE
25 m 1 Sector	38.05799	0.10499	2.469524	6.098547
25 m 8 Sector	21.83913	0.00549	0.748343	0.560018
80 m 1 Sector	5.419145	0.00563	0.818523	0.66998
80 m 8 Sector	5.440016	0.00229	0.814019	0.662627
200 m 1 Sector	5.361084	0.00043	0.823135	0.677551
200 m 8 Sector	1.900812	-0.0048	0.814034	0.662651

3.3 Cross-Validation Results of Stationary Geostatistical Models

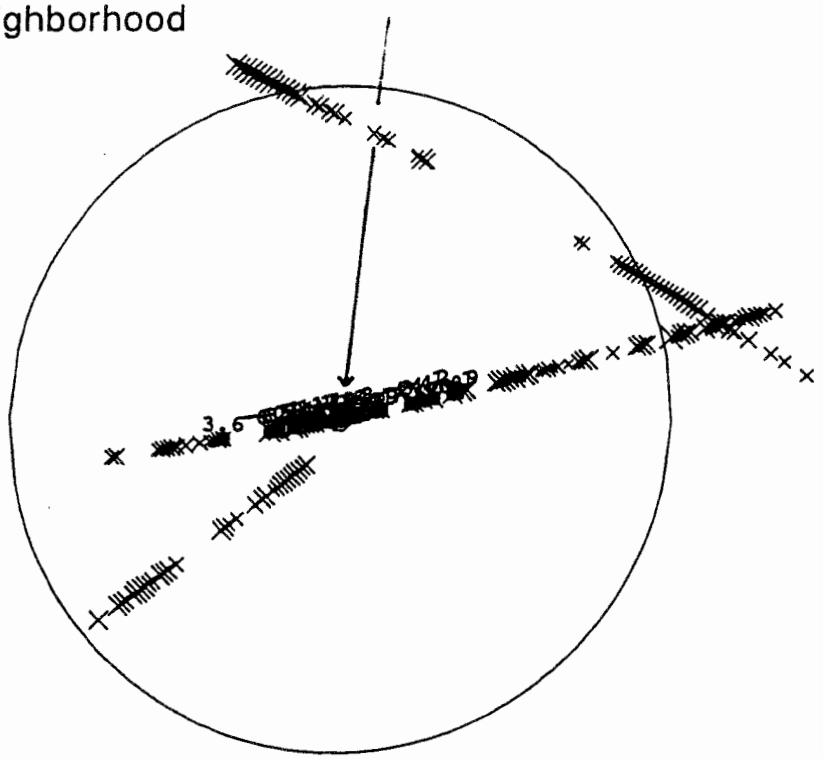
3.3.1 *"Delete-1" Cross-Validation Results*

Inferens 1.1 supports simple or nested variogram models. The nested models may consist of a combination of spherical and exponential functions (Geier, 1993b). Nugget effects are not allowed in the current release, but the raw 3 m variograms suggest that a nugget effect is a prominent constituent of the Äspö data (Fig. 3-2g), making up over 60% of the observed variance. Thus the stationary models evaluated consisted of alternative combinations of a Nugget effect, a Spherical and an Exponential variogram. Alternative models were first evaluated using the "Delete-1" approach. The model that had the best cross-validation, a combination of a Nugget Effect, a medium range spherical variogram and a nested exponential variogram of longer range was then used in "Delete-10%" and "Delete-50%" cross-validation exercises. The fitted variograms are shown in Figures 3-5a-d.

The results for the first set of cross-validation exercises, the "Delete-1" tests, are summarized in Table 3.2 and Table 3.3 and Figure 3-6. The results for the various combinations of nested structures are not entirely consistent; in some cases, one model is best in terms of one statistic but not as good in terms of another statistic.

Points selected from only 1 well with a single sector neighborhood

a



Points selected from multiple wells in a 4-sector neighborhood

b

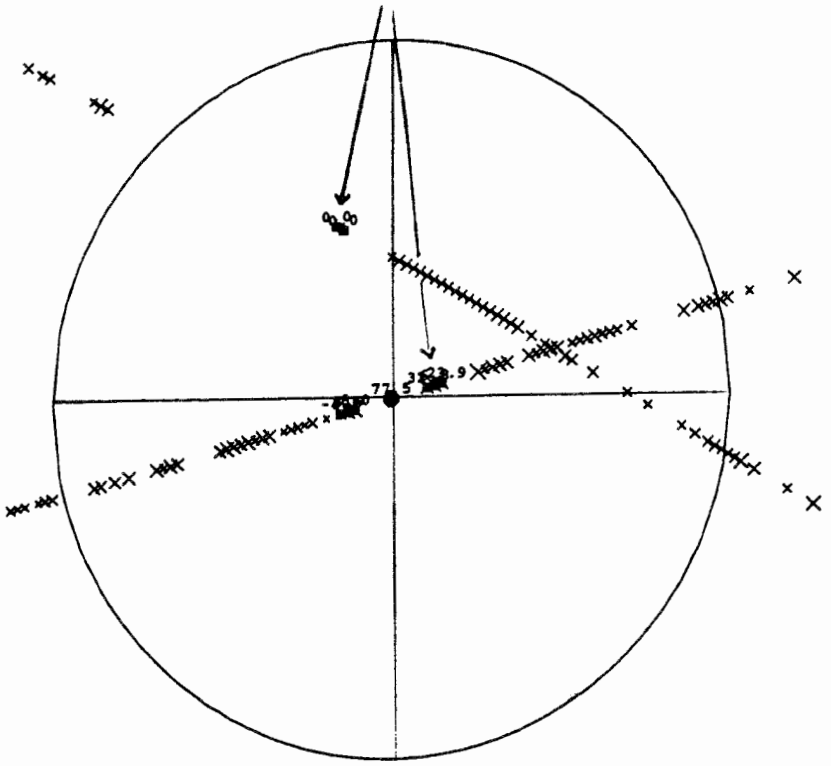


FIGURE **3-3**

ADVANTAGES OF USING A SECTORED NEIGHBORHOOD TO INCLUDE POINTS FROM MULTIPLE WELLS 4 SECTOR NEIGHBORHOOD

SKB/GEOSTATISTICS/SWEDEN

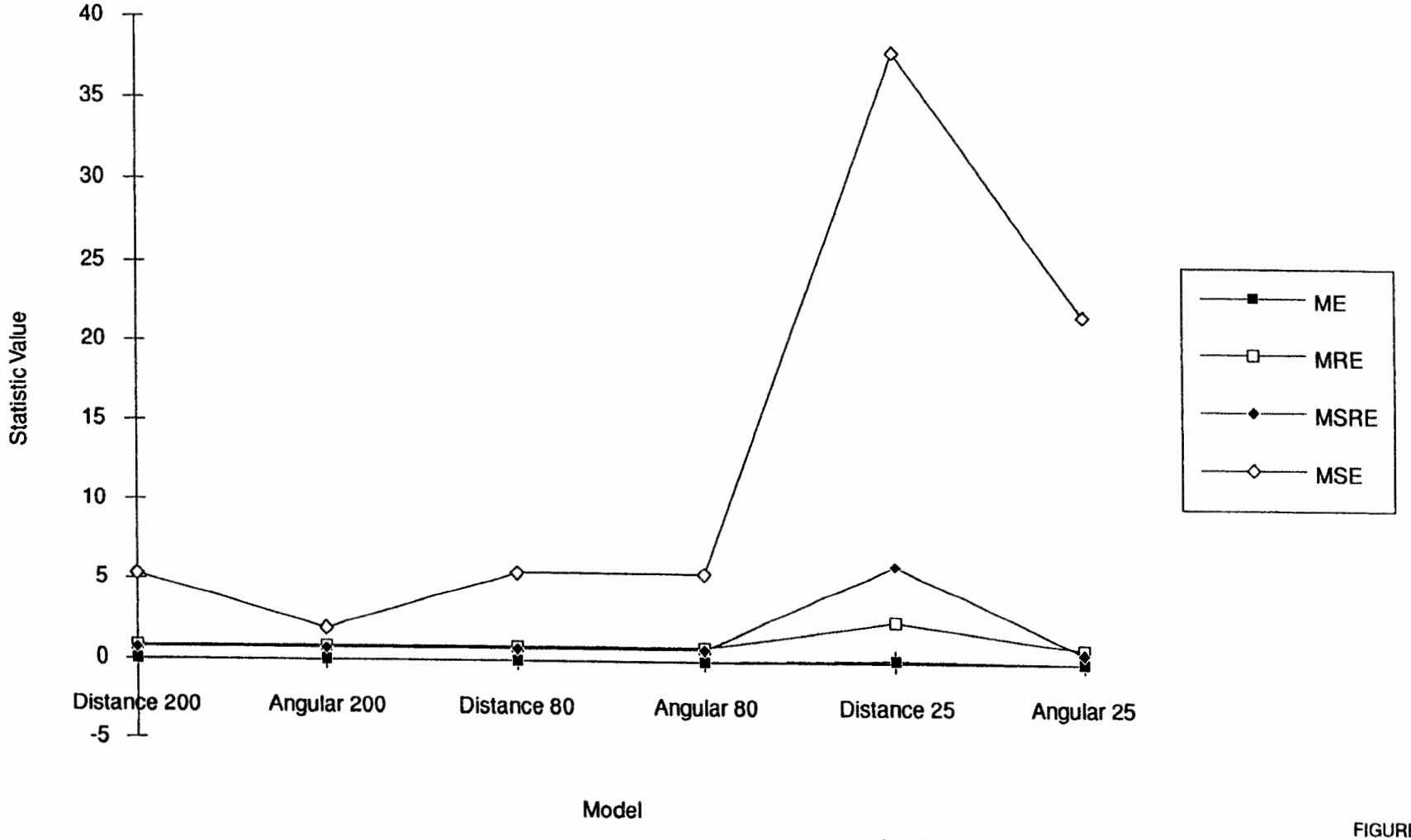


FIGURE **3-4**

"DELETE-1" CROSS-VALIDATION RESULTS FOR ALTERNATIVE NEIGHBORHOOD DEFINITIONS USING THE VARIOGRAM OF FIGURE 3-3

SKB/GEOSTATISTICS/SWEDEN

Table 3.2

	N	Range #1	Sill #1	Range #2	Sill #2
Nugget + Exponential	1.4866	200	0.9512		
Nugget + Spherical	1.4453	70	0.7654		
Nugget + Spherical + Exponential	1.2523	50	0.6802	200	0.2981
Nugget + Spherical + Spherical	1.2507	50	0.6827	200	0.2114

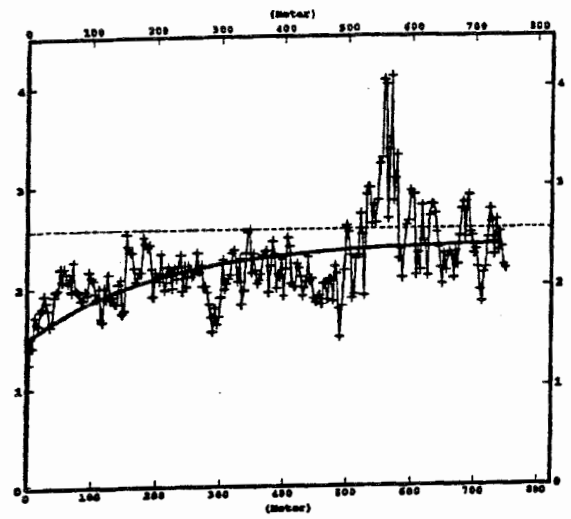
Table 3.3

	MSE	ME	MRE	MSRE
Nugget + Exponential	1.730457	-0.02092	1.014764	1.029747
Nugget + Spherical	1.72695	-0.01795	1.02202	1.044525
Nugget	1.687413	-0.01416	1.027333	1.055412
Nugget + Spherical + Spherical	1.688033	-0.01369	1.027503	1.05576

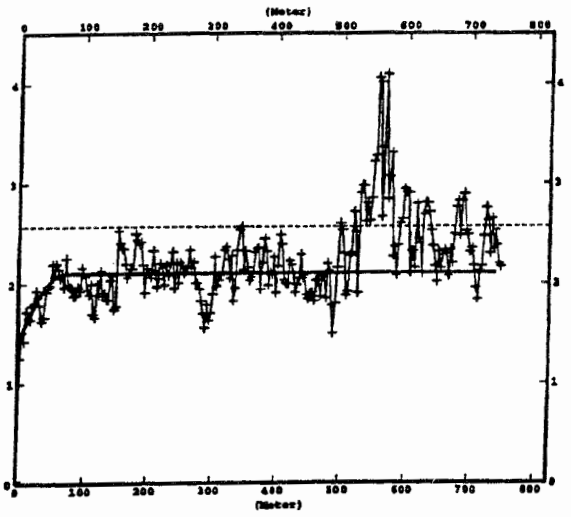
The four stationary models are not very different from each other in terms of the cross-validation statistics reported in Table 3.3. The model with the lowest MSE is the nested Nugget-Spherical-Exponential model, which also has the second lowest ME. The simpler models, consisting of a Nugget plus either a Spherical or Exponential covariance model produce slightly better MRE or MSRE statistics. There was insufficient time to exhaustively test all four of the models in the remaining two cross-validation tests. The basis for selecting a model was somewhat arbitrary. Only the Nugget +Spherical and the Nugget+Spherical+Exponential models were not the worst in any statistical category. Moreover, the Nugget+Spherical+ Exponential scored the best in the MSE category. In addition, although cross validation results for "Delete-1" testing may not depend much upon the accuracy of the model's fit to long distance lag pairs, such a match may improve the results from the more demanding "Delete-50%" cross-validation (Sec. 3.3.3). As a result, it was chosen for further cross-validation tests.

