Final Research Report

(Technical)

"Research Initiation: Optimal Sampling of Stochastic Processes"

National Science Foundation
Grant No. ECE-8503897

(Georgia Tech Project No. E-20-635)

Principal Investigator:
Shahrokh Rouhani, Ph.D.
School of Civil Engineering
Georgia Institute of Technology
Atlanta, GA 30332

NSF Program Official:
William A. Anderson, Ph.D.
National Science Foundation
Washington, DC 20550

Period Covered: 6/1/85 - 10/30/87

December, 1987

ABSTRACT

Universal kriging and variance reduction analysis are geostatistical procedures for mapping and sampling of spatial random variables. In this project these methods are expanded to the time-space domain in order to be applicable to spatiotemporal random variables. A large number of hydrological processes can be viewed as such variables. In the first phase of the project different versions of kriging are developed. They include non-negative universal kriging, universal time kriging and universal space-time kriging. The presented case studies are mapping of groundwater hydraulic variables in southwestern Georgia, estimation of drought lead time as a management tool for reservoir operations, and mapping of transient piezometric surface in southern Georgia. In the second phase, variance reduction analysis is expanded along universal kriging. It is used for the design of a groundwater quality monitoring network in the shallow aquifer of Dougherty Plain in southwestern Georgia, as well as, for the optimal sampling of transient piezometric heads in southern Georgia. In the final phase of the project, the resilience of variance reduction analysis is studied. It appears that variance reductoin analysis has an unstable parameter space, but a resilient action space. The above developments provide a reliable geostatistical sampling scheme for spatiotemporal random variables.

ACKNOWLEDGEMENTS

The PI would like to Thank all his scientific collaborators, especially, Prof. Myron Fiering, Mr. Tim Hall, and Mr. Ken Cargile for their help and cooperation. I am also indebted to Prof. Ram Arora of Georgia State University for his comments. Finally, I should thank the staff at the Centre de Geostatistique, Ecole des Mines de Paris, Fontainebleau, France for their support in the preparation of this report.

TABLE OF CONTENTS

List of Figures

List of Tables

Table No. Page

1. INTRODUCTION

A large number of natural and physical phenomena in hydrology can be viewed as stochastic processes. Variables such as, transmissivity, storativity, and piezometric heads are of this type. The sampled values of these variables usually exhibit complex behaviors, which at the first glance, appear to be totally random. However, points taken at neighboring locations in space and time reveal a degree of stochastic structure. Such structures can be represented by a variety of statistical models. Geostatistical techniques provide us with tools to study different problems associated with such random variables, including their spatiotemporal mapping and sampling.

As shown by Rouhani (1985), sampling of spatiotemporal variables can be studied in the framework of geostatistical procedures. In particular, the PI proposes to utilize a generalized scheme for optimal sampling, known as variance reduction analysis, which is based on the universal kriging (Rouhani, 1983 and 1985).

Universal kriging is a generalized Gauss-Markov interpolation method for estimation of non-stationary random variables. This procedure provides linear estimates of the variable of interest, as well as, a measure for the accuracy of these estimated values. This measure is given in the form of an estimation variance. Many authors, such as Matalas (1968) and DeMarsily (1979) propose to add sampling points at sites with highest estimation variances in order to minimize the regional variance. This approach, however, ignores the impact of a new sampling point on the accuracy of its neighboring zones. To resolve this problem Rouhani (1985) derives a measure for the relative influence of an arbitrary added measurement point on the estimation accuracy at another

location. Areal expansion of this measure then yields a regional indicator for the information efficiency of any potential sampling point. Two optimality criteria are utilized for the ranking of potential sampling sites. The first one directly depends on the variance reduction values and measure the amount of accuracy or informaiton gains, while the second one is proportional to the expected economic loss reductions due to new measurements. These two ranking functions are utilized to determine the best sequence of new added points.

Prior to this project, the above algorithm was generally applied to cases where the variables of interest were assumed to be Gaussian, and only spatially distributed. Such an approach excluded the application of variance reduction analysis to the important classes of non-negative spatiotemporal variables. These variables constitute a major group of stochastic processes in hydrology and water resources.

In response to the above problems, the PI proposed a research plan based on the expansion and extention of variance reduction analysis in the time-space domain. In the process of this project, which was initiated on 6/1/85 and effectively lasted until 10/30/87, the following major tasks were performed. In the first phase of the project, universal kriging was expanded and modified to perform the interpolation of non-negative random variables, time series, and finally spatiotemporal variables. For application purposes we utilized geohydraulic data in southwestern Georgia, streamflow data in western Georgia, and piezometric data in southern Georgia. Results of this phase provided more realistic interpolations, as well as, maps of forecasted and hindcasted values.

In the second phase of this project, variance reduction

analysis was expanded, along with universal kriging, to yield a more general sampling procedures for hydrological variables. In the first application we devised a scheme for groundwater sampling in The Dougherty Plain in Georgia. In our second attempt we studied the groundwater quantity sampling using the space-time variance reduction analysis. In the final phase of this project we looked into the question of the resilience of our proposed scheme under hypothetical conditions. The detailed discussion of the above findings is the topic of the following sections.

2. EXPANSION AND MODIFICATION OF KRIGING

Geostatitical methods such as kriging have been extensively used in the estimation of different hydrological phenomena. For example, auhtors like Delhomme (1979), DeMarsily (1979), and Ahmed and Demarsily (1987) utilize different version of kriging for mapping and simulation of transmissivity fields. Dunlap and Spinazola (1981), Rouhani (1983 and 1986), and Aboufirassi and Marino (1983) also use kriging for the spatial study of groundwater piezometric surfaces.

Others like Fogg et al. (1979), Dagan (1979 and 1982), and Chirlin and Dagan (1980) utilize geostatitical methods for the solution of stochastic groundwater equations. Another group of researchers work with kriging in the context of inverse problems, including Neuman and Yakowitz (1979), Neuman (1980), Neuman et al. (1980), and Kitanidis and Vomvoris (1982). Furthermore, a number of authors, such as Delhomme (1977) and Chua and Bras (1980) applied kriging procedures to precipitation data.

In all the above works the hydrological variables of interest are studied only in the space domain. Even with rainfall data, rather than using spatiotemporal algorithms, the authors have chosen a variety of spatial approaches. This is usually done by focusing on temporally averaged quantities at each point, such as monthly, seasonal, or annual values. Thus, the temporal structure of the data is only implicitly considered.

In the present work, however, we expand kriging to the time-space domain, so we can study the phenomena of interest as spatiotemporal variables. For this purpose, we have utilized universal kriging as the basis of our work, as one

of the more advanced versions of kriging.

Universal kriging is a Gauss-Markovian interpolation method for non-stationary random variables. In punctual estimation, given the measured values of a random variable Z at the measurement points X_i, i=1,....,N, universal kriging provides the best unbiased estimate of Z at X_o (the arbitrary location of an unmeasured site). The estimate is given in the following linear form:

$$\hat{Z}(X_o) = \sum_{i=1}^{N} \lambda_{io} Z(X_i) \qquad (2.1)$$

where,

$\hat{Z}(X_o)$ = kriging estimate at X_o;
$Z(X_i)$ = measured value at X_i, i=1,....,N; and
λ_{io} = kriging weights for $Z(X_i)$ to estimate $Z(X_o)$.

The λ_{io} are defined by two criteria. (1) Unbiasedness: $E(\hat{Z}(X_o)-Z(X_o)) = 0$, where $Z(X_o)$ is the true value of the variable at X_o, which is unknown. (2) Minimum squares error, which requires $E(\hat{Z}(X_o)-Z(X_o))^2$ to be minimum. This variance is also defined as the estimation or kriging variance, $Var((\hat{Z}(X_o))$.

Universal kriging views the process Z(X) as a random variable with the following structure:

$$Z(X) = M(X) + R(X) \qquad (2.2)$$

where M(X) is a slowly varying deterministic function known as the drift which is equal to the expected value of Z at point X. R(X) is a Gaussian stationary random variable with zero expectation. It is also assumed that R(X) has a covariance function $K(X_i,X_j)$, or simply a (semi-)variogram $\gamma(X_i,X_j)$, which depend only the distance vector between X_i

and X_j.

Kriging basically considers Z as a realization of a random function. Thus, in order to be able to estimate its statistical characteristics, it further assumes that the k^{th} order increment of Z, namely R, is stationary and satisfies the ergodic hypothesis. Stationarity implies that the probability distribution function of R(X) does not vary with X. Assumption of ergodicity indicates that the variability of Z is same as in the ensemble of realizations. This assumption is almost impossible to check in practice. Therefore, as DeMarsily and Ahmed (1987) note, these rather theoretical hypotheses are just working hypotheses to enable us to develope a model. It is never claimed that the variable of interest is stationary or ergodic in nature. They are simply used as a set of tools for parameter estimation, and must be checked to avoid inconsistency with data.

For automatic estimation of the covariance function, Matheron (1973) proposes to study Z as a realization of an intrinsic random function that can be made stationary by an incrementing process. First, it is assumed that M(X) admits a local presentation in the form of a polynomial of order k. Then λ_{io} are defined in such a manner that the linear combination $\sum_i \lambda_{io} Z(X_i)$, for i=0,1,....,N, filters out the mean, defined by $M(X_o)$. This approach leaves out the important step of estimating the actual parameters of the drift function.

For the case of an intrinsic random function of order 0, 1, or 2 in two-dimensions with Cartesian coordinates (x_i, y_i) the above incrementing requirements can be written as:

$$
\begin{array}{lll}
k = 0 & \sum_{i=1}^{N} \lambda_{io} = 1 & \\
k = 1 & \sum_{i=1}^{N} \lambda_{io} x_i = x_o & \sum_{i=1}^{N} \lambda_{io} y_i = y_o \\
k = 2 & \sum_{i=1}^{N} \lambda_{io} x_i y_i = x_o y_o & \sum_{i=1}^{N} \lambda_{io} x_i^2 = x_o^2 \\
 & \sum_{i=1}^{N} \lambda_{io} y_i^2 = y_o^2 & (2.3)
\end{array}
$$

The above constraints constitute the unbiasedness criterion of the universal kriging. They are also referred to as universality conditions. The estimation variance is then defined as:

$$
\mathrm{Var}(\hat{Z}(X_o)) = E(Z(X_o) - \sum_{i=1}^{N} \lambda_{io} Z(X_i))^2 =
$$

$$
= \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_{io} \lambda_{jo} K(|X_i - X_j|) \qquad (2.4)
$$

For added efficiency Matheron (1973) also suggests a family of admissible polynomial covariance functions for the two- and three-dimensional cases, as shown in Table 2.1.

At this stage λ_{io} are estimated by minimizing the estimation variance (2.4), subject to the incrementing constraints (universality conditions) given by (2.3). Using Lagrange multipliers, μ_{po}, for each constraint, one obtains the following set of equations, known as the kriging system:

$$
\sum_{j=1}^{N} \lambda_{io} K(|X_i - X_j|) + \sum_{p=1}^{l(k)} \mu_{po} f_p(X_i) = K(|X_o - X_i|) \quad i=1,\ldots,N
$$

DRIFT	k	f_p in R^2	f_p in R^3	MODELS OF GC
CONSTANT	0	1	1	$K(h) = C\delta(h) + a_1 h$
LINEAR	1	1,x,y	1,x,y,z	$K(h) = C\delta(h) + a_1 h + a_3 h^3$
QUADRATIC	2	$1,x,y,xy,x^2,y^2$	$1,x,y,z,xy,xz,$ yz,x^2,y^2,z^2	$K(h) = C\delta(h) + a_1 h + a_3 h^3 + a_5 h^5$
CONSTRAINTS ON THE COEFFICIENTS		$a_1 \leq 0 \quad a_5 \leq 0$ in R^2: $a_3 \geq -\frac{10}{3}(a_1 a_5)^{1/2}$	in R^3: $a_3 \geq -(10\ a_1 a_5)^{1/2}$	

TABLE 2.1: SELECTED MODELS FOR GENERALIZED COVARIANCES

Source: Delfiner [1975]

Table 2.1 Selected Models for Polynomial Covariances.

$$\sum_{j=1}^{N} \lambda_{io} f_p(Xj) = f_p(X_o) \qquad p=1,\ldots,l(k) \qquad (2.5)$$

where $f_p(X)$ is the p^{th} monomial in the drift function at X, l is the number of such monomials in the drift function that depends on the order k of the polynomial drift.

It is clear that for kriging one needs to estimate the order of the polynomial drift, as well as, the parameters of the covariance function. This pre-kriging task is known as the structural analysis. For structural analysis we use the algorithm suggested by Delfiner (1975), and a series of computer programs developed by the PI on the basis of an earlier work by Kafritsas and Bras (1981). For a detailed study of these procedures readers are referred to Rouhani (1983).

As noted earlier, the above algrithms are basically applied to cases of Gaussian spatial random variables. It is our aim,. as described in the following sections to expand the application domain of universal kriging to non-negative spatiotemporal random variables. This goal is obtained in a step by step manner, first by obtaining a non-negative universal kriging, followed by universal time kriging, and eventually, by the space-time universal kriging.

2.1. Application of Universal Kriging to Non-Negative Random Variables

A significant number of variables in hydrology that are considered as random processes, cannot acquire negative values. This is due to either physical characteristics of the variable, or the way in which it is defined. For example, point rainfall by its nature cannot have a negative value, while the net rainfall can have both positive and negative values. Transmissivity is another example which by its definition is always positive. In contrast, the groundwater flow rate, depending on its direction can acquire both positive and negative values.

In order to deal with this problem, authors have suggested a number of alternative approaches. For instance, Szidarovsky *et al.* (1987) and Baafi *et al.* (1986) discuss procedures for forcing the kriging system to produce only positive λ_{io}. This is done by adding N non-negative constraints for each weight: $\lambda_{io} \geq 0$, for i=1,...,N, to the kriging system (2.5).

The above alternative creats two basic problems. The first one is of theoretical nature, and that is, in universal kriging, it is assumed that the variable of interest has a Gaussian distribution - capable of acquiring both positive and negative values. Thus, imposing such a non-negative constraints simply causes a contradiction with the basic underlying assumption of the kriging.

The second problem is of a practical nature, which is caused by the addition of non-negative constriants. The existance of such inequality constraints prohibits the use of simple lagrange multipliers optimization scheme. Indeed the above authors have suggested optimization procedures which

are significantly less efficient than the lagrange mulitipliers scheme. These reasons lead us to search for other alternative approaches.

While studying non-negative variables, the first observations indicate that the assumption of Gaussian distribution is simply not valid. So we should probably look for approaches that allow relaxation of this assumption. Disjunctive kriging (Journel and Huijbregts, 1978, Yates et al., 1986a and 1986b) offers an approach based on the approxmation of any non-Gaussian process by a truncated series of Hermite polynomials of normally distributed variables of the form:

$$Z(X) = \Phi(y(X)) = \sum_{j=0}^{J} C_j H_j(y(X)) \qquad (2.6)$$

where,

$H_j(y) = (-1)^j \exp(y^2/2) d^j(\exp(-y^2/2))/dy^j$;
C_j = j^{th} coefficient; and
y = a bivariate normal variable.

The estimation of $Z(X_o)$ is then given as:

$$\hat{Z}_{DK}(X_o) = \sum_{i=0}^{N} \sum_{j=0}^{J} f_{ij} H_j(y(X_i)) \qquad (2.7)$$

where,

$\hat{Z}_{DK}(X_o)$ = disjunctive kriging estimate of $Z(X_o)$;
N = number of measurement points;
i = sample set; and
f_{ij} = a constant which depends on i and j.

The above approach, despite its mathematical elegance, displays some practical limitations. For instance, in

pratice, it has a tendency to only estimate the first 2 or 3 elements of the expansion series in (2.6). If the variable of interest requires a more extensive expansion, disjunctive kriging may lose its advantage. This deficiency may force the disjunctive kriging to yield inferior results.

Another approach, which is utilized by many authors in geohydrology, is to assume the that the non-negative variable of interest is log-normally distributed, see Delhomme (1974), Freeze (1975), and Neuman (1982). In many instance there are physical evidence to support such an assumption. For instance, in the case of transmisivity data, as Ahmed and DeMarsily (1987) note, measured values usually exhibit a wide range of magnitudes, many orders of difference, while their histograms are close to lognormal. Furthermore, Matheron (1967) shows that in two dimensions and for parallel flow in a heterogenous medium, the correct average transmissivity is the geometric mean, which is given by the arithmetic mean of the logarithm of the transmissivty. So, for the above reasons we decided to utilize a log-normal approach in our study.

In this approach the variable of interest Z(X) is defined as the log-transform of the original variable Y(X), such as:

$$Z(X) = \ln(Y(X)) \tag{2.8}$$

Afterward, Z(X) is treated exactly as described in Section 2. Different properties of Y(X) can then be estimated as:

$$E(Y(X)) = \exp[E(Z(X)) + Var(Z(X))/2] \tag{2.9a}$$

$$Var(Y(X)) = [E(Y(X))]^2[\exp(Var(Z(X))-1] \tag{2.9b}$$

also,

$$m(Y(X)) = \exp[E(Z(X))] \quad (2.10)$$

$$Y_{\alpha}(X) = \exp[E(Z(X)) + z_{\alpha}\sqrt{Var(Z(X))}] \quad (2.11)$$

where,

$m(\ )$ = median;

$Y_{\alpha}(X)$ = the risk value of Y at X whose probability of exceedence is α percent; and

z_{α} = the standardizd normally distributed random variable with a probability of exceedence of α percent.

2.1.1. Case Study : Mapping of Geohydrological Parameters in the Shallow Aquifer of Dougherty Plain, Southwestern Georgia

This case study was jointly supported by a grant from U.S. Geological Survey (USDI/USGS Project G-1219(05)).

The Dougherty Plain which is located in the southwestern corner of Georgia, as shown in Figure 2.1.1.1, is a rapidly growing agricultural region. This area is underlain by a succession of sand, clay, and carbonate rocks to a depth of more than 5,000. ft., forming one the most productive multilayer aquifers in the country. In our study we focused on the shallow aquifer, which is the main recharge route to the principal artesian aquifer. This latter aquifer is the main source of groundwater in this region that has sustained the agricultural growth of the Dougherty Plain.

The above growth has been accompanied by a drastic increase in the use of fertilizers and pestcides, some of whose components are toxic to humans, long-lasting, and tend to accumulate in the hydrogeological system. This poses an obvious threat to the quality of groundwater from the

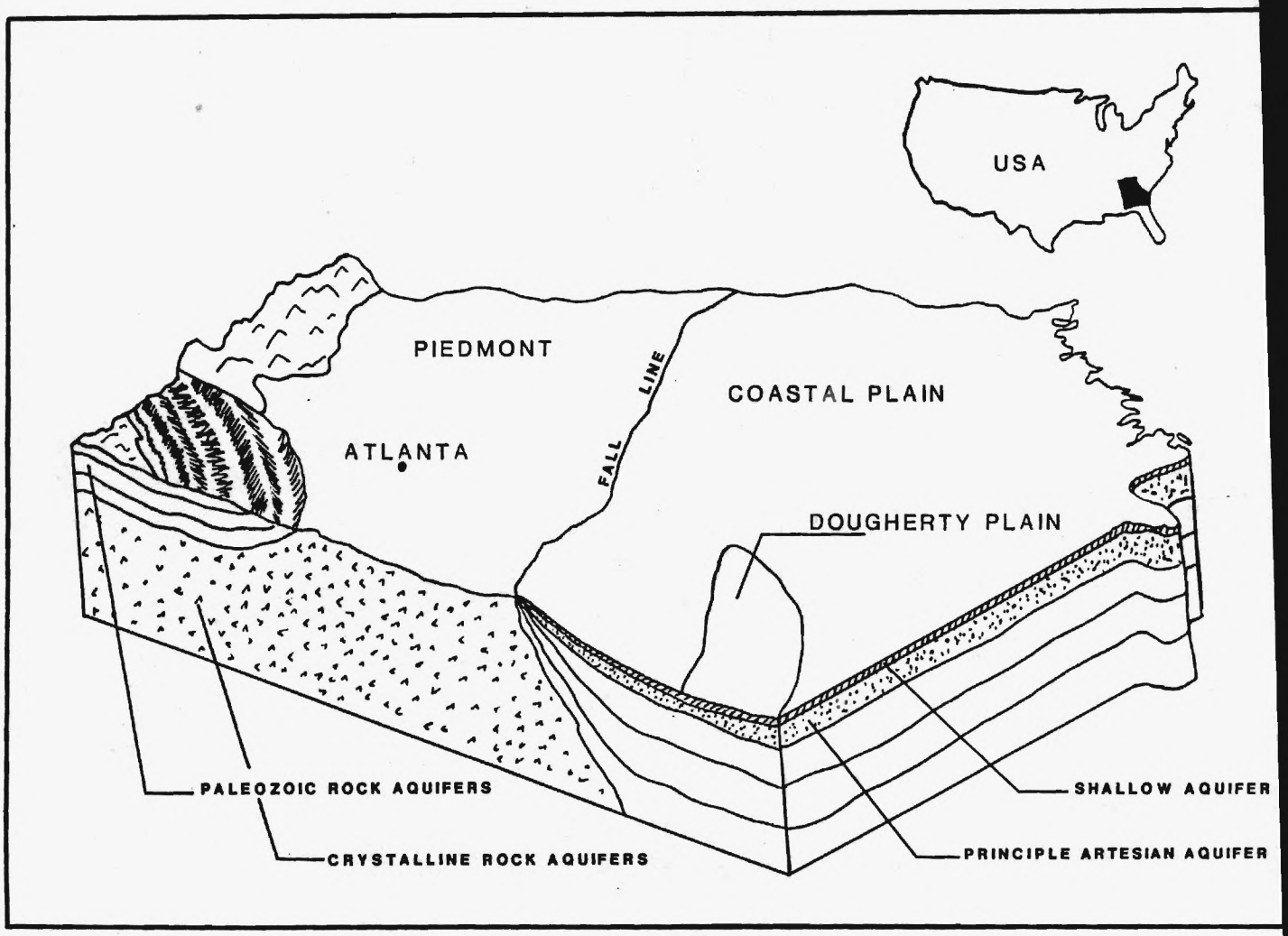


Figure 2.1.1.1 Dougherty Plain, Georgia

principal artesian aquifer.

An important step in the understanding of this phenomenon is the study of the hydraulic characteristics of the shallow aquifer. The available values are given in Table 2.1.1.1. As seen, all the variables exhibit wide spatial variations, which at some instances amount to differences of many order of magnitude.

In addition to the above lateral variations, vertical conductivity also dispalys a decreasing trend with respect to the depth. Test drillings indicate that permeable sand layers occur more commonly in the upper half of the shallow aquifer than its lower half. On this basis, Hayes et al. (1983) suggest to consider the lower half of the shallow aquifer as the leaky layer between the shallow and the principal artesian aquifer. To derive a measure for the potential recharge from the shallow into the principal artesian aquifer, they further define leakance, L, as the rate of recharge per unit horizontal area per unit hydraulic head difference:

$$L(X) = K_v(X)/b(X) \qquad (2.12)$$

where,

L(X) = leakance at X in (ft/d)/ft;
$K_v(X)$= vertical hydraulic conductivity of the shallow aquifer at X in ft/d; and
b(X) = thichness of the lower half of the shallow aquifer at X in ft.

Estimated values of leakance vary from .00001 1/d to 0.36 1/d, with a median of 0.00172 1/d. Our objective is to utilize universal kriging to determine the regional structures of these variables.

WELL NUMBER	WELL NAME	NORTH (miles)	EAST * (miles)	LEAKANCE (ft/d/ft)
--BAKER COUNTY----------				
38	T. RENZE RW	50.0	38.3	0.00762
39	JO-SU-LI TW1	42.0	38.1	0.00002
--CALHOUN COUNTY----------				
24	B. JORDAN TW1	59.4	41.8	0.00003
--DECATUR COUNTY----------				
44	DPG	30.3	38.3	0.00019
45	J. HALL TW2	27.0	46.2	0.00010
46	G. BOLTON TW2	23.0	33.4	0.00038
47	A. NEWTON	16.7	25.1	0.00755
--DOUGHERTY COUNTY----------				
69	SCHOOL BUS ROAD TW1	61.1	59.1	0.00006
70	GAME AND FISH TW1	66.3	49.1	0.00042
71	NILO TW3	58.5	54.3	0.00800
72	USMC SUPPLY TW1	64.3	66.4	0.00007
--EARLY COUNTY----------				
45	I. NEWBERRY TW2	51.9	30.6	0.00005
46	V. EVANS	38.0	12.4	0.00013
--LEE COUNTY----------				
40	PIED. PLANT FARM TW1	74.3	62.3	0.00013
41	S. STOCKS TW1	69.2	63.3	0.36000
42	B. KING TW1	83.0	64.3	0.00008
43	H. USRY TW1	86.0	59.1	0.00001
--MILLER COUNTY----------				
16	DP3	37.0	19.9	0.00002
33	J. FLEET TW2	53.0	29.2	0.00049
--MITCHELL COUNTY----------				
34	H. MEINDERS TW2	31.2	43.4	0.00034
35	C. BOLTON TW2	43.5	50.4	0.00010
36	H. DAVIS TW1	34.8	47.6	0.00010
39	DP12	45.8	49.1	0.00003
--SEMINOLE COUNTY----------				
27	RODDENBERRY TW2	18.4	16.0	0.00003
28	D. HARVEY TW2	31.2	11.3	0.00007
--SUMPTER COUNTY----------				
22	E. STEPHENS TW1	91.5	70.0	0.00005
--TERRELL COUNTY----------				
14	A. VANN TW1	71.0	63.3	0.00001
--WORTH COUNTY----------				
5	DP9	75.5	66.4	0.00250
9	C. ODOM TW1	81.0	74.8	0.00003

* The origin corresponds to 30° 38' North, 85° 10' West.

Table 2.1.1.1 Hydraulic Data for the Shallow Aquifer Test Wells (Source, Hayes et al., 1983).

At the first step we have to select an appropriate distributio function for our variables. Freeze (1975) states that most field studies have indicated that the log-normal distribution is a suitable function to describe the statistical variations of the transmissivity data. Ahmed and DeMarsily (1987) also note a number of studies which confirm the above. The same argumet can also be applied to the hydraulic conductivity.

Considering that the transmissivity is the product of the hydraulic conductivity and the saturated thickness, we can infer that the average saturated thickness is also log-normally distributed. This is based on the principle that the products or ratios of log-normally distributed variables are also log-normally distributed (Benjamin and Cornell, 1970). By extension, the leakance which is the ratio of vertical hydraulic conductivity and the confining layer of the shallow aquifer (defined as its lower half), is also log-normally distributed.

Our initial statistical structrual analysis and mapping based on the normal distribution assumption produced unreasonable results which were interpreted as an indicator that these variables are not normally distributed. In contrast, the log-normal assumption yielded reasonable results that confirm our theoretical argument that these varaibles appear to be log-ormally distributed.

The results of structural analysis are given in Table 2.1.1.2. As shown, all the three variables have constant drifts with linear covariance functions, which is equivalet to the cases of stationary random fields with linear varigrams. The statitical similarities among these variables stem from their explicit and implicit geohydrologic relationships.

Hydraulic Conductivity	Estimated Statistical Properties					$\hat{\rho}$
	K	Covariance				
		C	a_1	a_3	a_5	
Leakance (L)	0	0	-1.1552	0	0	1.4822
Vertical Conductivity (K_v)	0	0	-1.2366	0	0	1.5176
Transmissivity (T)	0	0	-1.3433	0	0	1.7656

Table 2.1.1.2 Results of Structural Analysis.

As a measure for the goodness-of-fit of the estimated covariance functions, we have utilized a jackknife estimator for ρ: the ratio of the actual sum of squared errors of the estimation and the theoretical sum of krigig variances (Rouhani, 1983). A perfect fit results in a value of 1 for ρ. The indicated values in Table 2.1.1.2, thus, display a reasonable and satisfactory degree of goodness-of-fit.

For mapping puposes the Dougherty Plain area is divided into a 20 x 22 grid with 5 mile increments. The actual maps are then produced by the contour program of DISSPLA version 9.2, which is available at Georgia Tech's OCS Cyber computer. For the sake of brevity and due to the similarities among the variable of interst only maps associaed with leakance are presented. For a detailed presetation of these maps readers are referred to Rouhani and Hall (1987).

The produced maps are analyzed in four different categories: maps of expected values, maps of medians, maps of estimation variances, and maps of risk values. Figure 2.1.1.2 displays the map of expected leakance, which has a relative uniform value throughout a large portion of the Dougherty Plain. However, a sudden rise is indicated in the southern tip of this region. At this region, leakance is about 4 to 10 times larger than the leakance in other parts of the Plain. So the southern tip should be considered an area with high recharge potentials.