3.3.2 *"Delete-10%" Cross-Validation Results*

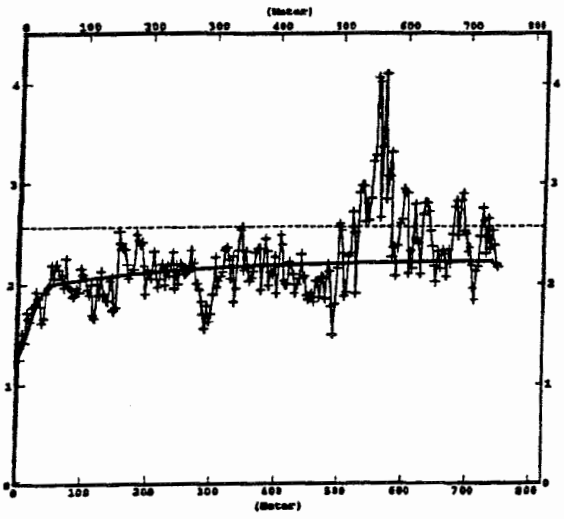
The "Delete-10%" test results are summarized in Table 3.4 and Figure 3-7.



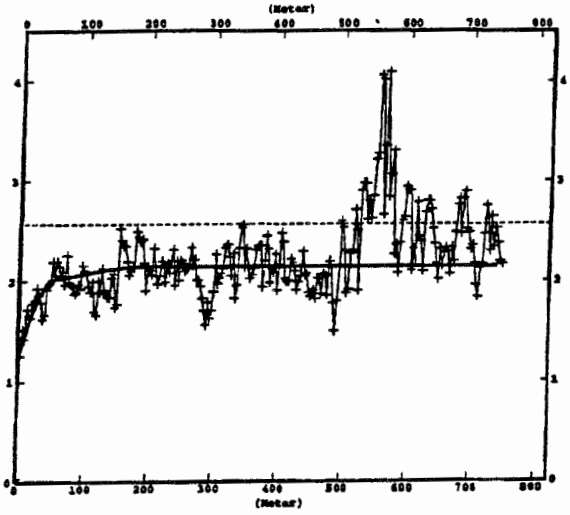
3-5a Nugget + Exponential



3-5b Nugget + Spherical



3-5c Nugget + Spherical (short range)
+ Exponential (long range)



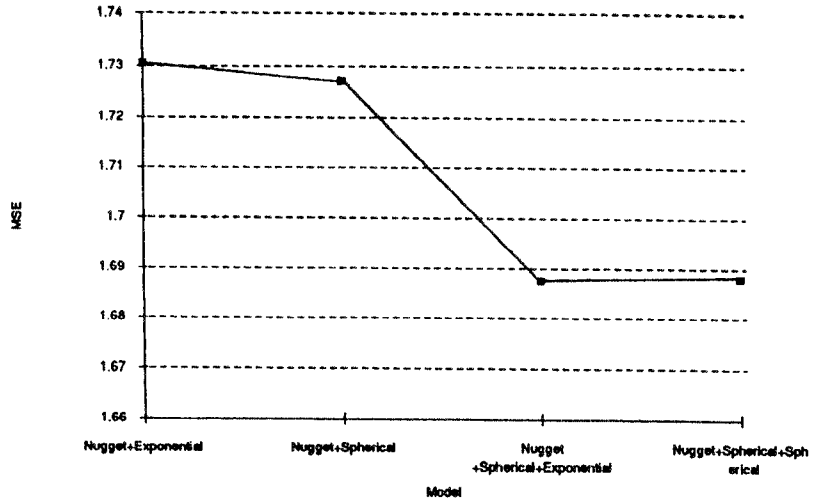
3-5d Nugget + Spherical (short range)
+ Spherical (long range)

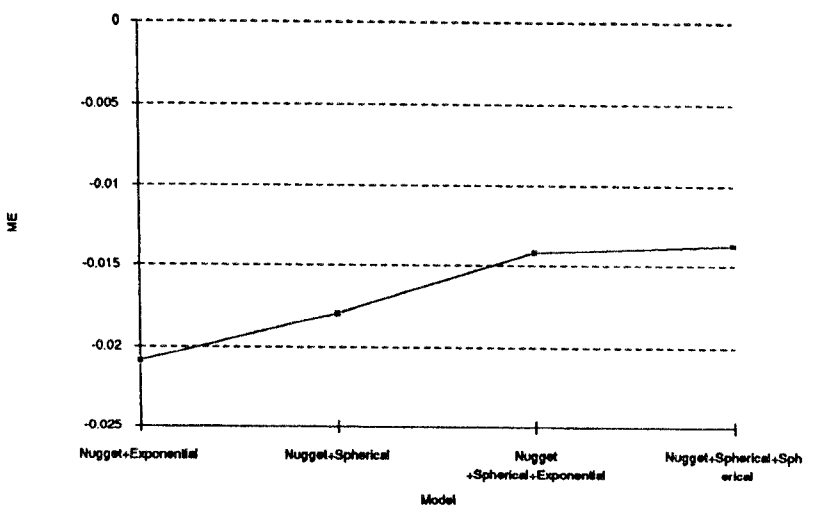
FIGURE **3-5**

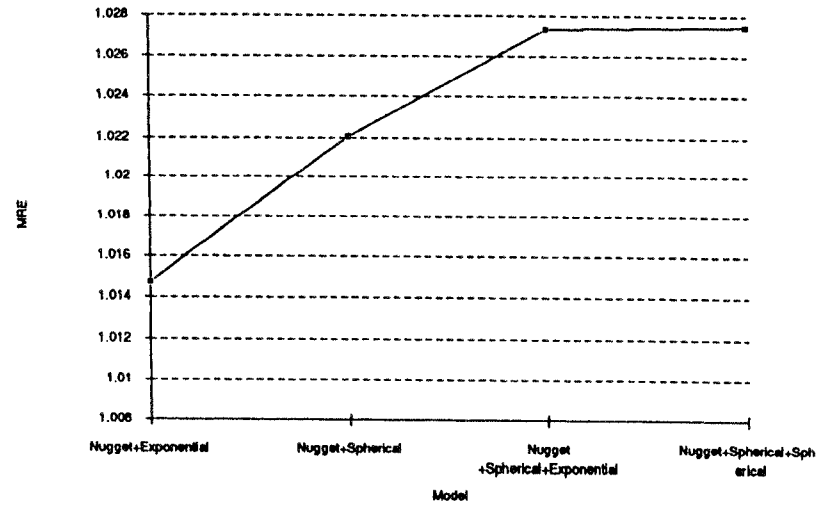
FITTED STATIONARY VARIOGRAMS FOR THE "DELETE-1" CROSS VALIDATION TESTS

SKB/GEOSTATISTICS/SWEDEN

PROJECT NO. IP3 1129 DRAWING NO. 41073 DATE 11/8/94 DRAWN BY TB







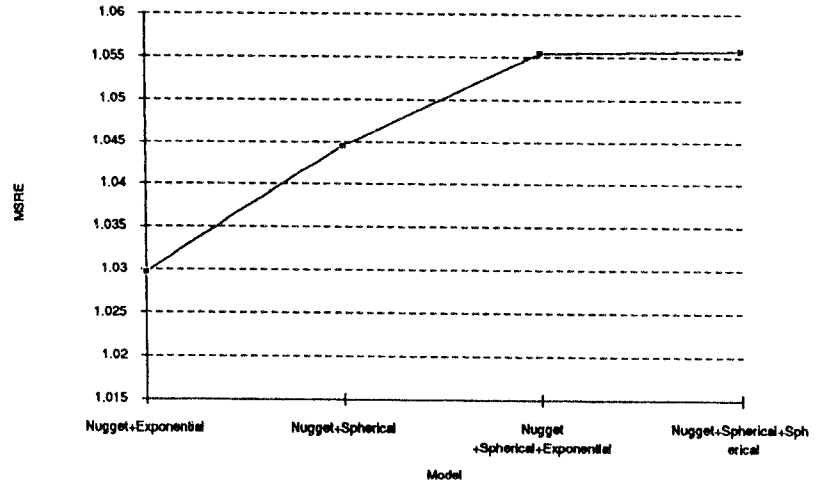


FIGURE **3-6**

TEST STATISTICS FOR "DELETE-1" CROSS-VALIDATION OF FOUR STATIONARY MODELS

SKB/GEOSTATISTICS/SWEDEN

PROJECT NO. IC3 1129.000 DRAWING NO. 49881 DATE 6/29/94 DRAWN BY CB

Golder Associates

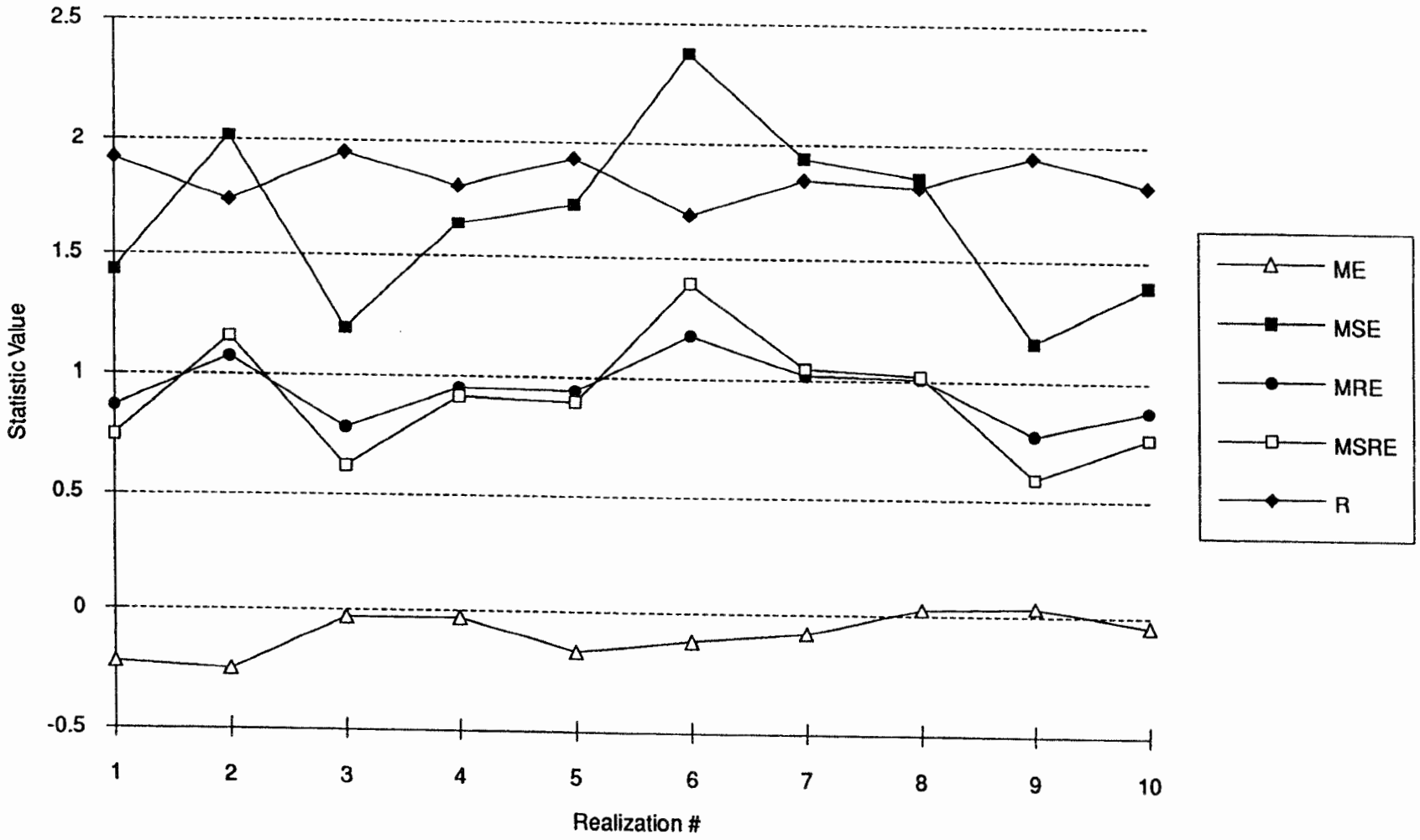


FIGURE **3-7**

TEST STATISTICS FOR 10 REALIZATIONS OF THE "DELETE-10%" CROSS-VALIDATION OF THE STATIONARY MODEL

SKB/GEOSTATISTICS/SWEDEN

PROJECT NO. IC2 1129.000 DRAWING NO.