It must be noted that due to the asymmetry of the log-normal distribution some of the indicated values in the above map are excessively high. Meyer (1975) states that : "the expectation value is not so useful for an asymmetric distribution. Often, more significant are such measures as the median, mode, and the geometric mean." So we propose to utilize the median map as the representative map of the above hydraulic variables. For such log-normally distributed

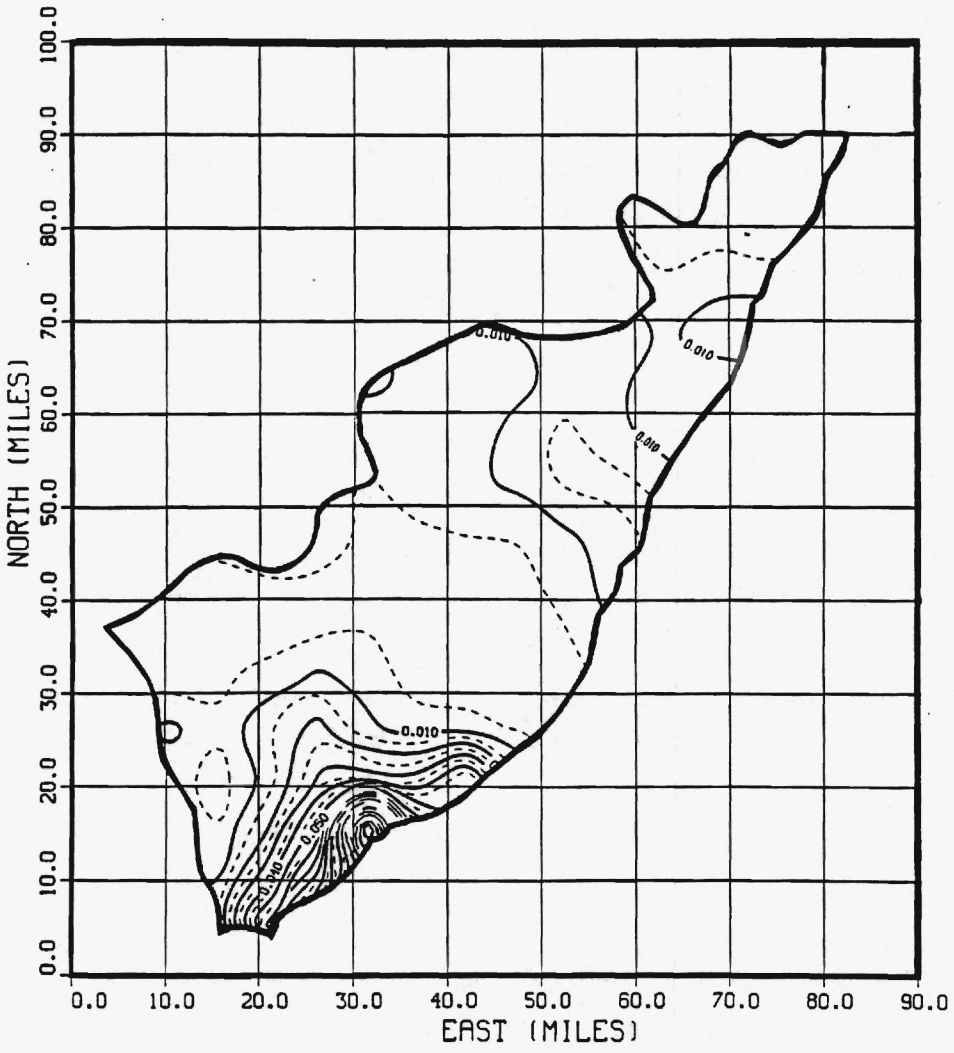


Figure 2.1.1.2 Map of Expected Leakance (1/d).

varaibles, median is equivalent to the geometric mean.

Figure 2.1.1.3 illustrates the median map for leakance. In contrast to the expected map it has more moderate values, and displays more spatial variations. This figure also points to the southern tip and the upper centeral region as the areas of high leakance.

As noted earlier, kriging provides a measure for the accuracy of its estimates. Figure 2.1.1.4 shows the estimation variance of the log-leakances. The boundry region have higher variances, due to the fact that these points are generally extrapolated, and thus, contain more uncertainty. As expected, the southern and the northern tips show higher levels of uncertainty. On the other hand, the middle of the upper half portion of the plain displays lower variances of estimation which is due to a higher concenration of measurement points in this region (see Figure 2.1.1.1).

Finally, we study the risk values, as defined by (2.11), as a measure that contains both the magnitude and the accuracy of the estiamted values. Figure 2.1.1.5 is the 10% risk map of leakance. Comparison of this map with the median map indicates a smoothing of small local variations in the risk values. This is due to the fact that at some points we have values with small leakances, but highly inaccurate, while at others we have the opposite condition. This tends toward more uniform risk values. It is also possible that inclusion of the estiamtion variance in the risk value may cover some of the local fluctuations of the median value.

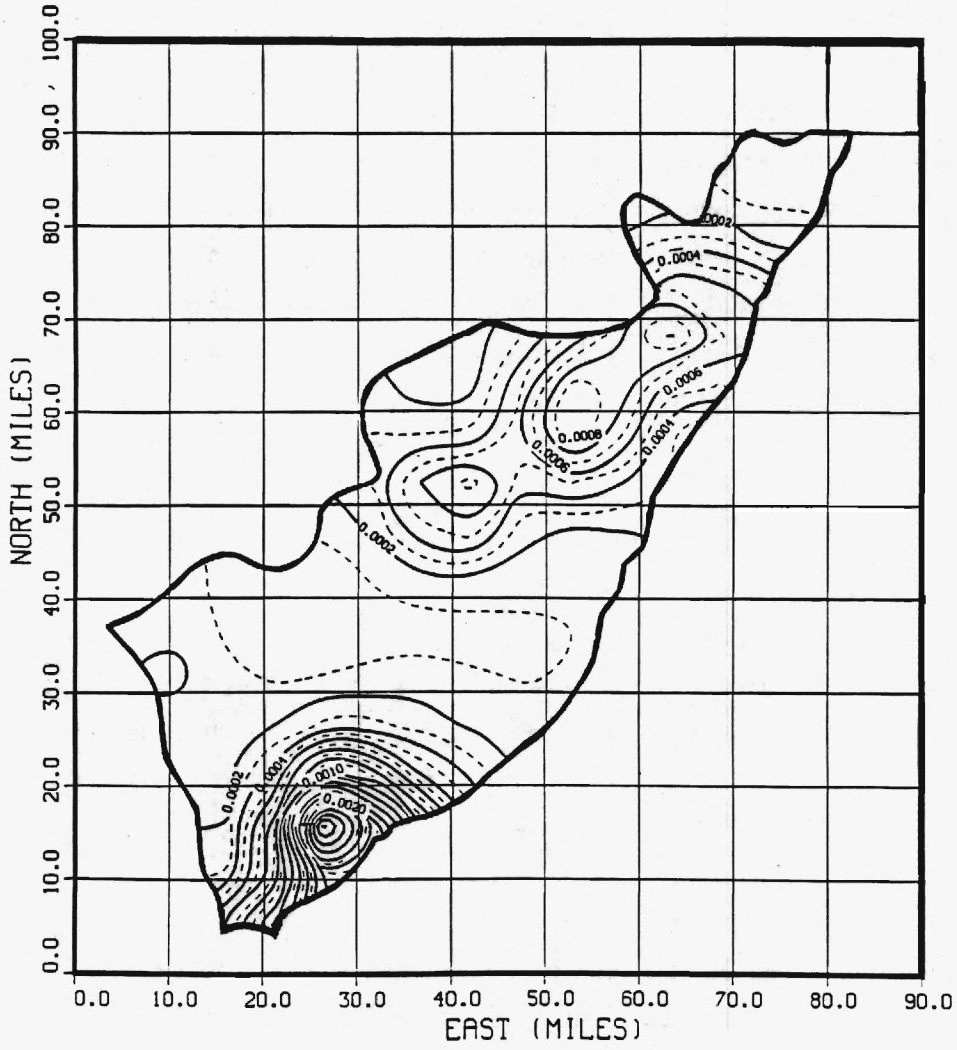


Figure 2.1.1.3 Map of Median Leakance (1/d).

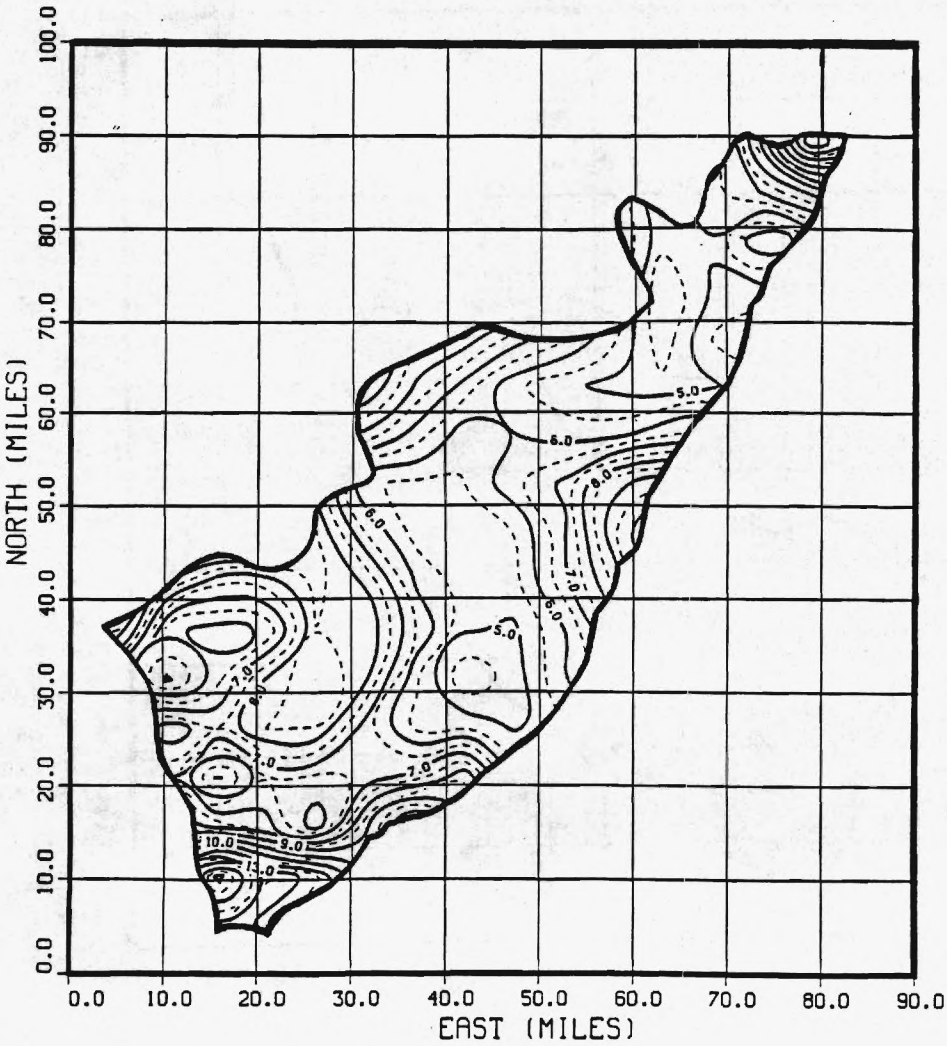


Figure 2.1.1.4 Map of Variance of Log-Leakance.

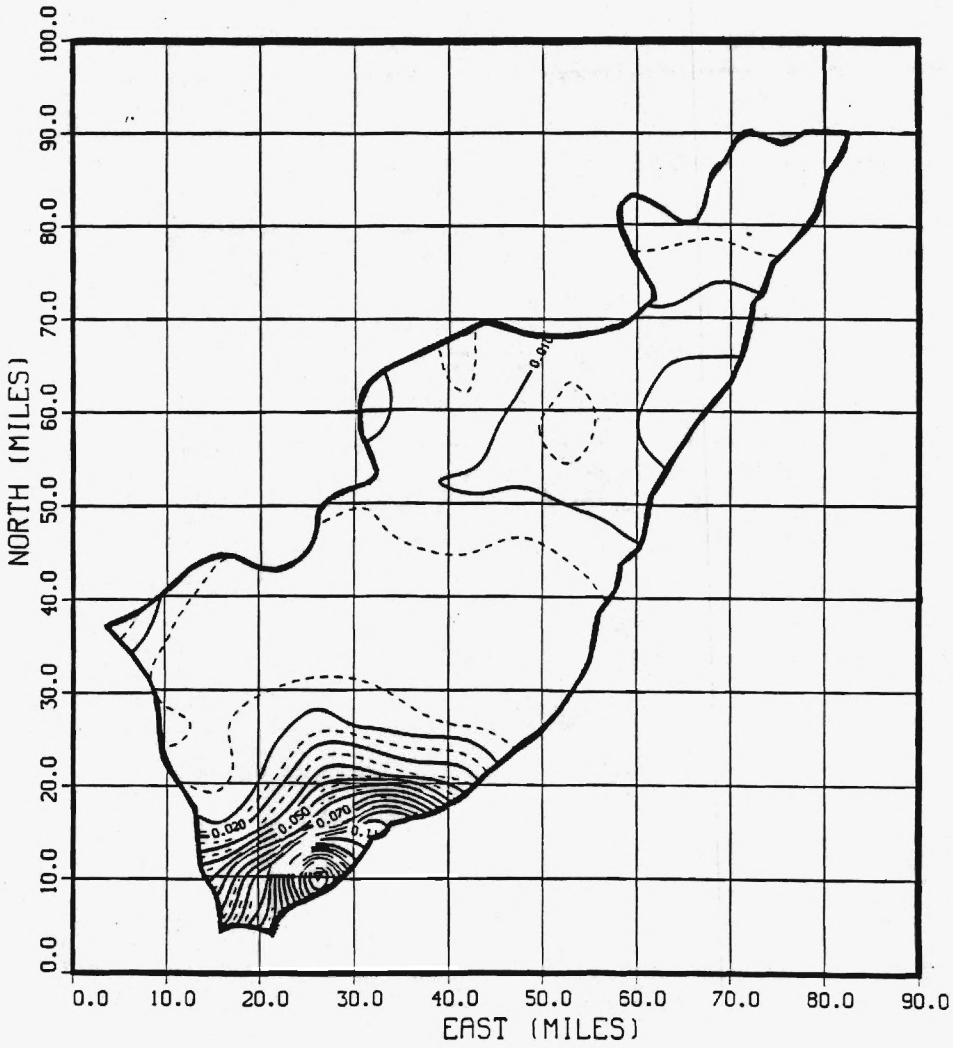


Figure 2.1.1.5 Map of 10% Risk Value of Leakance (1/d).

2.2 Time Kriging

Up to now, geostatistical procedures, including kriging, have been almost exclusively applied to spatial data. This is due to a variety of reasons, such as the fact many variables involved in mining (the original field of geostatistical applications) are only spatially distributed, or the fact that alternative estimation procedures are available for time series analysis. However, in our quest for working with spatiotemporal variables, it was essential to expand universal kriging into the time domain, which required certain modifications. First of all the two or three dimensional space vector X has to be converted into a one-dimensional vector T with coordiante (t_i). The universality conditions (2.3) are thus reduced to:

$$\begin{aligned} k = 0 \quad & \sum_{i=1}^{N} \lambda_{io} = 1 \\ k = 1 \quad & \sum_{i=1}^{N} \lambda_{io} t_i = to \\ k = 2 \quad & \sum_{i=1}^{N} \lambda_{io} t_i^2 = t_i^2 \end{aligned} \qquad (2.13)$$

Furtermore, the positive-definitness criteria of the polynomial covariance function, as indicated in Table 2.1, changes due to the one-dimensionality of time. Based on the work of Matheron (1973), for a polynomial covariance with the following form:

$$K(h) = C\delta(h) + \sum_{p=0}^{k} a_{2p+1} h^{2p+1} \qquad (2.14)$$

the positive-definitness conditions for the one-dimensional space, are:

$$a_1 \leq 0,$$

$$a_5 \leq 0, \text{ and}$$

$$a_3 \geq -2/3(30a_1a_5)^{1/2} \qquad (2.15)$$

where,

K( ) = covariance function;
C = nugget effect;
δ( ) = Dirac delta function;
a_i = i^{th} coefficient in the covariance function;
h = length of the distance vector (lag time in the case of time kriging); and
k = order of the polynomial drift function.

The above isotropic covariance function is composed of two parts. The first part is the nugget effect which represents small scale fluctuations and measurement errors. The second part is the sturctured portion that reflects the regional or large scale structure of the random variable of interest.

In order to implement the above changes in the universality conditions and the positive-definitness criteria of the covariance, some detailed modifications in the original spatial universal kriging computer program were made. For a detailed study see Appendix 5 (TKRIG).

2.2.1. Case Study: Time Kriging for Drought Management in Western Georgia

Drought is a reoccuring event in many parts of the world. Drought has been termed a creeping phenomenon. It is generally difficult to accurately predict either the onset or the end of a drought, or to even know if a drought is occuring. It is also difficult to determine the severity of

a drought, which depends on its magnitude, its duration, and its geographical extent.

Considering the above we propose to utilize time kriging to predict stream flow fluctuations in a drought prone area. This will provide a management tool for reservoir operators. The scheme is basically composed of the following steps: (1) structural analysis of the stream flow, upstream of the reservoir of interest; (2) prediction of stream flows during the critical periods; (3) applying the predicted values as inputs into an appropriate reservoir routing model; and finally, (4) calculating the length of time gap between the beginning of the critical period and the beginning of the predicted drought. This lenght is defined as the drought lead time, which can be used as a warning measure by reservoir operators and regulatory agencies.

For this purpose, we have selected a location in western Georgia,as shown in Figure 2.2.1.1, as the site of a hypothetical reservoir. In recent years this area has experinced some sever meteorological droughts, which has resulted into a significant drop in its agricultural production.

In our study, we designed a hypotetical reservoir on Brier Creek, located 6.7 miles south of Thompson, Georgia, with a drainage basin of 56. sq. miles. The basis of our design was the monthly streamflow data from a USGS gaging station just downstream of the proposed site. Using other relevant information concerning the use of water in this region, it was assumed that the normal release rate of the reservoir is equal to 40. cfs. Furthermore, we assumed that the reservoir has an operating policy, as follows:

$$R_t = S_{t-1} + I_t, \text{ if } S_{t-1}+I_t \leq T_t;$$

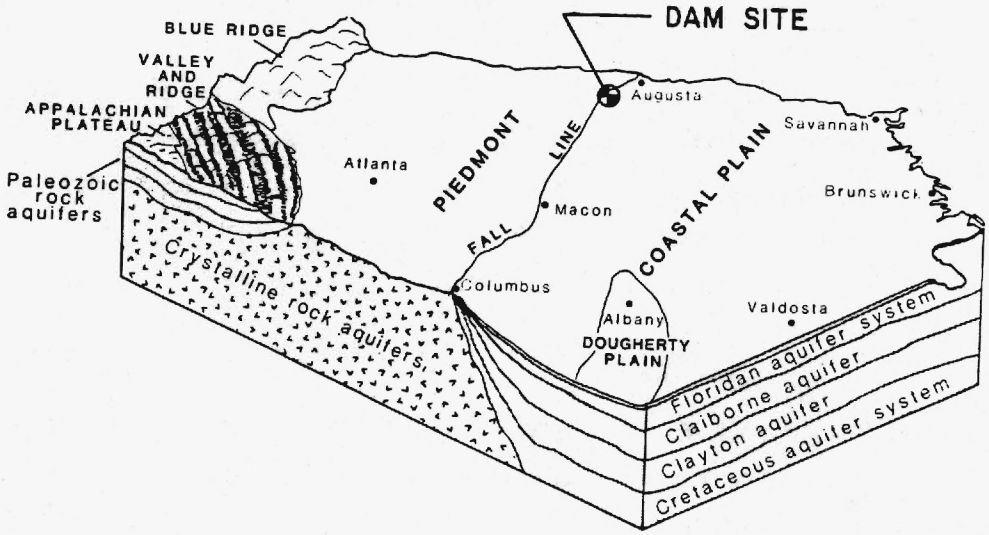


Figure 2.2.1.1 Proposed Dam Site

$$R_t = T_t, \text{ if } S_{max} \geq S_{t-1}+I_t \geq T_t; \text{ and}$$

$$R_t = S_{t-1}+I_t-S_{max}+T_t, \text{ if } S_{t-1}+I_t \geq S_{max} \qquad (2.16)$$

where,

R_t = release volume during the t^{th} month;
S_t = reservoir storage at the end of the t^{th} month;
I_t = inflow volume during the t^{th} month;
T_t = target release volume during the t^{th} month; and
S_{max} = maximum reservoir storage.

We then defined the drought in the context of water management, as the condition, at which the reservoir release falls short of the target. For other definitions of the drought readers are referred to White and Glantz (1985).

Available records of average monthly streamflows are provided by the U.S. Geological Survey, for the period of January, 1973 to December, 1982, as displayed in Figure 2.2.1.2. The critical period is chosen for the eight month period from March to October, which is the low-flow season. It is during this period that water management drought is most likely to occur.

For the selection of the appropriate distribution function for streamflow data, we first drew the log-streamflow histogram, as shown in Figure 2.2.1.3. This graph indicated a non-normal bi-modal distribution. For this reason log-normal distribution is rejected.

We then examined the monthly fluctuation data, which are defined as the differences between the actual flow in any month and the long-term average of flow for that month. The resulting histogram, as displayed in Figure 2.2.1.4, shows a nearly normal distribution. So, we selected fluctuation data

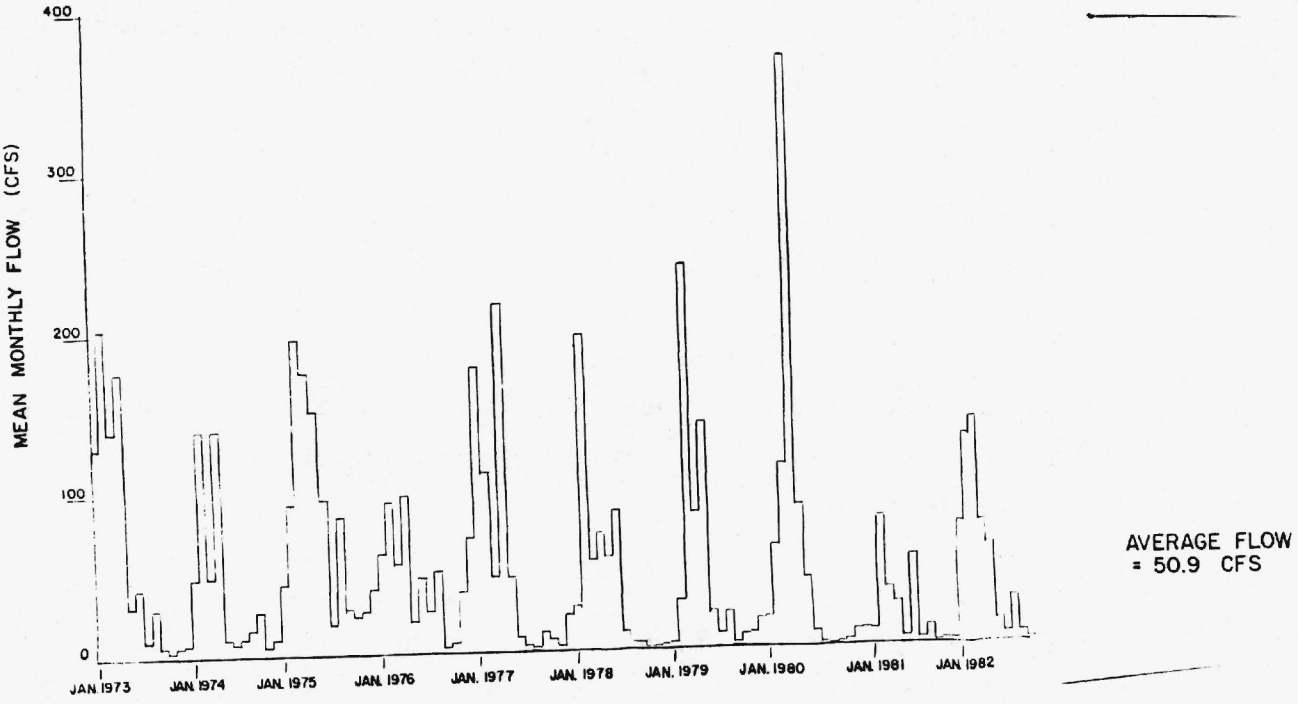


Figure 2.2.1.2 Hydrograph of Mean Monthly Streamflows (cfs).

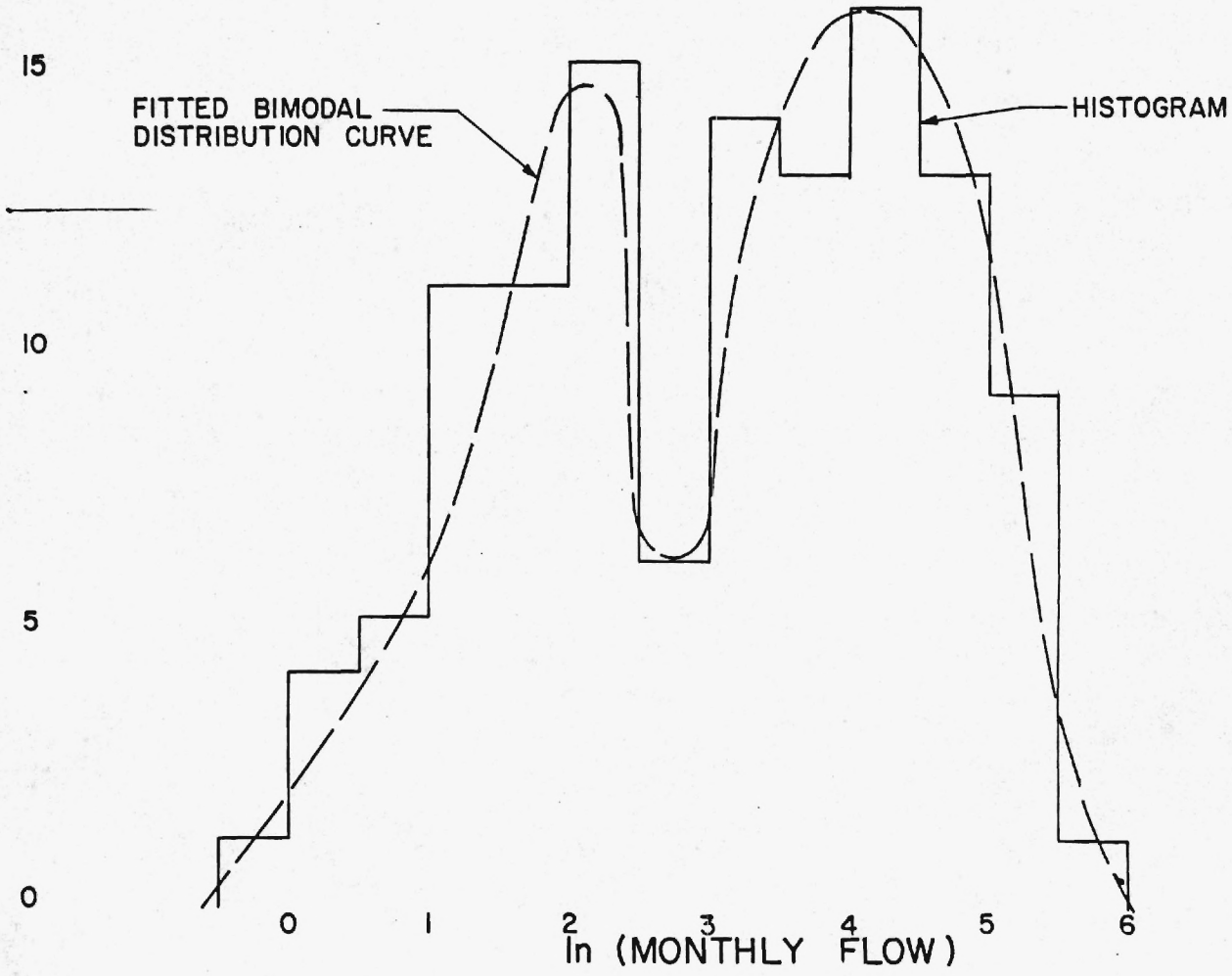


Figure 2.2.1.3 Histogram of Mean Monthly Streamflows

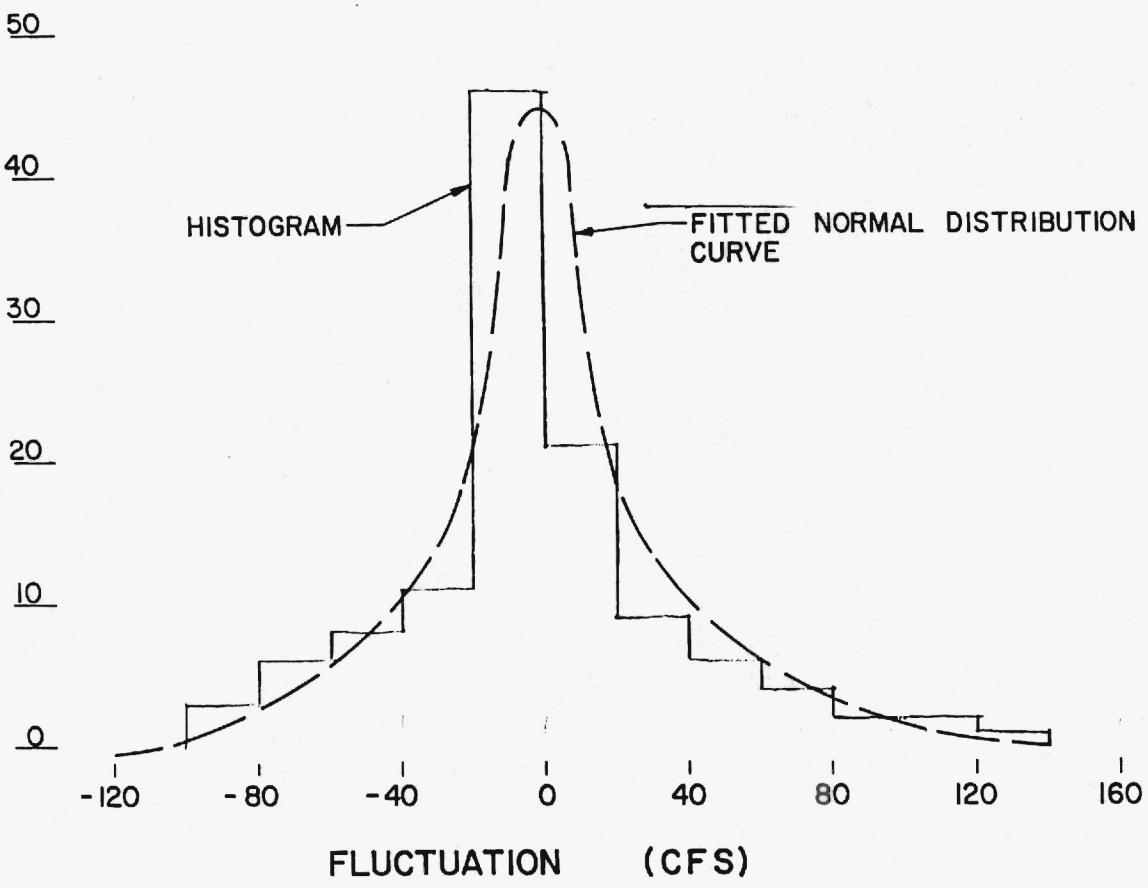


Figure 2.2.1.4 Histogram of Monthly Fluctuations

as the basis of our analysis.