Table 3.4

Realization #	MSE	ME	MRE	MSRE	R
1	1.434909	-0.21293	0.864587	0.74751	1.91958
2	2.015544	-0.24462	1.077501	1.161008	1.73603
3	1.203644	-0.02255	0.785619	0.617197	1.95017
4	1.650222	-0.02423	0.954431	0.910939	1.81156
5	1.730329	-0.15822	0.945872	0.894674	1.93403
6	2.375166	-0.11264	1.184028	1.401922	1.69422
7	1.935587	-0.07275	1.023074	1.046681	1.84926
8	1.856363	0.029523	1.011035	1.022191	1.81606
9	1.157548	0.044404	0.771455	0.595143	1.94499
10	1.397565	-0.03205	0.874265	0.764339	1.82846
Average	1.675688	-0.08061	0.957163	0.91616	1.84843

The mean error (ME) should theoretically be 0.0 for a perfect model. The value of ME for the ten realizations varies from -0.24 to +0.04, with an average of -0.08061. MRE, MSRE and R should all equal 1.0 for a perfect model. MRE varies from 0.77 to 1.18, with an average of 0.96. MSRE varies from 0.60 to 1.41, with an average of 0.92, while R varies from 1.74 to 1.95, with an average of 1.85. Recall that for this type of cross-validation, the R statistic is biased, so it probably less accurately reflects the model's usefulness. MSE varies from 1.16 to 2.4, with an average of 1.68. For a perfect model, this error would be 0.0. Figure 3-7 shows that the statistics can vary from realization to realization, and it is hoped that the average values for the ten realizations more accurately reflect the model's estimation usefulness.

3.3.3 *"Delete-50%" Cross-Validation Results*

The final set of cross-validation calculations consisted of randomly dividing the packer data into two equal subsets. A model fitted to the first subset was used to estimate the second and vica-versa. This procedure produces two sets of error statistics for each realization. Renard (1988, pg. 586) shows how the biased estimate of the R statistic can be reduced in this situation according to the formula:

$$\rho = 2r - \frac{n_1 r_1 + n_2 r_2}{n_1 + n_2} \tag{3.1}$$

where r_1, r_2 are the ratios estimated for subsets 1 and 2,

n_1, n_2 are the number of data points in subsets 1 and 2

r is the biased estimate of the ratio obtained by combining both subsets, and

ρ is the jackknife estimator with reduced bias.

Table 3.5 and Figure 3-8 summarize the results of the cross-validation tests run on ten realizations of the 3 m packer test data. The values reported in the table for each realization are the average values calculated for the two subsets, with the exception of the R statistic, which was calculated according to Eq. 3.1.

Table 3.5

Realization #	MSE	ME	MRE	MSRE	R
1	1.838127	-0.00379	0.994006	1.001479	1.772505
2	1.869077	-0.05514	1.047085	1.098143	1.692308
3	1.815765	-0.01489	1.002333	1.007383	1.798029
4	1.834063	0.003526	1.003378	1.007758	1.813946
5	1.783045	-0.01547	1.027333	1.068933	1.622201
6	1.902242	-0.04205	1.011991	1.035853	1.791049
7	1.788827	-0.02121	0.974057	0.94884	1.88554
8	1.818458	-0.03619	1.018813	1.039138	1.743592
9	1.774343	-0.01155	1.026962	1.056559	1.672293
10	1.961931	-0.03841	1.07977	1.174817	1.634208
Average	1.838588	-0.02352	1.018573	1.04389	1.742567

Interestingly, the statistics for the "Delete-50%" case are generally better and less variable than for the "Delete-10%" case. The MSE is the only statistic that did not

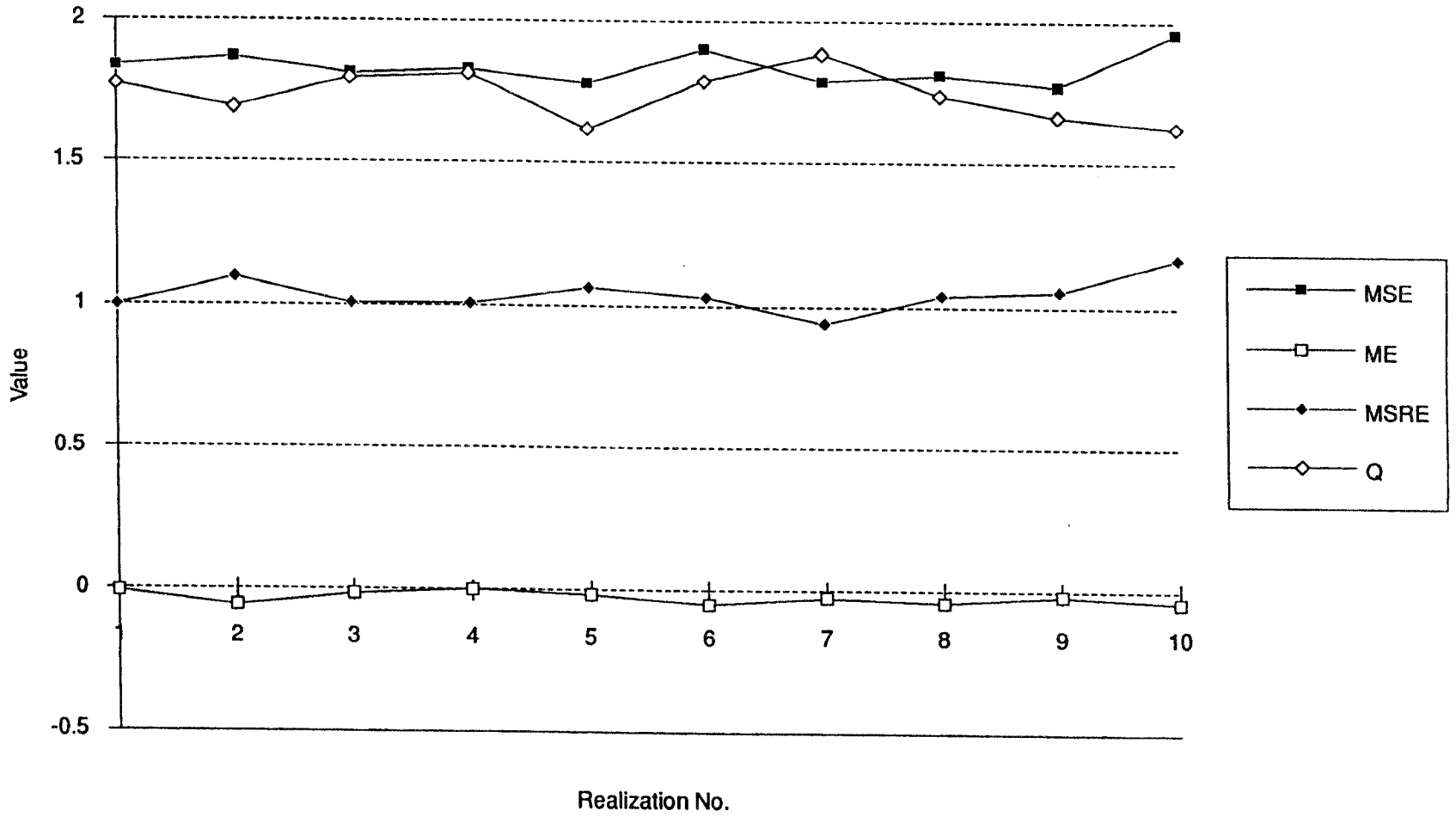


FIGURE **3-8**

TEST STATISTICS FOR 10 REALIZATIONS OF THE "DELETE-50%" CROSS-VALIDATION OF THE STATIONARY MODEL

SKB/GEOSTATISTICS/SWEDEN

PROJECT NO. IC3 1129.000 DRAWING NO. 49883 DATE 6/29/94 DRAWN BY CB

Golder Associates

improve. The reasons for this are not clear, but the lower variability is possibly due to the averaging of the statistics for the two subsets in each realization.

3.4 Cross-Validation Results of Non-Stationary Geostatistical Models

3.4.1 *"Delete-1" Cross-Validation Results*

Geostatistical interpolation or conditional simulation requires specification of three components: the drift, the correlation model and the neighborhood. In this context, drift is used to describe the form of the "large" amplitude fluctuations in the data in the local region surrounding the data. In this sense, it is distinct from the concept of "trend", which is a model to approximate the global large amplitude structure of the data. The drift thus defined corresponds to the amplitude structure of the neighborhood used for Kriging or interpolation. The correlation model is the smaller scale model that describes the spatial structure of the data after the drift has been removed. Thus, the drift, the definition of the neighborhood and the variogram models are all interrelated and not separable.

The drift function for intrinsic random functions of order k (IRF-k) are represented by polynomials of different degree Renard (1989). A zero-th order drift function is a constant that represents the local neighborhood's large amplitude fluctuations. A first order drift function is one in which these fluctuations are modeled by a linear combination of a constant term and the local X, Y or Z coordinates. A second order drift function includes the additional terms relating to X^2, Y^2, Z^2, XY, XZ, YZ or some subset of them. A complete drift function consists of a linear combination of all terms of that order and lesser orders. In other words, all six second-order terms, all three first-order terms and the constant term would appear in a complete second-order drift function. Unless there are compelling geological arguments, such as a strong and evident grain in the correlation structure of the data, it is probably best to always use complete drift functions. For purposes of this study, only complete functions

were used. Four drift functions were evaluated: zeroth-, first-, second- and third-order polynomials.

Each drift function can be evaluated in terms of mean squared error (MSE) and a relative ranking measure based on the MSE. ISATIS 1.3 employs the following procedure for evaluating drift models.

A variety of test points are selected according to an algorithm which tries to cover the region as regularly as possible, to de-clusterize clusterized information, and also to include isolated points. The goal is to test with as representative points as possible. Next, the samples belonging to the neighborhood for each test point are determined. The points in the neighborhood are sorted with respect to their distance from the test point, and then split into two subsets, S1 and S2. All of the odd numbered points go into one subset; all of the even numbered points into the other. This helps ensure that each subset is an equally valid independent realization of the neighborhood. The selected drift models are fitted to each subset of data using standard least-squares techniques. Each drift model fitted to the S1 data is then used to predict the samples in S2, and each drift model fitted to data in S2 is used to predict samples in S1. The MSE is computed for each drift model and for all of the test points. The drift model that produces the lowest overall MSE is not necessarily the best choice, since these errors are highly sensitive to anomalies. To provide a ranking procedure more robust against outliers, ISATIS 1.3 employs an ordinal ranking procedure for each test point. In other words, the MSE's are computed for all of the drift models at a particular test point. Then, the drift models are ranked according to their respective MSE's. This is done for all test points and the overall average ranks for each model are computed.

The drift analysis results are shown in Table 3.6. The six combinations selected for drift analysis correspond to the neighborhoods previously discussed. Since the stationary geostatistical analysis revealed a very short range structure, another with a range from 50 to 80 m, and a third with a range greater than 200 m, these radial distances were selected for drift evaluation. Each distance neighborhood was also

divided into a single sector and an 8 sector neighborhood. Table 3.6 shows that the size and type of neighborhood effects which order of drift is optimal. Short range models favor a linear or zero-th order drift function. Longer range neighborhoods favor first or second order drift functions. In the cross validation results, single sector neighborhoods have an insignificant tendency to favor lower order drift.

Table 3.6

Order of Drift				
Neighborhood Model	0	1	2	3
25 m 8 Sector	1.34	1.28	1.46	1.92
25 m Single Sector	1.34	1.27	1.47	1.92
80 m 8 Sector	1.56	1.47	1.38	1.59
80 m Single Sector	1.53	1.46	1.41	1.6
200 m 8 Sector	1.59	1.52	1.43	1.47
200 m Single Sector	1.55	1.46	1.46	1.53

It is worthwhile to point out that the optimal drift is a function of the actual data used in the cross-validation, independent of the neighborhood model. As a result, different "Delete-10%" and "Delete-50%" realizations may produce different optimal orders of drift than the "Delete-1" tests.

The algorithms for determining optimal models for different intrinsic variogram models are reasonably complex, but are explained clearly by Renard (1989). For the Moving Neighborhood, the general idea of the algorithm is to select a subset of sample points as test points. One or more elementary covariance models are selected for testing either as separate models or as linear combinations. The goal is to calculate model parameters that produce the smallest jackknifed regression statistics in terms of ordinal rankings. The statistic used to evaluate a model consists of the ratio

between the experimentally calculated variance and the theoretical variance for the data subsets. In practice, the variances are not known, but are estimated from the samples, leading to a biased estimate of the ratio. To reduce this bias, the jackknife estimator of the ratio is computed according to equation 3.1. The covariance model that consistently produces the lowest ranked jackknife estimate is the model selected.

Table 3.7 summarizes the test statistics for the "Delete-1" cross-validation. Comparing these statistics with those in Table 3.3 shows that the stationary models all produce better MRE, MSRE and MSE statistics but worse ME values. Overall, the "Delete-1" cross-validation results strongly favor the stationary models over the non-stationary models.

Table 3.7

MSE	ME	MRE	MSRE
1.90	-0.0048	0.81	0.66

3.4.2 *"Delete-10%" Cross-Validation Results*

The results of the automated drift calculations for the ten realizations (Table 3.8) tend to favor a linear drift model slightly over a zero-order model. Regardless of the selected order of drift, the most frequently selected covariance model to go with the drift model is one of zero order.