The structural analysis also confirms that the order of the drift function for the data is 0. Three covariance functions are then derived as alternatives. The first one is a pure nugget effect (C=2181.7, a_1=0) which results in equal λ_{io} for all the measured values in the estimation process, regardless of their temporal vicinity to the estimated time. The second one does not have any nugget effect (C=0, a_1=3149.7) which gives the whole weight to the nearest measured value and ignores all the rest. This is an extreme case of shadow effect which reduces the impact of data beyond the first ring of measurements (David, 1977). The third one contains both the nugget effect and the correlated portion (C=1696.53, a_1=-56.06). This third option yields more realistic weights by giving the highest weight to the nearest data and then decreasing as the time lag increases. This covariance was also confirmed by an experimental variogram. The range of the variogram is approximately estimated as 8 months. This means that fluctuations with 8 or more months of lag between them, have no significant covariance.

Using the third covariance, at each March (beginning of the critical period) we estimate a sequence of eight monthly fluctuations, based on the available data prior to March. Each estimated fluctuation is then added to the long-term monthly average flow for that month to produce a sequence of eight monthly flows. These values are then used as inflow data, I_t, in the routing procedure, to calculate the release rates, R_t, during the critical period. The first month that indicates an R_t less than the target release is identified as the beginning of the drought period. The lag between this month and the begining of the proceeding March is defined as the drought lead time- a measure of warning for reservoir operators about the possibility of a drought.

In order to test the validity of our procedure the above scheme is repeated using the actual data for the critical periods between 1972 and 1980, and their results are compared to time kriging results. Figure 2.2.1.5 shows the results of the predicted flows. Generally, they are poor estimates. This is due to two reasons. First, the predicted values are in fact extrapolated values, and thus, contain a significant amount of uncertainty. The second reason is particular to our data set which shows a poor correlation between the fluctuation values. However, it will be seen that these results still provide a reasonable estimate of the drought lead time, whose estimation is our main target.

At the next stage, the routing procedure is repeated for both the estimated data, and the actual data, using initial storage values of 0, 500, 1000, 1500, 2000, 2400, 5000, 8000, 10,000, 15,000, 20,000, 25,000, and 29,040 acre-feet. Thus, we cover the whole range of possible initial storages, from 0 to S_{max}, which produced 208 trials.

In the above runs, as expected, the drought lead time showed a positive correlatoin with repect to the initial storage S_o. Furhtermore, as S_o increases, the correlation between the actual and predicted drought lead time (N) increases as well. As indicated in Figure 2.2.1.6, for S_o of 8000 acre-feet and up, this correlation is 1, which means a perfect estimation. Figure 2.2.1.7 displays the actual and predicetd drought lead times for different initial storages. The correlation is quite satisfactory.

The drought lead time appears to be a useful tool for providing a warning system for reservoir operators. The time kriging plays a pivotal role by providing predictions for the inflows. For a more detailed study of this case study readers are referred to the M.S. special research of K.A. Cargile (1987), conducted under the supervision of the PI.

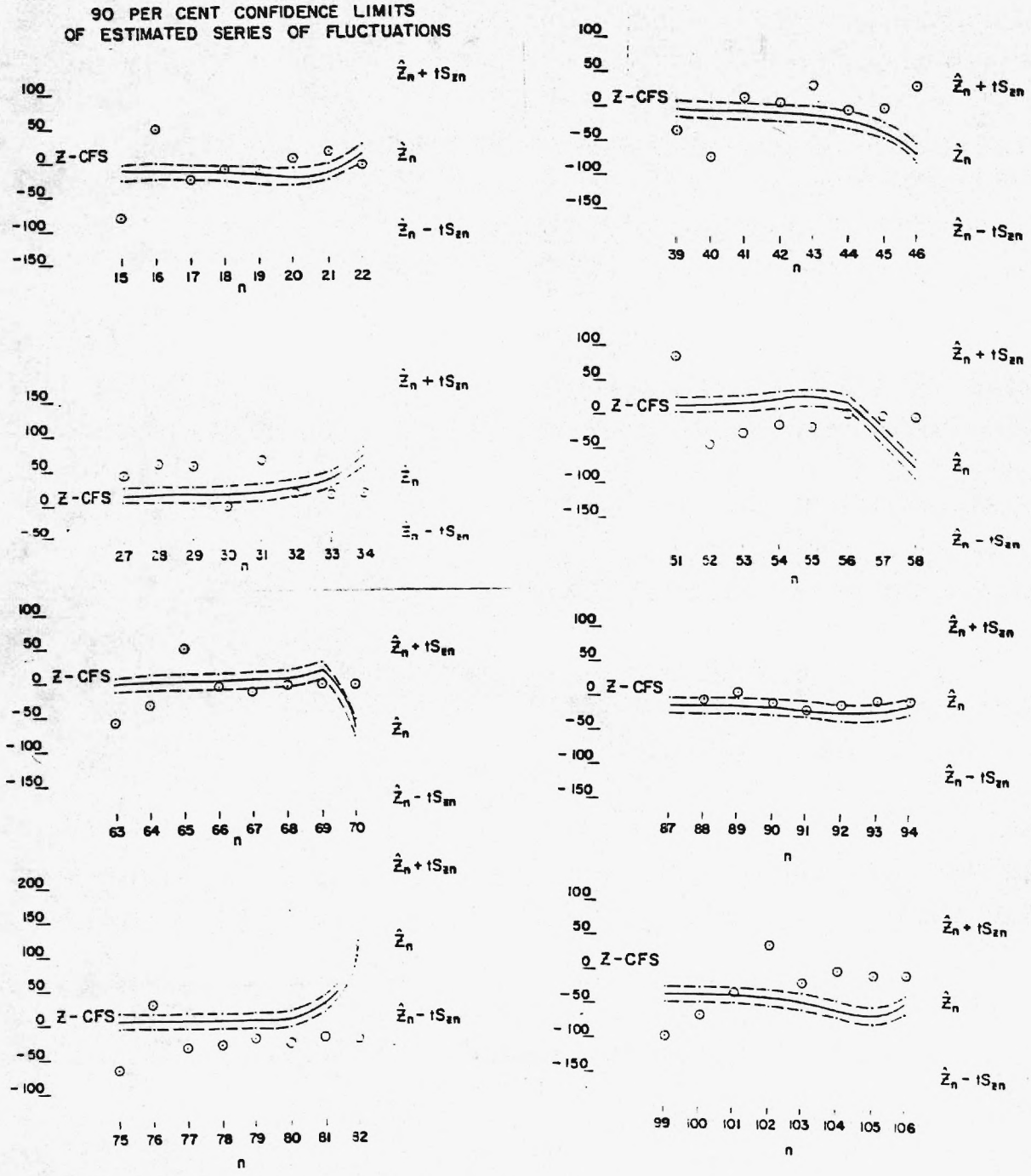


Figure 2.2.1.5 Estimated vs. Actual Flows

LEGEND

⊙ ACTUAL FLUCTUATION DATA
—— PREDICTED FLUCTUATION DATA
– – 90 % CONFIDENCE LIMITS

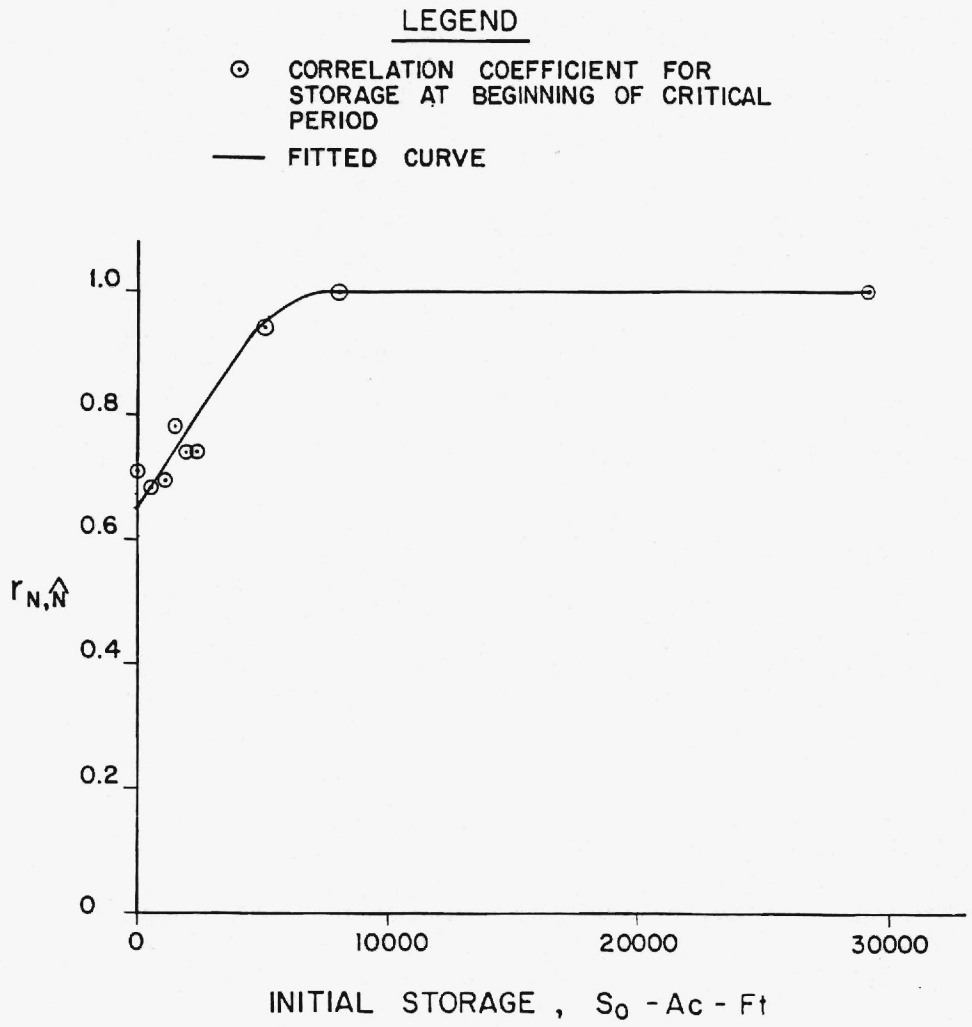


Figure 2.2.1.6 Initial Storage vs. Correlation between Estiamted and Actual Drought Lead Times.

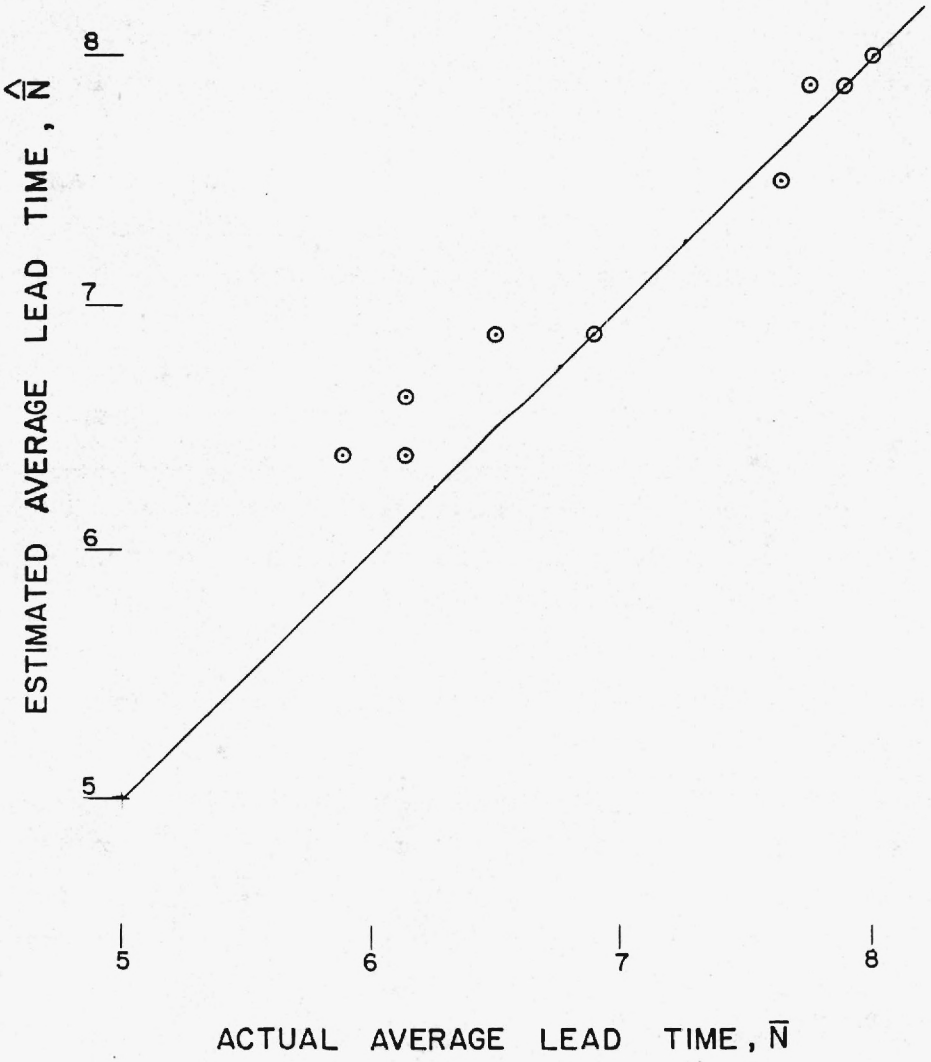


Figure 2.2.1.7 Estimated vs. Actual Drought Lead Times.

2.3 Universal Space-Time Kriging

As noted earlier, our eventual objective is to expand universal kriging to the space-time domain. This expansion has a multitude of benefits for hydrologists. First of all it enables us to utilize our space-time data efficiently. As noted before, in a number of past studies, authors have to divide their data base into segments for each time interval, and then apply spatial geostatistical techniques to each subsection (Delhomme, 1977), or use temporally integrated values as the basis of their analysis (Chua and Bras, 1980).

Space-time universal kriging, on the other hand, provide us with a tool to study the spatiotemporal variable of interest, without forming artificial divisions or integrated values based on temporal occurances. It also allows hindcasting and forecasting, which are not possible with spatial kriging. Finally, as it will be shown, in most cases, the maps based on space-time kriging are more accurate than the ones by spatial kriging. This is due to the fact that space-time kriging is able to utilize a broader range of information than the spatial kriging.

Our survey reveales that only few authors have ever attempted to use kriging simultaneously in both the space and time dimensions. Most notably, Bilonick (1985, 1987) has exapnded kriging into time space by utilizing anisotropic variograms, where the time is treated as the $(n+1)^{th}$ dimension. He has applied this prosedure for mapping of sulfate and ion deposition in the northeastern U.S. To our knowledge, this project is one of the first attempts to expand and uitlize space-time universal kriging in hydrology.

Expansion of universal kriging into the time-space domain requires two important steps. In the first step we propose to exapnd the universality conditions by including

both the spatial and temporal drifts. We then search for an appropriate family of spatiotemporal covariance functions.

For the expansion of the universality conditions, we assumed that similar to the space domain, the random variable of interest may have a temporal polynomial trend of some order. We define orders k_s and k_t, as the order of the polynomial drift in the space and time domain, respectively. We expand the universality conditions by combining (2.3) and (2.13). For the case of k_s=2 and k_t=2, the universal kriging system with N data point for estimation at X_o with coordinates (x_o, y_o, t_o) is:

$$\begin{bmatrix} 0 & 0 \dots 0 & 0 & 0 & 0 & 1 \dots & 1 \\ 0 & . & . & 0 & 0 & x_1 & x_N \\ . & . & . & . & . & . & . \\ . & . & . & . & . & . & . \\ 0 & 0 \quad 0 & 0 & 0 & 0 & y_1^2 & y_N^2 \\ 0 & 0 \quad 0 & 0 & 0 & 0 & t_1 & t_N \\ 0 & 0 \quad 0 & 0 & 0 & 0 & t_1^2 & t_N^2 \\ 1 & x_1 \dots . & y_1^2 & t_1 & t_1^2 & K_{11} & K_{1N} \\ 1 & x_2 \dots . & y_2^2 & t_2 & t_2^2 & K_{21} & K_{2N} \\ . & & & & & . & . \\ . & & & & & . & . \\ 1 & x_N & y_N^2 & t_N & t_N^2 & K_{N1} \dots & K_{NN} \end{bmatrix} \begin{bmatrix} \mu_{1o} \\ \mu_{2o} \\ . \\ . \\ \mu_{6o} \\ \mu_{7o} \\ \mu_{8o} \\ \lambda_{1o} \\ \lambda_{2o} \\ . \\ . \\ \lambda_{No} \end{bmatrix} = \begin{bmatrix} 1 \\ x_o \\ . \\ . \\ y_o^2 \\ t_{o_2} \\ t_o^2 \\ K_{1o} \\ K_{2o} \\ . \\ . \\ K_{No} \end{bmatrix}$$

(2.17)

where, K_{ij} is defined as the covariance between X_i and X_j, λ_{io} is the kriging weight for the i^{th} data point, and μ_{po} is the Lagrange multiplier for the p^{th} monomial. In the above case there are eight monomials as: 1, x, y, xy, x^2, y^2, t, and t^2.

Our initial search for an appropriate family of covariance functoins provided us with a number of

alternatives (Rouhani, 1986x). For example, in our pilot study, we constructed a composite polynomial covariance function to cover both the space and the time dimension. This approach is similar to deriving covariance functoins for anisotropic random variables, or for spatial fields with a nested statistical structures. It is based on the following properties of positive definite functions (Journel and Huijbregt, 1978), that: (1) every linear combination of covariances with positive coefficient is a covariance, and (2) any covariance product is also a covariance. For variograms only the first rule is applicable.

The above rules indicates that there are numerous models, as well as, their combinations that can be used as models for our study. The question that arises is which one of these is more suitable? For instance, we use polynomial covariance functions for the sake of operational efficieny. However, in the course of our study we encountered some difficulties, which were mainly due to the fact that covariance estiamtes based on actual data display a lack of robustness. Small changes in the data set causes significant fluctuations in the estimated parameters of the covariance function. Furthermore, there are some tendency in the proposed algorithm by Delfiner (1975) to yield covariances with large nugget effects. Consequently, the choice of the best fitted polynomial covariance becomes rather subjective, which in turn, reduces the efficiency of the scheme.

Rouhani (1985 and 1986) and Rouhani and Fiering (1986) discuss the above problems in detail. They conclude that despite fluctuations in the estimated covariance functions, kriging estimates show a high degree of stability. Journel and Huijbregt (1978) go even further and state that: "the results of the geostatitical calculations prove to be robust in relation to the choice of the (covariance or variogram) model - provided that the parameters of this model are

correctly estimated."

After the above considerations, we decided to continue using the family of polynomial covariance functions as the basis of our analysis, as defined in (2.15), such that:

$$K(h,t) = K_s(h) + K_t(t) \qquad (2.18)$$

where,

K = spatiotemporal covariance;
K_s = spatial polynomial covariance;
K_t = temporal polynomial covariance;
h = space lag; and
t = time lag.

Structural analysis is then performed by adding a tolerance limit to both the space and the time dimensions. For instance, while estiamting the parameters of K_s, we assumed that any data point located at $t_i \pm \varepsilon_t$ can be considered at the same time interval as the i^{th} data point. Likewise, in the estiamtion of K_t, the points located at a radius of ε_s of each other are considered to belong to the same time series. This allowed us to simultaneously conduct the spatiotemporal stuctural analysis. For a detailed description of the above algorithm, readers are referred to the M.S. special research of T.J. Hall (1987), conducted under the supervision of the PI. The program itself is given in the Appendix 5 (STVARED).

2.3.1. Case Study: Space-Time Mapping of Groundwater Data in Southern Georgia

The data for this study is from a U.S. Geological report by Clark et al. (1985). We selected a study area of 110 miles by 80 miles in southern Georgia with eight sampling

Figure 2.3.1.1 Study Area in Southern Georgia.

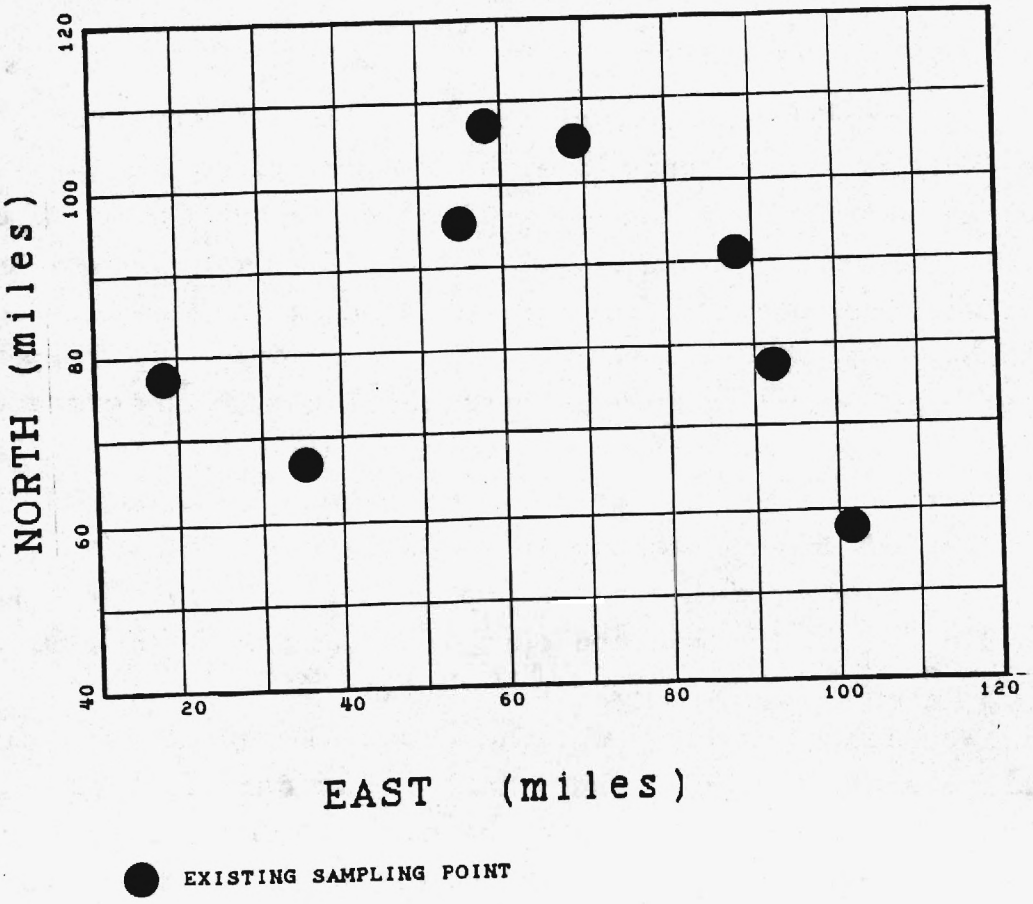


Figure 2.3.1.2 Location of Existing Sampling Points.

points, as shown in Figures 2.3.1.1 and 2.3.1.2. The sampling points are not uniformly distributed. The upper-centeral portion of the study area has a relatively higher concentration of points, while the lower-centeral portion has no data points.

The basis of our study are monthly averages of water table elevations in 1984 for each of these points. In general, water table elevations range from less than 100 feet above MSL in the lower portion of the study area, to well over 200 feet above MSL in the upper portions.

The pre-kriging structural analysis is conducted on both the time and the space domain. However, we encountered a problem which was caused by ill-conditioned matrices in the kriging system, given by (2.17). This was caused by the similarities between rows associated with sampling points with same spatial locations. In other words, differences in time lags were relatively insignificant when compared to spatial lags. To solve this problem, we use a set of scales to create more homogenous values for space and time. After a series of trials, we choose to divde spatial coordinates by 10, and the temporal coordinates by 1.2. In this way all the coordinates of the measurement sites vary between 0 to 10.

We first estimated the spatial structure of the variables. It is found that the order of the spatial polynomial trend is 2. The best fitted polynomial covariance is also determined to have the following coefficients ($C=5$, $a_1=-28$). The jackknife estimator of ρ for the chosen function is 1.0233, which indicates a satisfactory fit.

We then analyzed the temporal structure of the data. Our analysis indicates a linear drift. As indicated in the Section 2.2.1 the covariance function should contain a nugget effect, as well as, the structured portion. This way the

resulting weights give a gradually decreasing value as the lag time between the sampled point and the estimated point increases. After some trials, we selected a linear covariance function with the following coefficients as: C=5 and a_1=-0.5. The jackknife estimator of ρ is 1.0053, which is an indication of a very good fit.

At the next stage we performed universal kriging for a 12 x 8 grid with 10 miles increments. For the sake of brevity we only present some of the results. Figure 2.3.1.3 displays a hindcasted map for the middle of March (3.5 months), while Figure 2.3.1.4 shows its variance of estimation. It must be noted that with spatial kriging, such maps cannot be estimated. Universal space-time kriging is also capable of forecasting. Figures 2.3.1.5 and 2.3.1.6 exhibit forecasted map for 14^{th} month and its estiamtion variance. The similarity of the variance maps is due to the fact that in our initial studies we imposed a weight to give more preference to temporal data. We later abandon this approach by using scale factors, instead. This scheme produced more reasonable results. It must be noted that forecasted maps usually lack the desired accuracy. However, for short term predictions they provide a reasonable map, as indicated in our example.

Finally, we present the results of spatial mapping with and without the use of temporal data. Figures 2.3.1.7 and 2.3.1.8 show the estimated map of piezometric surface for the first month along its variance map, using only the eight available values for the first month. Now compare these maps to Figures 2.3.1.9 and 2.3.1.10 which display the same, however, based on universal space-time kriging. The most striking feature of this comparison is the significant improvement in estimation variances after the inclusion of the temporal data. In short, universal space-time kriging provides a tool for forecasting and hindcasting, as well as,

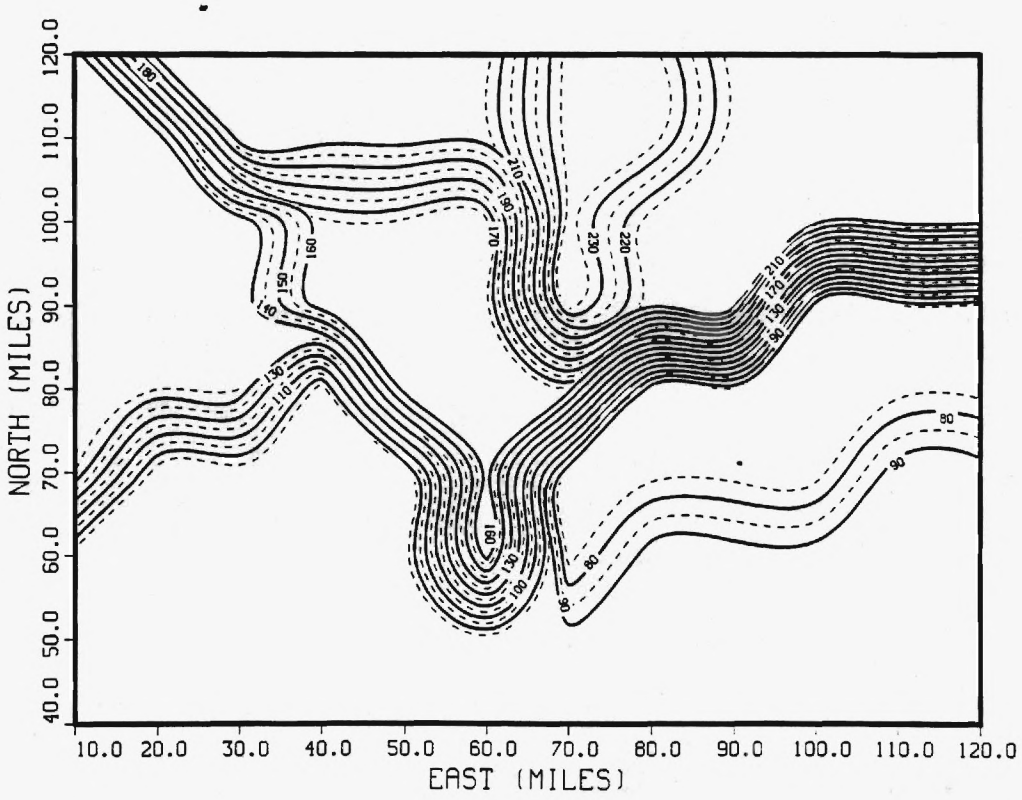


Figure 2.3.1.3 Piezometric Surface at Mid-March, 1984.