Table 3.8

	Order of Drift Ranking			Selected Drift	Optimal
Realization #	0	1	2		
1	1.01	0.98	1.02	1	N
2	1.04	0.95	1.01	1	N
3	0.99	0.99	1.03	0	N+3rd Order
4	1.01	0.93	1.05	1	N
5	1.02	1.02	0.97	1	N
6	0.95	1.02	1.02	0	N
7	1	0.94	1.06	1	N
8	0.99	0.99	1.02	1	N
9	0.99	1.01	1	0	N
10	1.01	0.92	1.06	1	N
Average	1.001	0.975	1.024	1	N

*N = Nugget Effect; G.C. = Generalized Covariance

Cross validation results for the "Delete-10%" tests are summarized in Table 3.9 and Figure 3-9. For this neighborhood definition, the optimal order of drift was generally of zeroth or first order, and the covariance consisted of zeroth or third order intrinsic functions. The mean error (ME) and R statistics are better than for the stationary models, but the MSE, MRE and MSRE are not as good. Even without additional statistical tests, it is clear that the MRE, MSRE and R statistics indicate that the model would be rejected with a high level of confidence.

Table 3.9

Realization #	ME	MSE	MRE	MSRE	R
1	-0.16031	1.799445	1.164301	1.355597	1.327419
2	-0.15286	2.486622	1.353156	1.831032	1.358044
3	-0.04848	1.211047	1.067958	1.140534	1.061825
4	0.067297	1.711717	1.202929	1.447038	1.18291
5	-0.22473	2.603512	1.16651	1.360746	1.913297
6	-0.22018	2.498767	1.53355	2.351777	1.062502
7	-0.05538	2.507085	1.206239	1.455013	1.723067
8	0.086772	2.071585	1.301154	1.693001	1.223617
9	0.131666	1.360024	1.131381	1.280022	1.062501
10	0.066036	1.859356	1.173321	1.376681	1.350607
Average	-0.05102	2.010916	1.23005	1.529144	1.326579

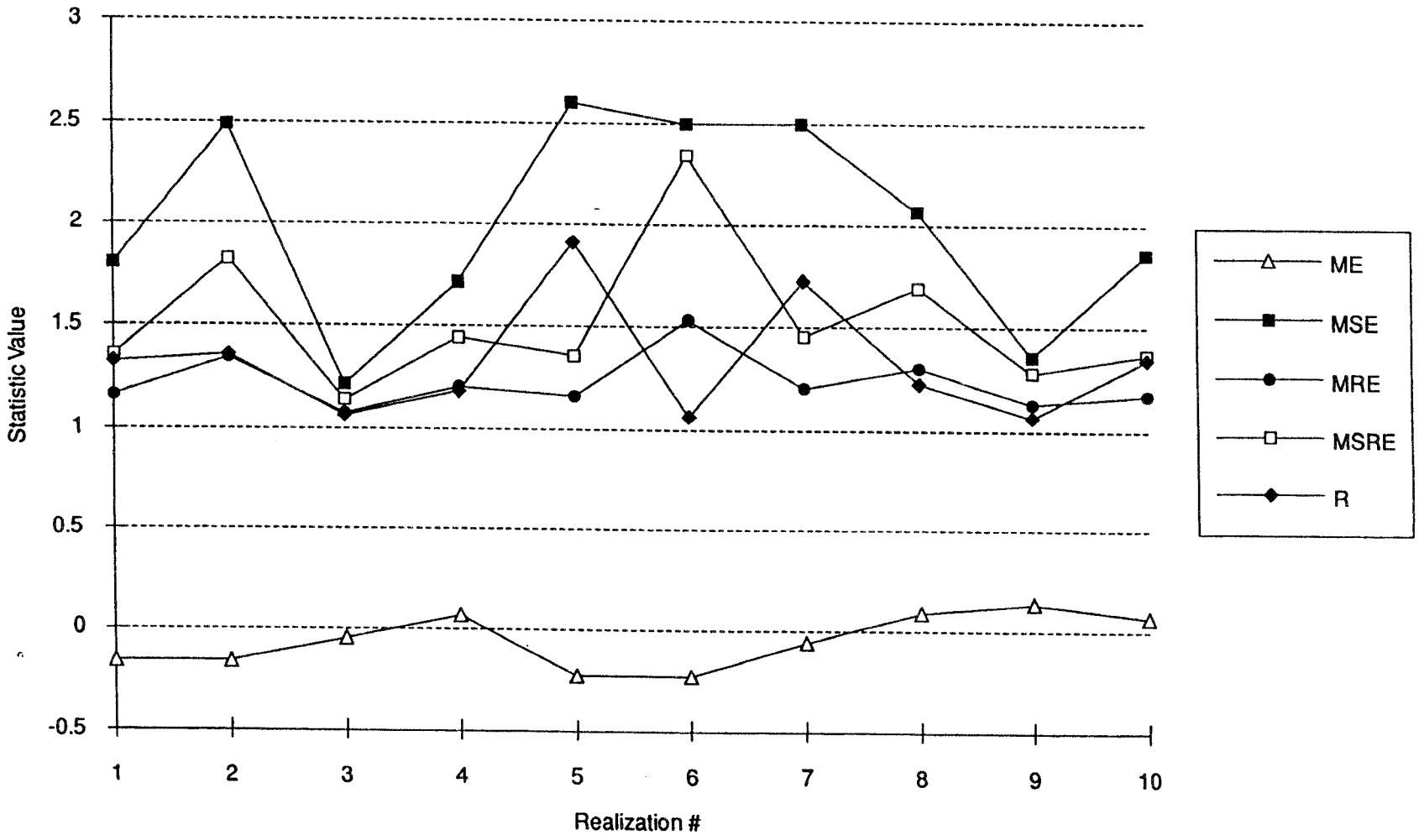


FIGURE **3-9**

TEST STATISTICS FOR 10 REALIZATIONS OF THE "DELETE-10%" CROSS-VALIDATION OF THE NON-STATIONARY MODEL

SKB/GEOSTATISTICS/SWEDEN

PROJECT NO. IC3 1129.000 DRAWING NO. 49884 DATE 6/29/94 DRAWN BY CB

Golder Associates

3.4.3 *"Delete-50%" Cross-Validation Results*

The results of the automated drift calculations for the ten realizations (Table 3.10) also favor a linear drift model over a zero-order model. The most frequently selected covariance model to go with the drift model is one of zero order, although in a few instances when the drift selected is zero order, the optimal generalized covariance is a combination of zero order and third order functions. The similarity in the ranking statistics shown in Table 3.10 for a particular realization and among all of the realizations shows that the degree of spatial correlation in the data is not strong. Most of the variability depends upon the specific data used to estimate the model.

Table 3.10

	Order of Drift Ranking			Selected	Optimal Covariance
Realization #	0	1	2		
1a	0.95	0.92	1.13	1	N
1b	1.06	0.93	1	1	N
2a	0.99	0.98	1.03	1	N
2b	1.03	0.9	1.08	1	N
3a	0.97	0.98	1.05	0	N+3rd Order G.C.
3b	0.99	0.94	1.07	1	N
'4a	1.02	1.05	0.93	2	N
4b	1.07	0.96	0.96	1	N
5a	1.06	0.98	0.96	2	N
5b	0.86	1	1.14	0	N+3rd Order G.C.
6a	0.95	0.92	1.13	1	N
6b	1.06	0.93	1	1	N
7a	1.01	0.85	1.14	1	N
7b	1.01	0.97	1.02	1	N
8a	1.02	0.99	0.99	1	N
8b	1.04	0.99	0.97	2	N
9a	1.01	0.97	1.02	1	N
9b	0.99	0.91	1.11	1	N
10a	1.09	0.91	1	1	N
10b	1.02	0.92	1.06	1	N
Average	1.01	0.95	1.039	1	N

N = Nugget Effect; G.C. = Generalized Covariance

Table 3.11 and Figure 3-10 summarize the results of the "Delete-50%" cross-validation calculations for the non-stationary model. As in the case for the stationary model, the statistics are less variable across realizations. The ME improves over the value for the "Delete-10%" tests and is better than for the stationary models. The optimal intrinsic models were similar to those in the previous cross-validation tests in that the order of drift and the intrinsic random covariance function were a zeroth and/or third order. As in the case for the "Delete-10%" tests, the non-stationary models yielded MRE statistics quite close to 0.0. However, the values for MRE, MSRE and R are clearly too high, and would most certainly lead to rejection of the model with high confidence.

Table 3.11

Realization #	MSE	ME	MRE	MSRE	R
1	2.526769	0.085342	1.43726	2.076227	1.994527
2	2.118209	-0.05647	1.372585	1.883995	1.865363
3	2.192901	0.009331	1.400369	1.962779	1.935382
4	2.175224	0.032876	1.359052	1.851688	1.770803
5	2.439993	0.006482	1.493856	2.232286	1.429724
6	2.178163	-0.05461	1.396012	1.951538	1.912692
7	1.973006	-0.01736	1.330871	1.771623	1.741522
8	2.247469	0.032906	1.367303	1.87245	1.817523
9	2.173507	-0.07914	1.398609	1.963565	1.89912
10	2.140087	-0.01949	1.386395	1.922105	1.113406
Average	2.216533	-0.00601	1.394231	1.948826	1.748006

3.5 Discussion of Results

Figure 3-11a-e and Figure 3-12a-e compare statistical results between the stationary and non-stationary models for the "Delete-10%" and "Delete-50%" cases. The cross-validation statistics for stationary models are as good or significantly better than results from non-stationary models for all cross-validation cases. For both models, the ME statistic is near 0.0, suggesting that the errors are unbiased. The stationary model produces MRE and MSRE values much closer to 1.0 than does the non-stationary model. This indicates that the non-stationary model is making

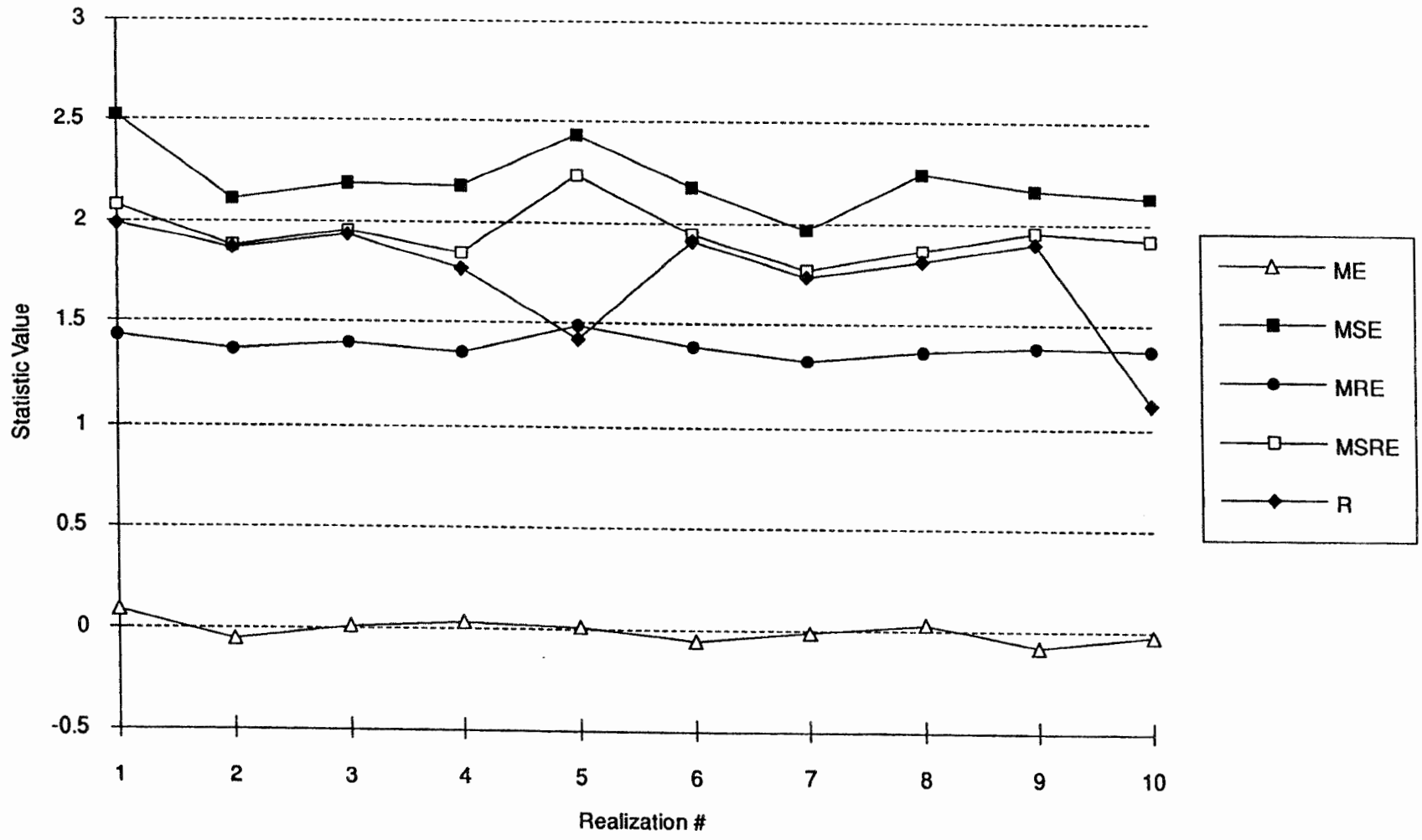
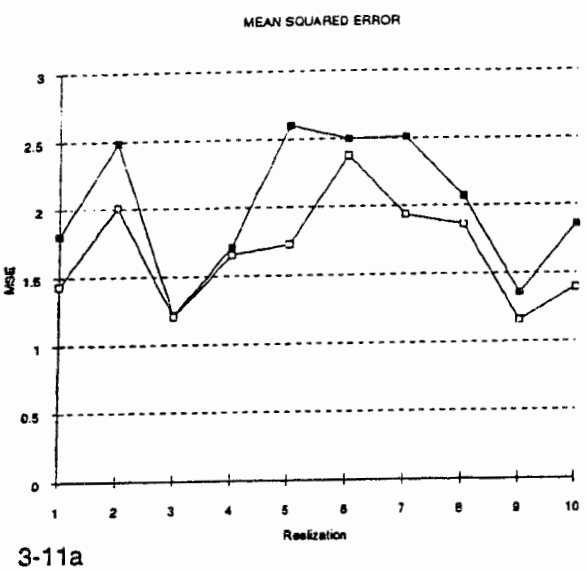


FIGURE **3-10**

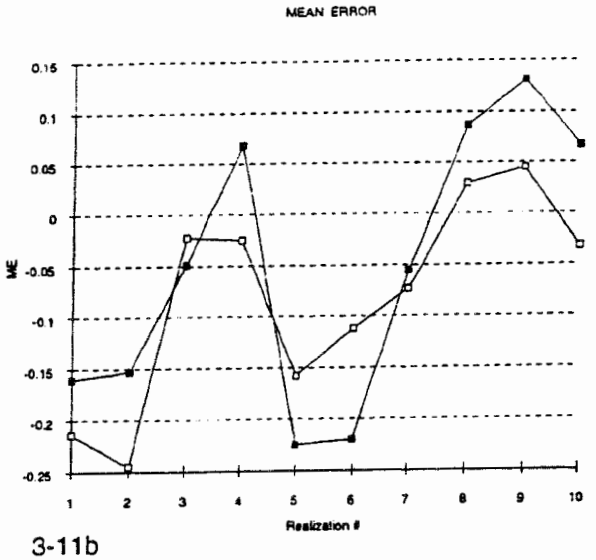
TEST STATISTICS FOR 10 REALIZATIONS OF THE "DELETE-50%" CROSS-VALIDATION OF THE NON-STATIONARY MODEL

SKB/GEOSTATISTICS/SWEDEN

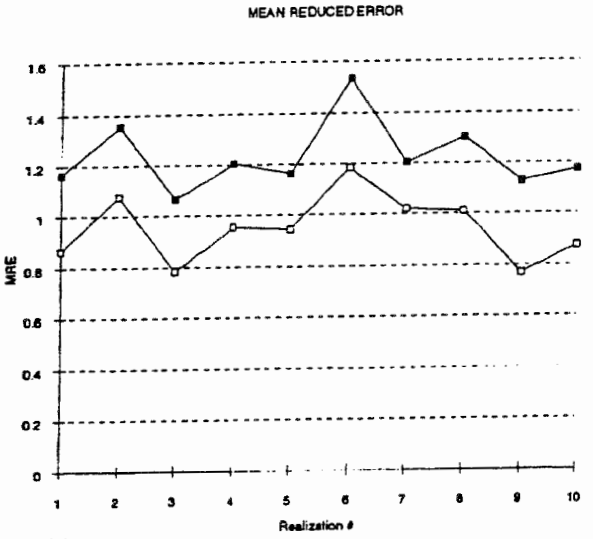
PROJECT NO. IC3 1129.000 DRAWING NO. 40885 DATE [illegible]



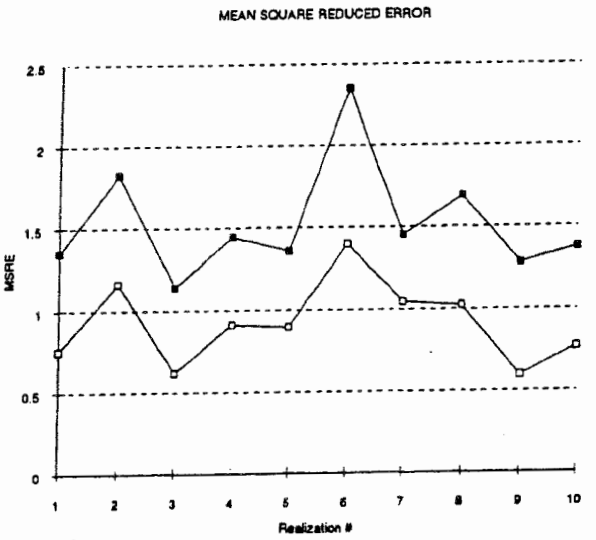
3-11a



3-11b



3-11a



3-11b

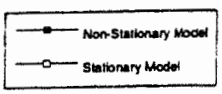
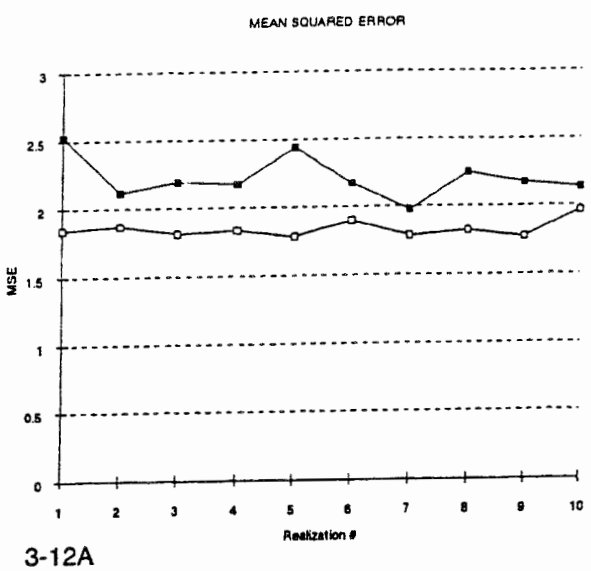


FIGURE **3-11**

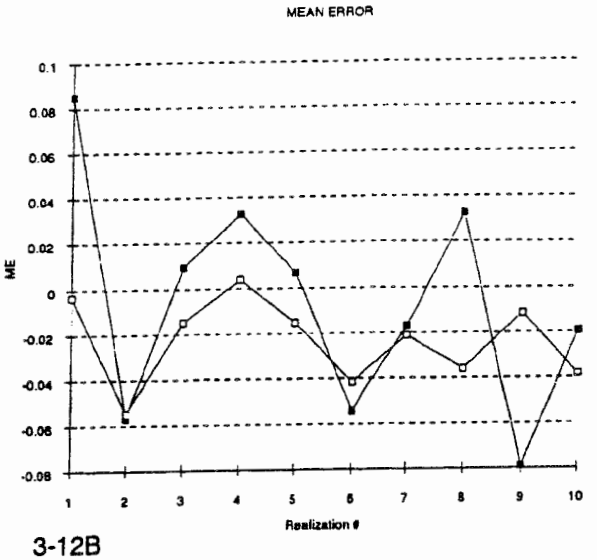
COMPARISON OF CROSS-VALIDATION STATISTICS FOR THE "DELETE-10%" CROSS-VALIDATION TESTS

SKB/GEOSTATISTICS/SWEDEN

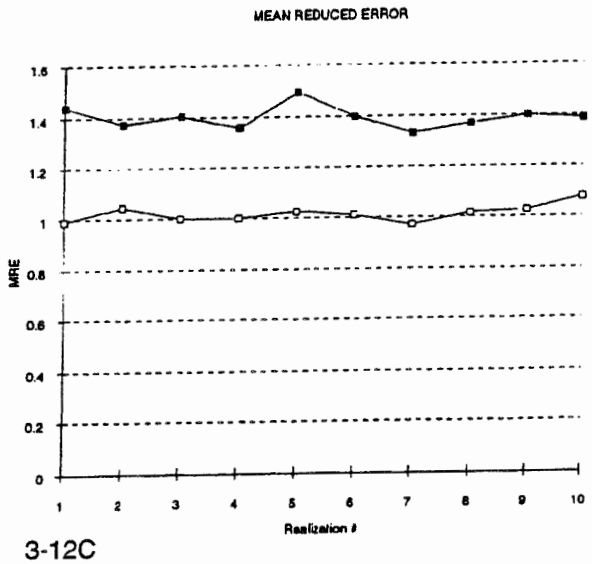
PROJECT NO. IP3 1129 DRAWING NO. 41069 DATE 11/1/94 DRAWN BY TB



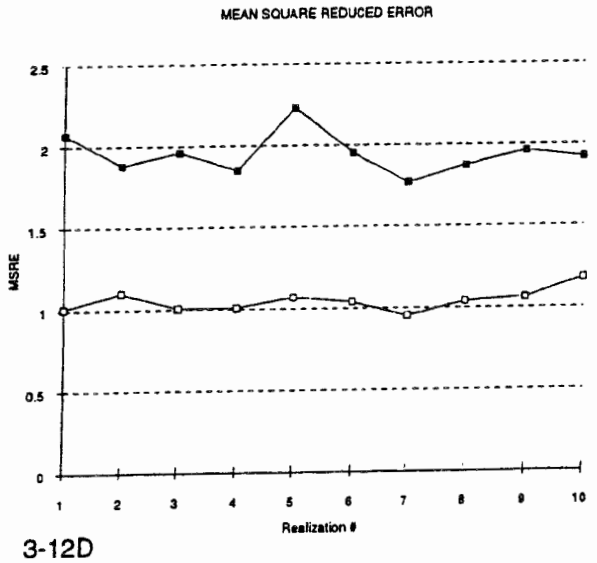
3-12A



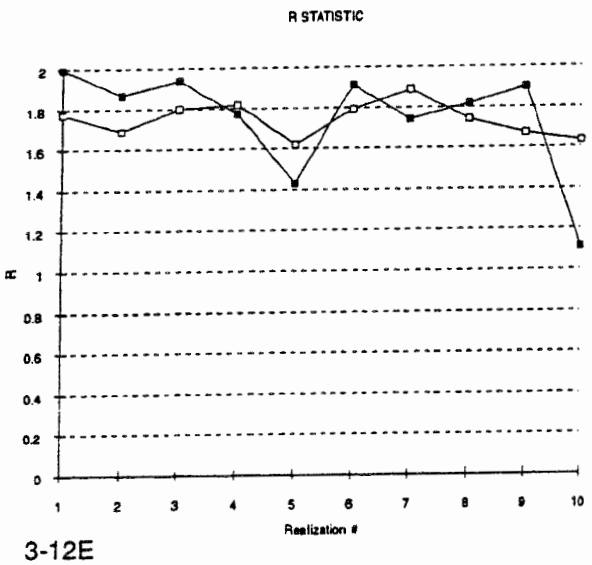
3-12B



3-12C



3-12D



3-12E

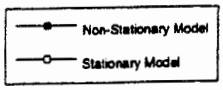


FIGURE **3-12**

COMPARISON OF CROSS-VALIDATION STATISTICS FOR THE "DELETE-50%" CROSS-VALIDATION TESTS

SKB/GEOSTATISTICS/SWEDEN

PROJECT NO. IP3 1129 DRAWING NO. 41070 DATE 11/7/94 DRAWN BY TB

predictive errors with a greater variance than would be expected if the model and the data were consistent. A plot of the reduced squared error for both the stationary and non-stationary models shows that the variance for the non-stationary models is much greater. Many of the bigger errors are located at the end or beginning of a well. This is because well tests at ends or beginnings are estimated more through extrapolation rather than interpolation. It is not uncommon to see "edge effects" like these in kriged maps. A possible reason for the greater variance might be related to the previous observations in Section 2 that the packer test data at Äspö appears to be drawn from a reasonably spatially stable population. Hence there is no advantage for a non-stationary algorithm. In this situation, the elementary generalized covariance functions may not represent the local spatial correlation as well as a stationary model. As Figure 3-2g indicates, it is possible to obtain a very good match to the first 70 meters of raw variogram data with a stationary model.

Isatis 1.3 does not have a convenient method to sample three-dimensional simulated data sets in the manner of Inferens 1.1. As Norman (1992a) points out, the assumption that the cross-validation Kriging error is normally distributed and uncorrelated is probably wrong, so that the calculation of confidence intervals for MRE, MSRE and MSE must rely upon simulations. For this reason, it was not possible to calculate confidence limits for the alternative models. Even without the rigorous calculation of confidence limits, it is clear that the non-stationary models produce poorer results for the Äspö 3 m data than do the stationary models.