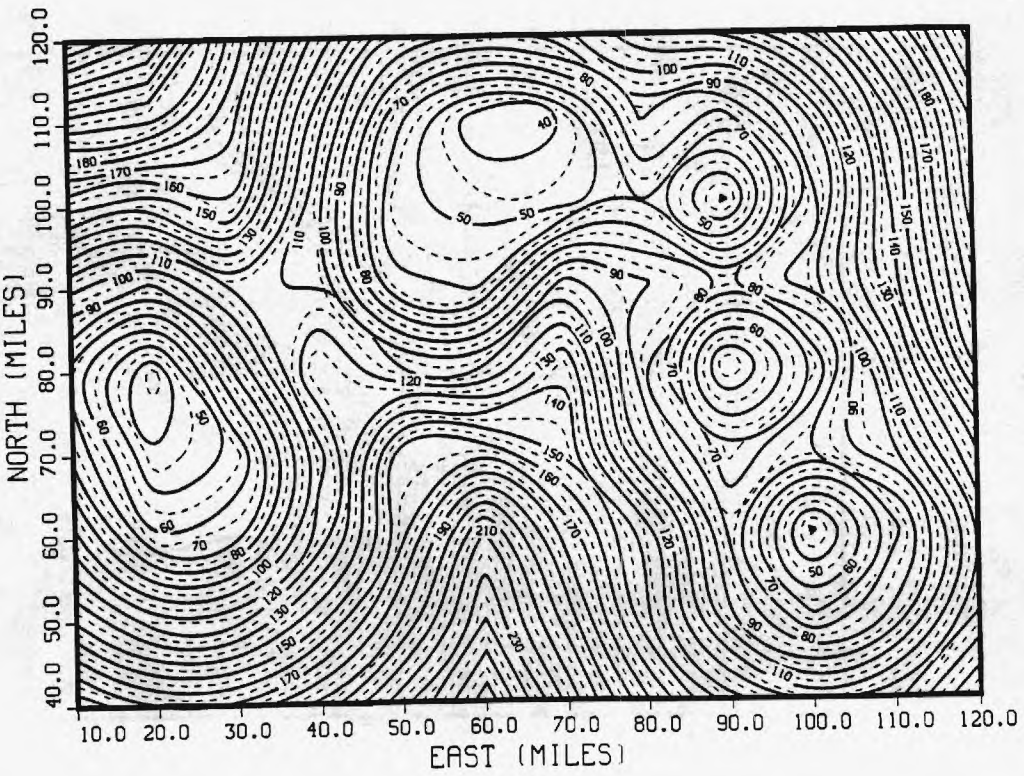


Figure 2.3.1.4 Estimation Variance at Mid-March, 1984.

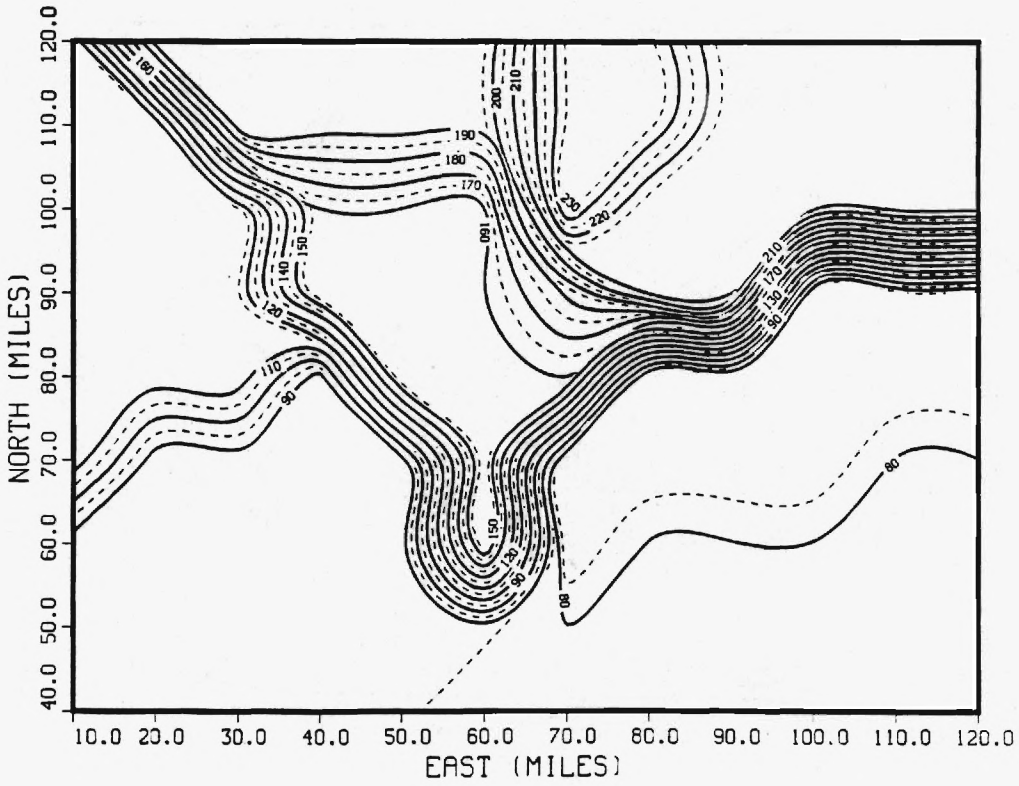


Figure 2.3.1.5 Piezometric Surface at 14 Months (Feb. 1985).

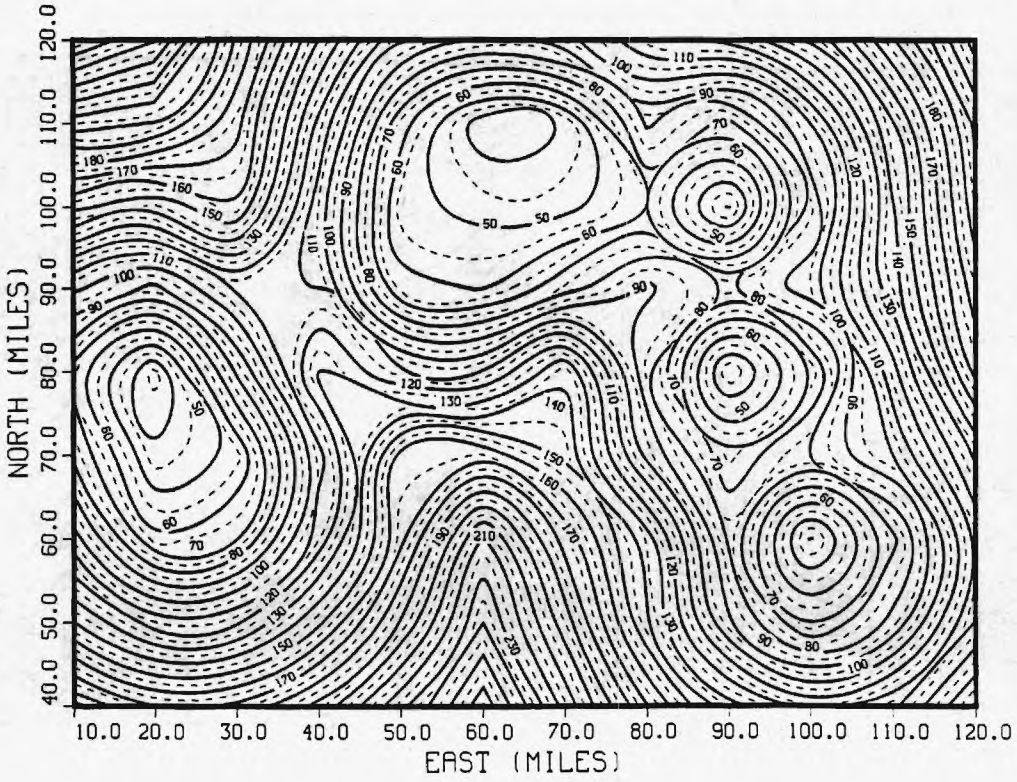


Figure 2.3.1.6 Estimation Variance at 14 Months (Feb. 1985).

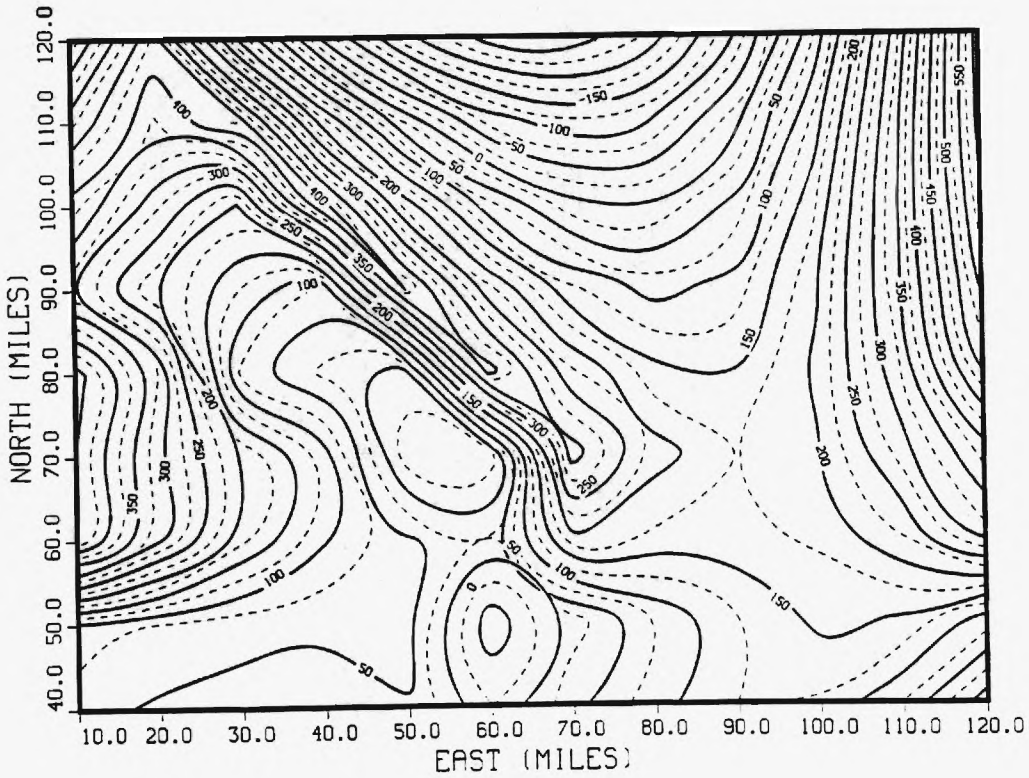


Figure 2.3.1.7 Spatial Kriging of Piezometric Surface at 1 Month (Jan. 1984).

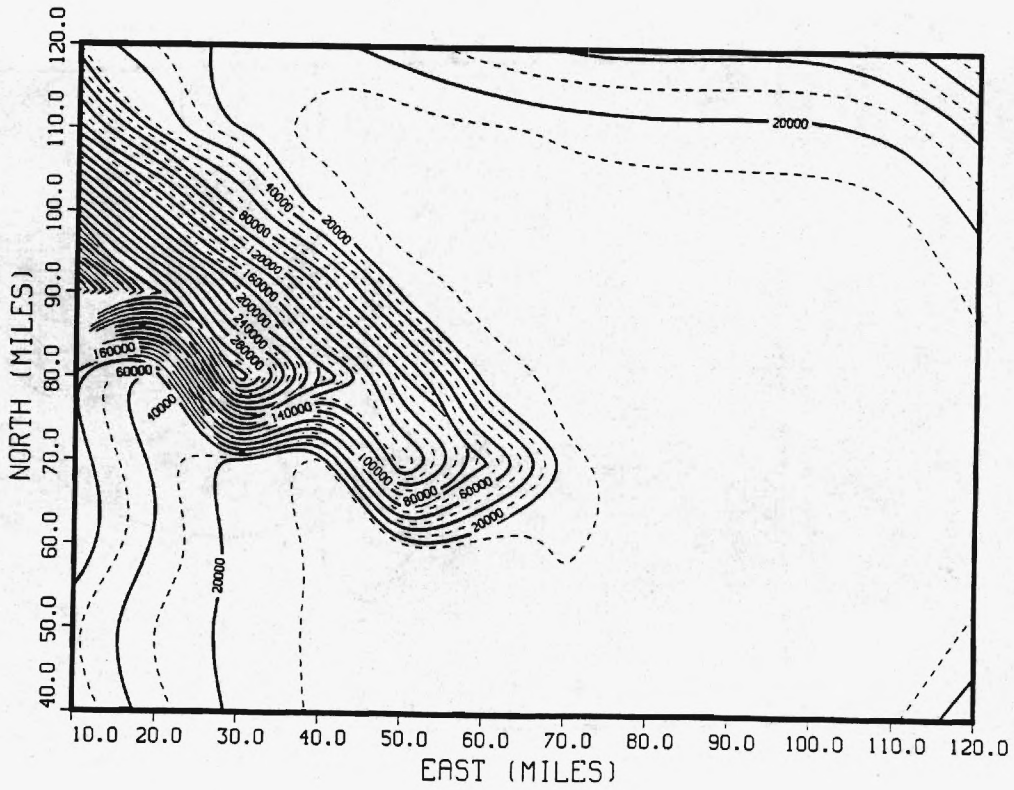


Figure 2.3.1.8 Estimation Variance Using Spatial Kriging at 1 Month (Jan. 1984).