Cross-validation statistics that fall within acceptable confidence limits do not guarantee that a model will be useful or successful for its intended purpose. Cross-validation is a test of a model's self-consistency. That a model is consistent with a small set of data does not clarify whether the inherent model parameter uncertainty creates unacceptable estimation uncertainty in modeling results. In other words, does the model sufficiently constrain the uncertainty so that models using the data will come to the same conclusions for a high percentage of realizations? For the same model, would different model realizations change the way in which a site is evaluated

or influence construction design alternatives? The typical mean squared errors (Tables 3-11a, 3-12a) made by the stationary and non-stationary models for the 3 m Äspö data (in log-transformed values) are on the order of 1.5 to 2.5. Recall that the mean-squared error measures the raw mis-estimation of deleted data (Fig. 1-2). If the mean mis-estimation is 2.0 for the log-transformed interval test values, then the actual error is on the order of two orders of magnitude for the untransformed values. This suggests that estimation of conductivity values for HYDRASTAR or other model input away from well tests will be at least this much in error, *on average*. However, this may or may not be a problem. Is it sufficient to error by as much as two orders of magnitude in the conductivity assigned to a model grid cell away from the wells from the standpoint of site evaluation or construction planning? Answers to these questions are beyond the scope of cross-validation statistics, and must be answered through a combination of stochastic hydrologic modeling and its link to the decision making process appropriate for the Äspö site.

3.6 Regularization and Upscaling of Packer Tests

3.6.1 Regularization Strategy

Hydrological modeling requires that hydraulic conductivity values be specified for volumes of rock larger than the typical packer test interval. A strategy imbedded in Inferens and Hydrastar (Norman, 1992a,b) is to average a series of shorter interval packer tests contained within a longer interval according to an upscaling formula developed by Norman (1992a). Geostatistical analysis is then performed on the estimates for the larger intervals. The geostatistical models for these larger intervals are then used in a Turning Bands procedure to simulate data for the flow simulator Hydrastar.

Upscaling of measured hydraulic conductivity or permeability data is an area of intense, on-going research in reservoir simulation in the petroleum industry (e.g.

Wattenbarger and others, 1993), and it is not yet clear how best to approach this problem. Upscaling is a type of regularization. Regularization refers to a change in the statistical support of the data under analysis. The goal of hydrological modeling is to regularize 3 m packer test data to larger three-dimensional volumes of rock. Experience in reservoir simulation in fractured reservoirs indicates that the hydraulic conductivity calculated from a 50 m long well test has little to do with the effective block-scale hydraulic conductivity in a 50 m by 50 m by 50 m block. This suggests that scaling 3 m packer tests to equivalent 50 m tests, and then stochastically estimating the hydraulic conductivity of a 50 m test in an untested portion of the rock mass may not solve the problem, since the link between the 50 m test and the block-scale conductivity is unknown.

A more flexible approach to estimating the block scale conductivity for models is to make a series of estimates for each block at the scale of the measured data. In terms of the Äspö data, this would mean estimating a series of 3 m packer test for each block. Various types of upscaling could then be tried and tested without having to re-do the geostatistical interpolation for each new upscaling method. It is unlikely that any one upscaling method will work best for all sites and conditions, so there is efficiency in separating the geostatistical component from the upscaling component and performing the upscaling after the interpolation.

Previous geostatistical studies at Finnsjön (Norman, 1992a; Geier, 1993a) regularized the data to larger scales, usually 36 m, according to a moving average process, and then performed the geostatistical analyses. The current study did not adopt this method for analysis of the Äspö data, for several reasons.

1) First, the wells at Äspö are deviated, and the deviation directions are different for different wells. As described in La Pointe (1994), such a moving average process leads to a variogram with a high degree of correlation for lag spacings less than the regularization length. As Norman (1994) correctly points out, this is due to the overlapping support of the moving average windows. There is nothing

mathematically incorrect about this method if care is taken in the Kriging algorithms to implement the formulation for non-disjoint support (Journal and Huijbregts, 1978, pg. 311). However, it will tend to obscure anisotropy and condition the fitted variogram to a portion of the raw variogram that is not as important for hydrological modeling.

To understand how this moving average regularization obscures anisotropy, consider that data was derived from a series of horizontal and vertical wells in a data set composed of nothing but Gaussian noise. The horizontal variogram calculated from the horizontal wells will shows a pronounced correlation for distances less than the regularization scale. However, if the same horizontal variogram were calculated from the vertical wells, there would be no correlation. If these two sets of data were then treated as a single set, some composite variogram would emerge which could be fit with great accuracy, but would not properly represent either the correlation structure of overlapping tests or the randomness of disjoint tests. In actual application, this problem may not be too severe, since well spacing is usually not at the scale of the regularization length.

2) A more practical problem is that the goal of the geostatistical analysis is to find a model which accurately predicts hydraulic conductivity at a distance from the wells. This distance is much greater than the regularization length. The moving average regularization creates a stronger degree of correlation over the lag pairs less than the regularization scale. The degree of correlation drives the automatic variogram fitting process, so that the result is that the variogram automatically fitted conforms to the short lag pairs. It may be a very poor fit to the data at scales greater than the regularization length. However, it is these larger scales that are most important to the actual interpolation accuracy.

3) A final problem with averaging the data prior to geostatistical analysis is that it introduces unwanted noise into the data. Figure 3-13 shows the results of applying the regularization formula described in Norman (1992a, pg 12) to regularize 3 m

packer test data to 30 m lengths. The hole radius is assumed to be 28 mm. This leads to a correction factor of approximately 1.462. In order to check on the accuracy of this formula for Äspö, it is possible to compare the regularized 30 m intervals with identical interval in wells KAS02, KAS03 and KLX01 over which actual 30 m packer tests were conducted. The regularization tends to overestimate the 30 m tests at Äspö introducing a bias. Bias in this instance is manifest in Fig. 3-13 in that many more points lie above the line than below it, and that the points above the line are often at a greater distance than those below. An unbiased regularization would over-predict and underpredict with equal probability. The same number of points would lie above and below, and their cumulative distance above and below would be equal. Fig. 3-14 is better than Fig. 3-13 in these respects. However, the bias would have no effect on the variogram analysis since multiplication of the data by a constant does not alter the variogram.

More important is the scatter about the line on the figure which represents a perfect match between the actual and the regularized prediction. In order to make the scatter more apparent, Figure 3-14 shows the data shifted down by computing the arithmetic average of the 3 m tests to better center it about the line. This does not change the variance from the previous plot (Fig. 3-13), since the only difference is in the scalar constant by which the arithmetic average is multiplied. Errors in predicting the 30 m tests are around one to two orders of magnitude, which is similar in magnitude to the mean squared error reported for the best models in the previous section. It is not clear from the figures whether the error is spatially correlated, but it would certainly introduce substantial non-spatially correlated noise into the data sets, which would obscure the true spatial correlation structure.

For all of these reasons, the 3 m data was not regularized and then analyzed. Rather, several calculations were performed on the 30 m packer test data for wells KAS02, KAS03 and KLX01. The results are described in the next section.

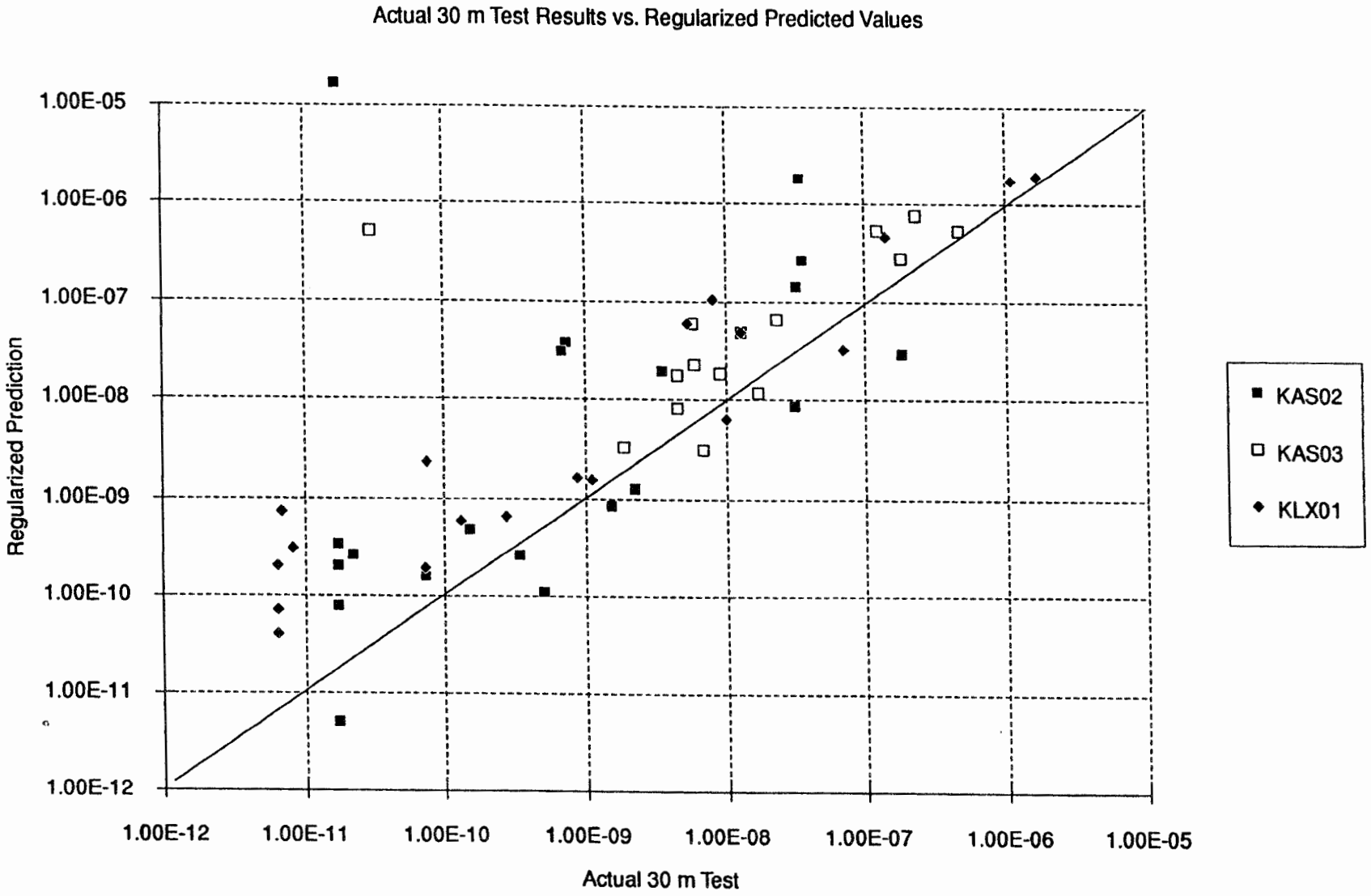


FIGURE **3-13**

3 M DATA REGULARIZED TO 30 M SCALE USING NORMAN'S (1992a) FORMULA COMPARED TO MEASURED 30 M RESULTS

SKB/GEOSTATISTICS/SWEDEN

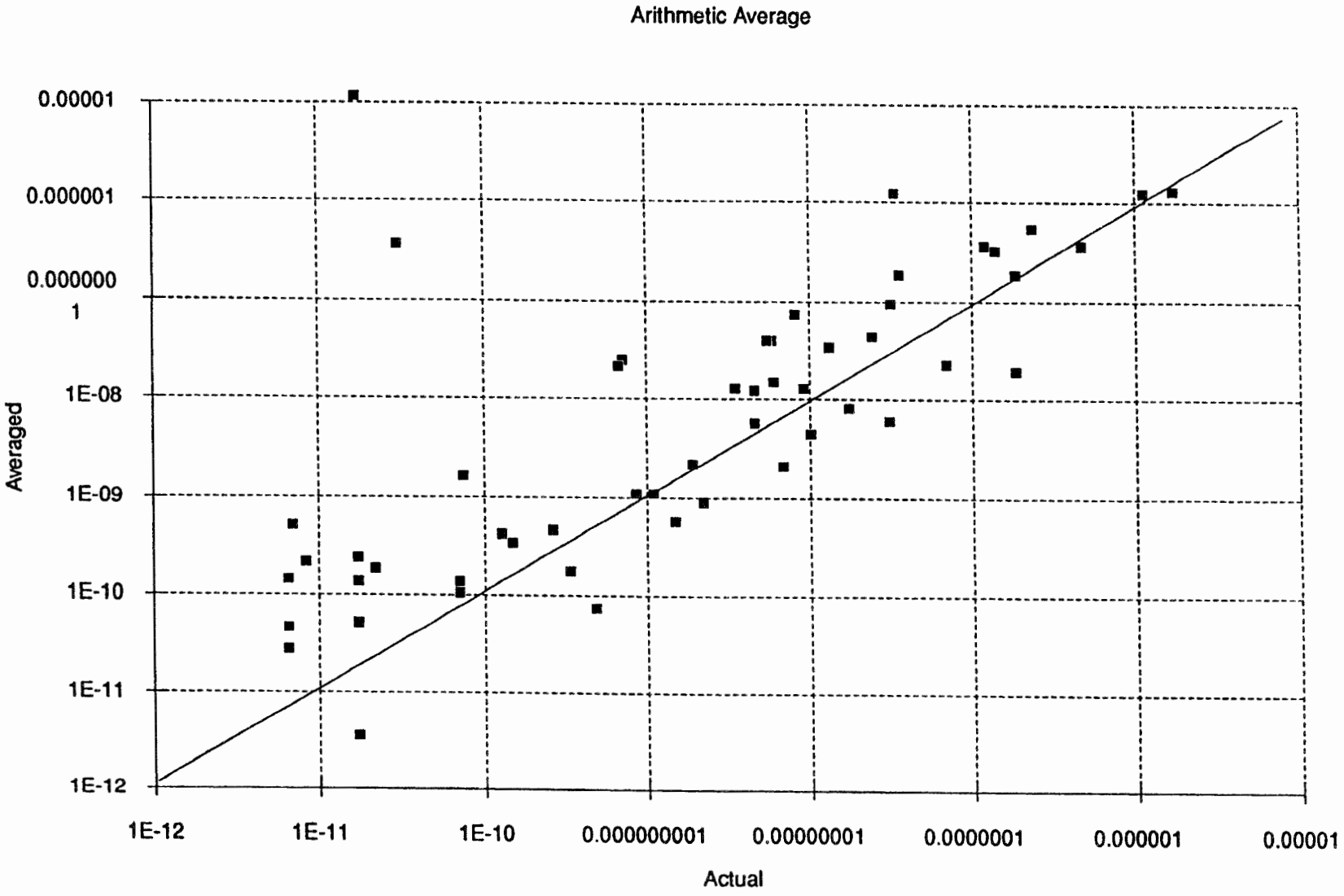


FIGURE **3-14**

3 M DATA REGULARIZED BY ARITHMETIC AVERAGING COMPARED TO MEASURED 30 M RESULTS

SKB/GEOSTATISTICS/SWEDEN

PROJECT NO. IC3 1129.000 DRAWING NO. 49897 DATE 6/29/94 DRAWN BY CB

Golder Associates

3.6.2 Results of Geostatistical Analysis of 30 m Test Data

Figure 3-15 shows the raw variogram and the variogram cloud for the 30 m data from the three wells. There does not appear to be any systematic change in correlation with distance. This may be more a function of the small amount of data used to construct the variogram rather than a reflection of the lack of spatial correlation.

Automatic structural recognition fitted a pure Nugget Effect variogram to the data with the cross validation results shown in Table 3.12.

Table 3.12

30 m Interval Cross-Validation Statistics	
MSE	2.482669
ME	0.00928
MRE	1.514108
MSRE	2.292522

This table indicates that a model assuming no spatial correlation is unbiased but is less variable than the actual data, which is evident in the dispersion of the variogram cloud.

The mean error is near to 0.0, which implies that the model makes as many overestimates as underestimates, in other words, is unbiased. However, the other error measures that should theoretically equal 1.0 are much greater than 1.0, which indicates that the model makes bigger errors than would be expected were the model a perfect representation of the 30 m data.

The 30 m results are not entirely consistent with the 3 m results. The range of the 3 m variogram was on the order of 70 m. This distance corresponds to the second lag class of the 30 m variogram, which has already reached the sill (Fig. 3-15). Perhaps the lower semivariance value for the first lag pair in the 30 m variogram reflects the existence of some spatial correlation for the shortest distance ranges, but in any case, the degree of spatial correlation for short lags pairs is a very small component in

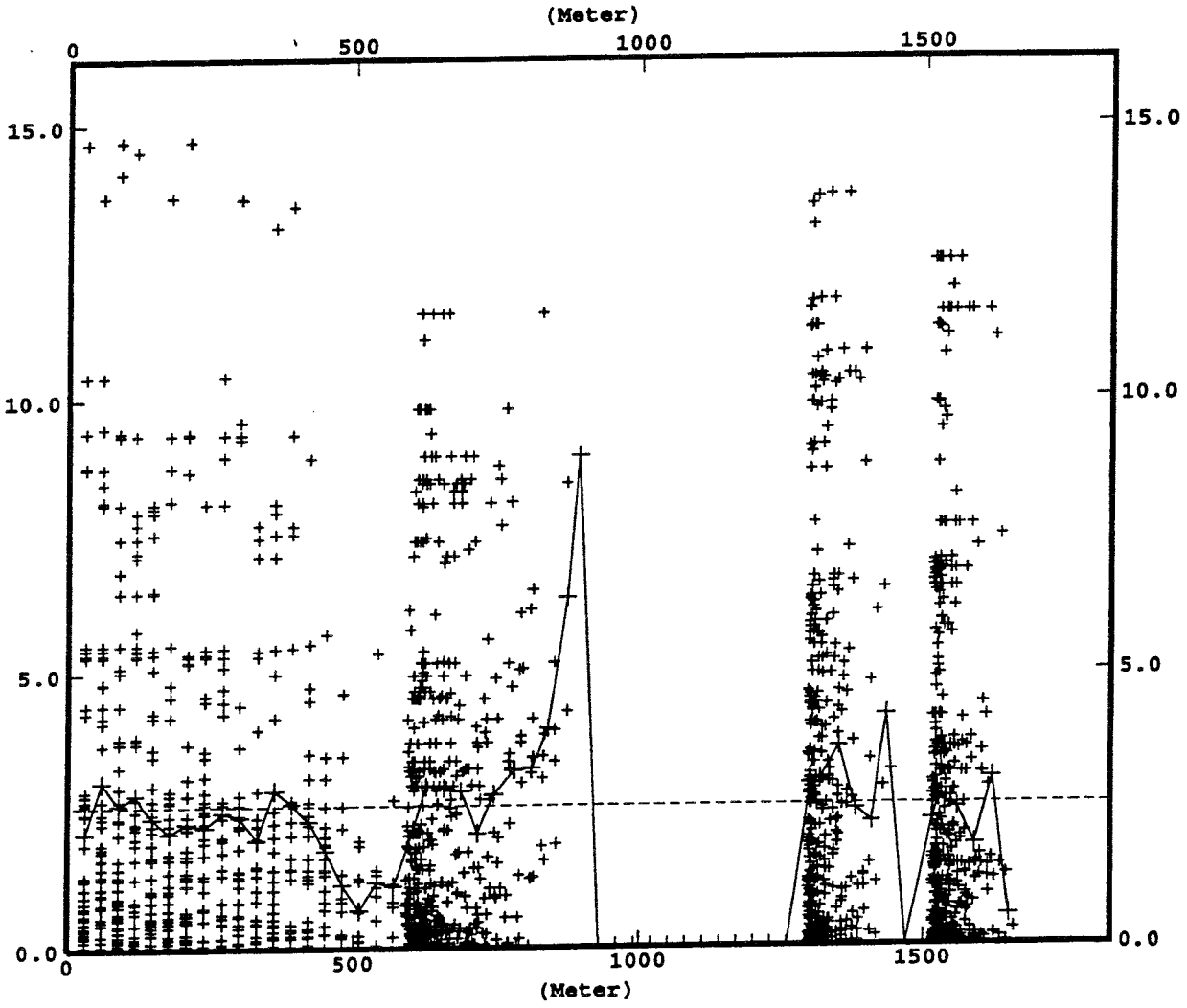


FIGURE **3-15**

RAW VARIOGRAM AND VARIOGRAM CLOUD FOR 30 M TEST DATA

SKB/GEOSTATISTICS/SWEDEN

proportion to the non-spatially correlated variability. The fact that both the 3 m and the 30 m variograms lose most of their spatial correlation by 70 meters, and that both have a high proportion of non-spatially correlated variability even a short lag pairs, is internally consistent.

There are some inconsistencies, however. It is interesting to compare the 3 m and 30 m variograms to further investigate how to regularize 3 m data. It is possible to calculate the range and variance of 3 m data regularized to 30 m, and then to compare it to the actual 30 m variogram. Formulas for regularizing data are given in Journal and Huijbregts (1978) for a variety of different variogram models. As shown in Sec. 3.3.1, a stationary model consisting of a Nugget of 1.4453 and a spherical variogram with sill equal to 0.77 is a good model for the 3m data. Journal and Huijbregts (1978) show that, for distances h $\geq$ L, where L is the length of the sample, the regularized variogram is:

$$\gamma_L(h)=\gamma(h)-\bar{\gamma}(L;L) \tag{3.2}$$

The term $\bar{\gamma}(L;L)$ for L $\leq$ a, where a is the range, is given by the expression:

$$\bar{\gamma}(L;L)=\frac{L}{2a}-\frac{L^3}{20a^3} \tag{3.3}$$

The reduction in variance due to regularization can be calculated from this formula. For the short range spherical variogram, it becomes:

$$\gamma_{reduced}=\gamma_3(h)-\gamma_{30}(h)=\bar{\gamma}(30;30)-\bar{\gamma}(3;3) \tag{3.4}$$

or approximately 0.19, reducing the sill for the regularized 30 m variogram from 0.77 to 0.58. The Nugget Effect, whether due to sampling error or microstructure, should decrease by the ratio 3/30 = 0.1 (Rendu, 1981). This means that the nugget effect should be 0.145. The range for the spherical variogram should increase approximately by 30 m, to about 100 m. This is not what the variogram presented in Figure 3-15 shows.

The 30 m variogram based upon 30 m well tests has a population variance virtually identical to the 3 m variogram - 2.61 for the 30 m and 2.59 for the 3 m (Table 3.13). While the variogram shown in Figure 3-15 might possibly show a range of 60 m or so, it is clear that the sill over the first 100 m is very much smaller than would have been predicted from the 3 m variogram, and the 30 m nugget effect is much larger.

Table 3.13 also shows that most of the first and second-order population statistics do not differ much between the 3 m and 30 m tests. In other words, the interval permeability tests do not scale between 3 m and 30 m. If the 3 m and 30 m well tests essentially tested similar volumes of rock, then this might explain why the variance of the tests does not appear to scale. Another possibility is differences in test conditions between the 3 m and 30 m tests, such that additional measurement error entered into the 30 m test interpretations.

Table 3.13

Parameter	3 m Tests (no KAS02)	30 m Tests
Mean	-8.05	-8.75
Standard Deviation	1.61	1.62
Variance	2.59	2.61
Minimum Value	-11.56	-11.19
Maximum Value	-4.01	-5.77
Number of Tests	553	63

4 CONCLUSIONS AND RECOMMENDATIONS

4.1 Conclusions Concerning the Spatial Statistics of the Packer Test Data

The following conclusions emerge from this study:

1) The packer test data, with the possible exception of measurements made in KAS02, can be treated as a single population from the standpoint of geostatistical analyses. At this time, there are no obvious correlation with mappable geological units or systematic trends with depth or areal location. Most of the variability of the packer data consists of three components: drift, the spatially correlated component and the non-spatially correlated stochastic component. The non-spatially correlated component constitutes a greater portion of the variability than does the correlated component.

2) There was no evidence found from analysis of directional variograms that the correlation structure was anisotropic on a regional scale.

3) Cross-validation statistics stationary models composed of a Nugget Effect plus either a spherical, exponential or spherical plus exponential models yield approximately the same results. This may be due to the fact that all of these models fit the shorter lag classes of the raw semivariogram equally well, and only differ at longer ranges which do not play a significant role in the cross-validation process.

4) Optimal Kriging neighborhoods are on the order of 200 m radius and partition the neighborhood into sectors.

5) For the non-stationary models using a neighborhood of 200 m radius and 8 sectors, the best order of drift was found to be zeroth or first order. The order of the intrinsic random functions was also of similar low order.

6) Cross-validation results for non-stationary models produced unacceptable MSRE and MRE and to a lesser extent, MSE statistics, while the ME statistics were the same or insignificantly better than for the stationary models. A possible explanation may be that the advantages of using non-stationary models may be lost at the Äspö site because the data appears to be very stationary. In such situations, stationary models may more accurately model the correlation structure.

7) Geostatistical models for the Äspö site should consist of stationary models. The raw 3 m packer tests are well-described by a nested model consisting of a Nugget Effect of 1.25 , a mid-range spherical variogram with range on the order of 70 m and Sill equal to 0.68, and a longer range exponential variogram with Sill equal to 0.30 and range equal to 200 m.

8) Regularized 3 m packer test data over-predicts 30 m tests conducted over the same intervals. A straight arithmetic average of the intervals produces less biased results, but both averaging methods introduce additional randomness into the data.

9) Variograms computed from 30 m tests show very little spatial correlation, which may in part be a function of the small amount of data. The amount of spatial and non-spatial variance expected by regularizing 3 m tests using standard geostatistical methods is not consistent with the statistics of the 30 m tests. The reason for this is not known, but suggests that standard geostatistical regulation methods may not work at Äspö.

4.2 Recommendations for Stochastic Hydrological Modeling at Äspö

1) Stationary models make errors of more than two orders of magnitude in the hydraulic conductivity of packer tests, even when the tests are adjacent to measured data. Whether this amount of error is tolerable requires the investigation of this variability of hydrological modeling predictions for a particular set of purposes. At

some stage, it would be important to carry out a series of Hydrastar simulations to evaluate whether the uncertainty would lead to design or evaluation changes.

2) If the errors discussed above are significant, it is unlikely that any stochastic inference method interpolating conductivity data from wells alone will much improve the situation. It will become essential to include other data that is more extensively distributed. For example, some component of geophysical data might relate to the hydraulic conductivity of the Äspö rock mass. If this type of additional "soft" data were available, then it would be possible to use it as a co-estimator in the form of co-Kriging (Journal and Huijbregts, 1979), as an external drift function (Geovariances, 1994) or as a trend surface. Non-geostatistical methods could also be employed, among them POCS (Projections on Convex Sets; Menke, 1991).

Co-Kriging is a method used for using more abundant soft data to co-estimate less abundant hard data when the spatial correlation for both variables are represented by stationary models. External drift plays a similar role in the non-stationary case, where the "shape" of the soft data replaces the standard drift functions in the automatic fitting of the intrinsic random functions. Neither method requires that the "hard" and "soft" data have the same spatial structure (Journal and Huijbregts 1978).

Co-Kriging and to a lesser extent, external drift, are well known in the literature. POCS is a very new algorithm with only a few publications describing it, but with much potential for stochastic modeling of flow properties (Malinverno, 1993). Because it may be useful for Äspö, it is worth some additional explanation. POCS efficiently conditions stochastic simulations to a variety of constraints in a stationary region. Simulations generated through POCS can have self-affine fractal and many types of geostastical constraints. The correlation constraints are imposed by adjusting the amplitude spectrum of the data. Other constraints are imposed by means of mathematical projections. Among the types of constraints that can be imposed in addition to spatial correlation are:

- The simulation can take on the same values as measured data in wells or elsewhere
- The simulated data can have the same mean, energy content and variance as the measured data

- The values can be bounded. For example, it is possible to constrain all of the hydraulic conductivities to being positive or to be greater than some threshold value or less than some maximum credible value. This constraint could also be used to incorporate as a constraint nearly half the Äspö well data below threshold.

- Faults or other discontinuities can be honored

- The data can be forced to monotonically increase or decrease along a curve **X(s)**. This feature makes it possible to replicate many types of drift.

- Spatial correlation properties may be orthotopic

POCS assumes that the data possesses second order stationarity throughout the simulation region, which appears to be clearly the case at Äspö. The advantage of a POCS approach is that it can enforce data constraints other than the spatial correlation, which the present study strongly suggests is insufficient in itself to make accurate estimation at Äspö. Another useful aspect of POCS is that it is an iterative procedure, and the order in which the constraints are imposed during each iteration effectively makes it possible to emphasize one constraint over another should there be an inconsistency or conflict between the model and the data. For example, imposition of the spatial correlation constraint first means that any later imposed constraint that conflicts with the spatial correlation model will be the one enforced.

However more appropriate the capabilities of POCS might be for Äspö, it too will be limited by the very sparse well data. Regardless of the stochastic inference algorithm employed, the well data must be supplemented with some more extensive "soft" data for estimation accuracy to improve.

3) The issue of regularizing the packer test data prior to geostatistical analysis should be evaluated and tested. It would be worthwhile to create several simplified discrete fracture models, simulate packer tests in them, and then evaluate various strategies for predicting block scale fluid flow properties. The mathematical types that could be tested include models with no spatial correlation, those with a self-affine fractal structure, and those whose with a stationary spherical or exponential variogram. A series of wells based upon the well geometry of the packer tests at Äspö could be inserted into the model, and packer tests simulated. Block scale hydraulic conductivity could be calculated by simulating flow under a constant gradient using a code like MAFIC. The goal would then be to predict the block scale conductivity from the packer tests through different types of regularization both before and after the geostatistical analysis.

These models could also help in evaluating different types of stochastic interpolation methods. It would be possible to compare performance of algorithms like POCS, or to evaluate the impact of including soft data through External Drift or Co-Kriging.

5 REFERENCES CITED

Feder, J. (1988). Fractals. Plenum Press, New York, 283p.

Geier, J. (1993a). Verification of the geostatistical inference code INFERENS, Version 1.1, and demonstration using data from Finnsjön. SKB Technical Report TR 93-09.

Geier, J. (1993b). Version 1.0 User's Guide to Inferens 1.1. SKB Arbetsrapport AR 93-24, 73p.

Geier, J. (1994). Geologic constraints for fitting geostatistical models using Inferens 1.1. Project memorandum, May 25, 1994.

Geovariances (1994). Isatis Version 1.3 Users Documentation. Geovariances, Avon, France.

Journal, A. G. and Ch. J. Huijbregts (1978). Mining Geostatistics. Academic Press, London, 600p.

LaPointe, P.R. (1980). Analysis of the spatial variation in rock mass properties through geostatistics, in Proc. of the 21st U.S. Symposium on Rock Mechanics, p. 570-580.

La Pointe, P. R. (1994). Suspected serious Inferens 1.1 problem. Memo to Anders Ström, 16 June, 1994.

Liedholm, M. (1991). Conceptual Modeling of Äspö. Technical Notes 1 - 17. SKB HRL Progress Report 25-90-16a.

Malinverno, A. and Rossi (1993) Applications of projections onto convex sets to stochastic inversion. Paper SPE 25659, *in* Proc. Society Petroleum Engineers Middle East Oil Technical Conf., Bahrain, 3-6 April 1993, pp. 515-523.

Menke, W. (1991). Applications of the POCS inversion method to interpolating topography and other heophysical fields. Geophys. Res. Letters, Vol. 18, pp. 435-438.

Nilsson, L. (1989). Hydraulic tests at Äspö and Laxemar. SKB HRL Progress Report 25-88-14, 86p.

Nilsson, L. (1990). Hydraulic tests at Äspö, KAS05-KAS08, HAS13-HAS17, Evaluation. SKB HRL Progress Report 25-89-20, 85p.

Norman, S. (1992a). Statistical inference and comparison of stochastic models for the hydraulic conductivity at the Finnsjön-site. SKB Technical Report TR 92-08. 113 p.

Norman, S. (1992b). HYDRASTAR - a code for stochastic simulation of groundwater flow. SKB Technical Report TR 92-12, 107p.

Norman, S. (1994) Comments to Memorandum "Suspected Serious Inferens 1.1 Problem". Memo to Anders Strom, 20 June, 1994.

Renard, D. (1989). Automatic structure recognition. *in* GEOSTATISTICS, Proc. of the 3rd International Geostatistics Congress, Sept. 5-9, 1988, Avignon, France. Kluwer Academic Publishers, Dordrecht. pp. 579-590.

Russo, D. and W. A. Jury (1987). A theoretical study of the estimation of the correlation scale of spatially variable fields,1, stationary fields. Water Resources Research, Vol 23, No. 7 pp 1257-1268.

Stearns, D. W. and M. Friedman (1972). Reservoirs in fractured rock. American Assoc.Petrol. Geol. Memoir 16, Stratigraphic Oil and Gas Fields, pp. 82-106.

Wattenbarger, R. C. K. Aziz and F. M. Orr, Jr. (1993). Optimal scales for representing reservoir heterogeneity. *in* Proc. 3rd International Reservoir Characterization Technical Conf., Tulsa, OK, Nov. 3-5, 1991. PennWell Publishing Co., Tulsa. pp. 195-225.

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

1990
TR 90-46
SKB Annual Report 1990
Including Summaries of Technical Reports Issued during 1990
Stockholm, May 1991

1991
TR 91-64
SKB Annual Report 1991
Including Summaries of Technical Reports Issued during 1991
Stockholm, April 1992

1992
TR 92-46
SKB Annual Report 1992
Including Summaries of Technical Reports Issued during 1992
Stockholm, May 1993

1993
TR 93-34
SKB Annual Report 1993
Including Summaries of Technical Reports Issued during 1993
Stockholm, May 1994

Technical Reports
List of SKB Technical Reports 1994

TR 94-01
Anaerobic oxidation of carbon steel in granitic groundwaters: A review of the relevant literature
N Platts, D J Blackwood, C C Naish
AEA Technology, UK
February 1994

TR 94-02
Time evolution of dissolved oxygen and redox conditions in a HLW repository
Paul Wersin, Kastriot Spahiu, Jordi Bruno
MBT Technología Ambiental, Cerdanyola, Spain
February 1994

TR 94-03
Reassessment of seismic reflection data from the Finnsjön study site and prospectives for future surveys
Calin Cosma[1], Christopher Juhlin[2], Olle Olsson[3]
[1] Vibrometric Oy, Helsinki, Finland
[2] Section for Solid Earth Physics, Department of Geophysics, Uppsala University, Sweden
[3] Conterra AB, Uppsala, Sweden
February 1994

TR 94-04
Final report of the AECL/SKB Cigar Lake Analog Study
Jan Cramer (ed.)[1], John Smellie (ed.)[2]
[1] AECL, Canada
[2] Conterra AB, Uppsala, Sweden
May 1994

TR 94-05
Tectonic regimes in the Baltic Shield during the last 1200 Ma - A review
Sven Åke Larsson[1,2], Eva-Lena Tullborg[2]
[1] Department of Geology, Chalmers University of Technology/Göteborg University
[2] Terralogica AB
November 1993

TR 94-06
First workshop on design and construction of deep repositories - Theme: Excavation through water-conducting major fracture zones Såstaholm Sweden, March 30-31 1993
Göran Bäckblom (ed.), Christer Svemar (ed.)
Swedish Nuclear Fuel & Waste Management Co, SKB
January 1994

TR 94-07
INTRAVAL Working Group 2 summary report on Phase 2 analysis of the Finnsjön test case
Peter Andersson (ed.)[1], Anders Winberg (ed.)[2]
[1] GEOSIGMA, Uppsala, Sweden
[2] Conterra, Göteborg, Sweden
January 1994

TR 94-08
The structure of conceptual models with application to the Äspö HRL Project
Olle Olsson[1], Göran Bäckblom[2], Gunnar Gustafson[3], Ingvar Rhén[4], Roy Stanfors[5], Peter Wikberg[2]
1 Conterra AB
2 SKB
3 CTH
4 VBB/VIAK
5 RS Consulting
May 1994

TR 94-09
Tectonic framework of the Hanö Bay area, southern Baltic Sea
Kjell O Wannäs, Tom Flodén
Institutionen för geologi och geokemi, Stockholms universitet
June 1994

TR 94-10
Project Caesium–An ion exchange model for the prediction of distribution coefficients of caesium in bentonite
Hans Wanner[1], Yngve Albinsson[2], Erich Wieland[1]
[1] MBT Umwelttechnik AG, Zürich, Switzerland
[2] Chalmers University of Technology, Gothenburg, Sweden
June 1994

TR 94-11
Äspö Hard Rock Laboratory Annual Report 1993
SKB
June 1994

TR 94-12
Research on corrosion aspects of the Advanced Cold Process Canister
D J Blackwood, A R Hoch, C C Naish, A Rance, S M Sharland
AEA Technology, Harwell Laboratory, UK
January 1994

TR 94-13
Assessment study of the stresses induced by corrosion in the Advanced Cold Process Canister
A R Hoch, S M Sharland
Chemical Studies Department, Radwaste Disposal Division, AEA Decommissioning and Radwaste, Harwell Laboratory, UK
October 1993

TR 94-14
Performance of the SKB Copper/Steel Canister
Hans Widén[1], Patrik Sellin[2]
[1] Kemakta Konsult AB, Stockholm, Sweden
[2] Svensk Kärnbränslehantering AB, Stockholm, Sweden
September 1994

TR 94-15
Modelling of nitric acid production in the Advanced Cold Process Canister due to irradiation of moist air
J Henshaw
AEA Technology, Decommissioning & Waste Management/Reactor Services, Harwell, UK
January 1994

TR 94-16
Kinetic and thermodynamic studies of uranium minerals. Assessment of the long-term evolution of spent nuclear fuel
Ignasi Casas[1], Jordi Bruno[1], Esther Cera[1], Robert J Finch[2], Rodney C Ewing[2]
[1] MBT Tecnología Ambiental, Cerdanyola, Spain
[2] Department of Earth and Planetary Sciences, University of New Mexico, Albuquerque, NM, USA
October 1994

TR 94-17
Summary report of the experiences from TVO's site investigations
Antti Öhberg[1], Pauli Saksa[2], Henry Ahokas[2], Paula Ruotsalainen[2], Margit Snellman[3]
[1] Saanio & Riekkola Consulting Engineers, Helsinki, Finland
[2] Fintact Ky, Helsinki, Finland
[3] Imatran Voima Oy, Helsinki, Finland
May 1994

TR 94-18
AECL strategy for surface-based investigations of potential disposal sites and the development of a geosphere model for a site
S H Whitaker, A Brown, C C Davison, M Gascoyne, G S Lodha, D R Stevenson, G A Thorne, D Tomsons
AECL Research, Whiteshell Laboratories, Pinawa, Manitoba, Canada
May 1994

TR 94-19
Deep drilling KLX 02. Drilling and documentation of a 1700 m deep borehole at Laxemar, Sweden
O Andersson
VBB VIAK AB, Malmö
August 1994

TR 94-20
Technology and costs for decommissioning the Swedish nuclear power plants
Swedish Nuclear Fuel and Waste Management Co, Stockholm, Sweden
June 1994

TR 94-21
Verification of HYDRASTAR: Analysis of hydraulic conductivity fields and dispersion
S T Morris, K A Cliffe
AEA Technology, Harwell, UK
October 1994