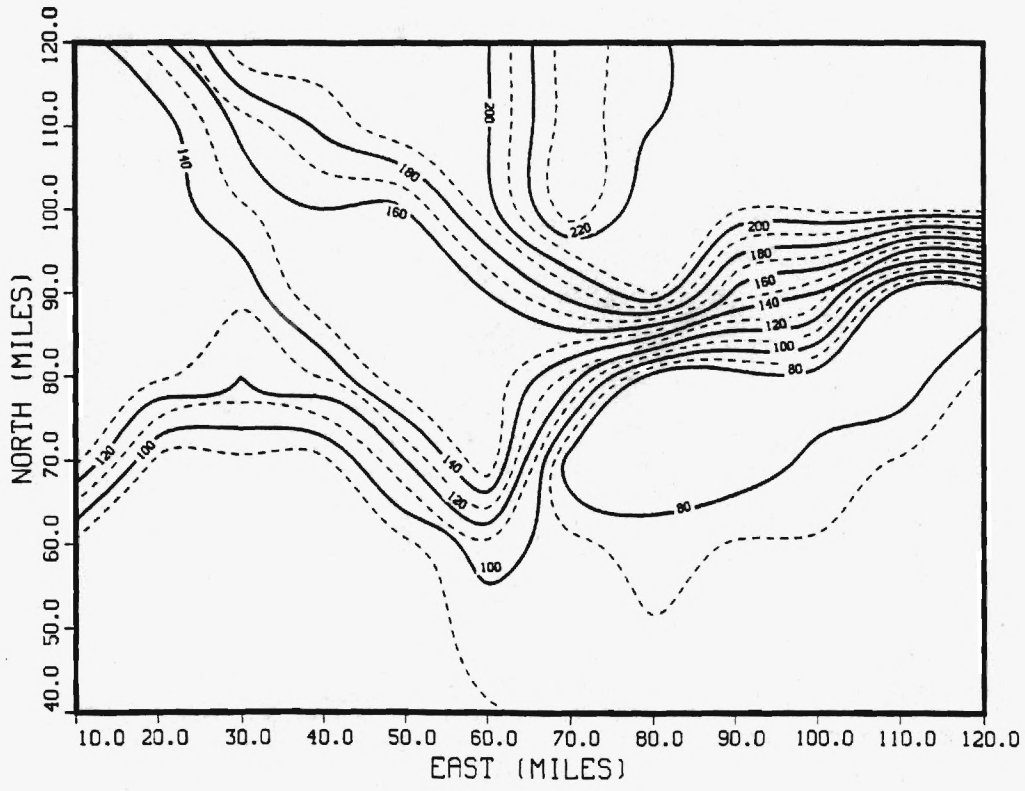


Figure 2.3.1.9 Space-Time Kriging of Piezometric Surface at 1 Month (Jan. 1984).

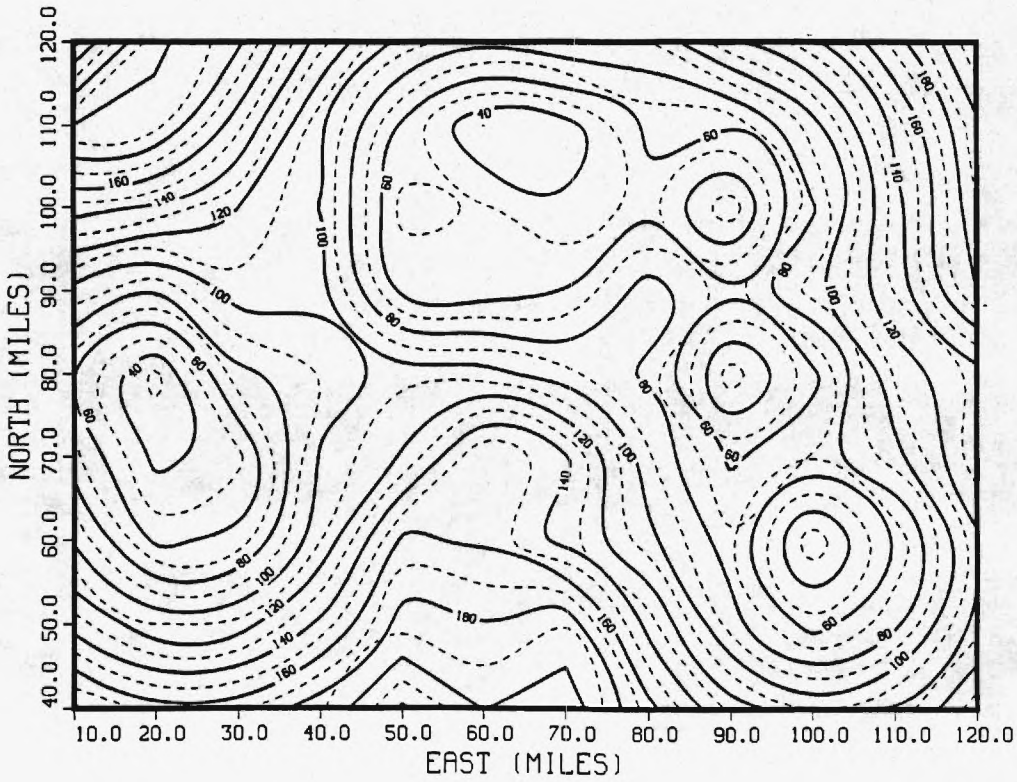


Figure 2.3.1.10 Estimation Variance Using Space-Time Kriging at 1 Month (Jan. 1984).

more accurate mapping of spatial data.

3. EXPANSION OF VARIANCE REDUCTION ANALYSIS

As noted, the variance of estimation, given by (2.4) is a measure for the accuracy of the estiamted value. Many authors have proposed the use of this measure as a guideline for sampling activities. For example, DeMarsily (1979) suggests that the location with the highest estimation variance should be selected as the next sampling point. This approach, however, does not consider the impact of a new data point on its neighboring region. It also ignores the fact that another point with lower estimation variance may be more effective in reducing the over-all uncertainty of the field.

In order to resolve the above problem, Rouhani (1983 and 1985) proposes the derivation of the magnitude of variance reduction at point X_o due to a sampling at X_* (the arbitrary location of a potential sampling site). This magnitude is denoted as VR_{o*}, which was determined through the concept of bordered matrices. It was also shown that VR_{o*} can be calculated without resolving the kriging sysytem. This allows us to evaluate the variance reduction potential of any point prior to its sampling.

VR_{o*} is determined to be:

$$VR_{o*} = (V_*(N))^{-1}\left(K_{*o} - \sum_{i=1}^{N} \lambda_{i*} K_{io} - \sum_{p=1}^{l(k)} \mu_{p*} f_p(X_o)\right)^2 \qquad (3.1)$$

where $V_*(N)$ is defined as the estiamtion variance at X_* prior to any sampling (i.e. using the N available data points). The rest of definitions are given in (2.5) and (2.17), assuming the condition of pre-sampling at X_*. This measure can then be expanded to the whole field to determine the total variance reduction due to a sampling at X_*, denoted as

TVR_*. In the continious case it is:

$$TVR_* = \int_S VR_{O*} dX_O \qquad (3.2)$$

or in the discrete case:

$$TVR_* = \sum_j VR_{O*} \qquad (3.3)$$

where S defines the continious estimated field, and j defines the set of discrete estimated point, such that $j \varepsilon S$.

TVR_* is a measure for information gain due to a sampling at X_*. It can be directly used by the planners to identify the location of the best sampling point. If a loss function is available we can calculate the economic gain due a sampling. For instance, Rouhani (1985) uses a two-piece linear loss function due to over- or under-estimation of the variable interest (in this case, piezometric levels). It is shown that the expected value of economic loss reduction is related to variance reductions as follows:

$$TLR_* = (C_U + C_O)(2\pi)^{-1/2}\left(\sum_j V_j^{1/2} - \sum_j (V_j - VR_{j*})^{1/2}\right) \qquad (3.4)$$

where,

TLR_* = total loss reduction due to a sampling at X_*;
C_U = loss per unit length of under-estimation;
C_O = loss per unit length of over-estimation; and
j = set of estimated points.

The above can be easily derived for the continious case.

TLR_* can be compared to the cost of sampling at X_*, in order to derive a value for the net worth of data. This

measure provides an alternative ranking for the selection of the best sequence of points. The identification of the best sequence is conducted, such that, at each round of kriging the point with the highest information gain ($TVR_{\star}$) or the highest economic gain ($TLR_{\star}$) is selected as the next added sampling point, which will then be added to our data set. This process continues until we have satisfied our infromation criteria, or our budget is exhausted. This yields a sequence of n points among m available points for further sampling. For a more detailed study of variance reduction analysis readers are referred to Rouhani (1983 and 1985).

3.1 Variance Reduction Analysis for Non-negative Variables

As noted earlier, the above procedure had been applied to cases of Gaussian random variables, only. In our first attempt we utilize this procedure for sampling in cases where the variable of interest is non-negative. The basis of this procedure is already outlined in Section 2.1. In other words, we assume that our variable of interest is log-normally distributed. We further propose to use the lograrithm of Y, denoted as Z and defined by (2.8), as the basis of our variance reduction analysis, which is outlined in Section 3.

To accomplish the above, it needs to be shown that the point with the highest information or economic gain for Z(X) is the same point for the original variable Y(X). This can be easily illustrated by considering (2.9a) and (2.9b) that give the expected value and the variance of an estimated Y(X), in terms of the expected value and estiamtion variance of the estimated Z at the same point. As seen, the variance of Y(X) directly depends on the Var(Z(X)) and E(Z(X)). However, E(Z(X)) remains unchanged, as long as, the added point agrees closely with its predicted value. Thus, the only variable is Var(Z(X)), which goes down as new data points are added. This means that whichever point that induces the highest varaince reduction for Z(X), yields the same for Y(X).

The magnitude of variance reduction for the estiamted $Y(X_o)$ due to a measurement at X_*, $VR_{o*}(Y)$, can be derived from the variance reduction at $Z(X_o)$ due to the same measurement, denoted as $VR_{o*}(Z)$, by taking the first derivitive of (2.9b) with respect to $Var(Z(X_o))$, and rewriting it to yield:

$$VR_{O*}(Y) =$$

$$VR_{O*}(Z)[\exp\{2E(Z(X_O))+Var(Z(X_O))\}][2\exp\{Var(Z(X_O))\}-1] \quad (3.5)$$

which further shows that the selected sequence of points based on variance reduction analysis of Z, produce the highest variance reductions for Y.

3.1.1. Case Study: Optimal Schemes for Groundwater Quality Monitoring in the Shallow Aquifer of Dougherty Plain, Southwestern Georgia

As described in Section 2.1.1, Dougherty Plain is a major agricultural center, whose growth is made possible by the groundwater from the principal artesian aquifer. The protection of this aquifer can be substantially improved by establishment of a water quality network in the shallow aquifer. This monitoring network can act as an early warning system for pollution control in the lower layers. It also allows time for the design and implementation of appropriate prevention plans.

The questions that immediately arise are: How should we design such monitoring netwrok? More specifically, what criteria should be utilized as the basis of our network design? Where are the best locations for sampling sites? In order to answer these questions we studied a number of schemes.

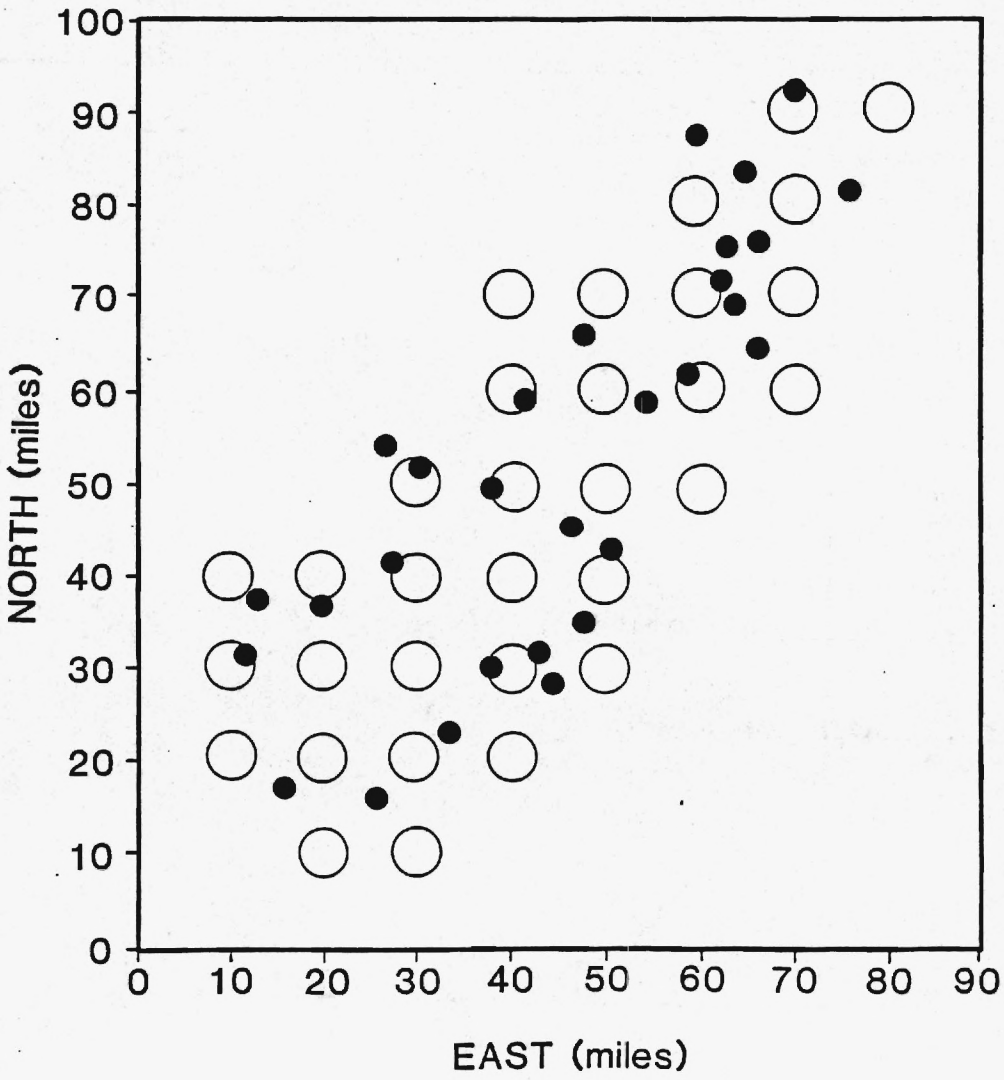
The common statitical approach to a sampling design is the maximization of incremental infromation subject to budget constraints; see Fiering (1965), Hughes and Lettenmaier (1981), Chou and Scheck (1984), and Rouhani and Fiering (1986). Variance reduction analysis is one such method.

These schemes generally give priority to points with high estiamtion variances, regardless of the magnitude of the estiamted values. Such a criterion is thus suitable for cases where the magnitude of the variable of concern is not of primary importance.

In our case study, however, the desired monitoring network is designed on the basis of leakance data. This choice is made in order to identify locations, where there are higher chances of surface pollution leakage into the principal artesian aquifer. So, we are not only interested to gain as much information as possible, but also to monitor areas with potentially high levels of recharge. This means that we should also explore other selection criteria which include both the accuracy and the magnitude of the estiamted values.

To accomplish the above, we have used three selection criteria. The first one is based on the maximization of incremental information, using variance reduction analysis. The second one is based on the ranking of median values of estiamted leakance. The third one uses the risk value as the basis of its selections, as defined by (2.11). This last criterion includes both the accuracy and the magnitude of the varaible of interest.

For sampling purposes we have defined the 32 points shown in Figure 3.1.1.1 as potential sampling sites. These points are scattered uniformly over the Dougherty Plain area. This figure also displays the location of existing sampling sites which are scattered throughout the plain and its vicinity. There are concentrations of data points in few zones, such as the middle of the upper portion. However, the distribution of these points can be considered as relatively uniform. Such a distribution allows us to examine the ranking of potential sampling sites without any implicit



3.1.1.1 Potential and Existing Sampling Points, Dougherty Plain (Potential sites are given as blank circles).

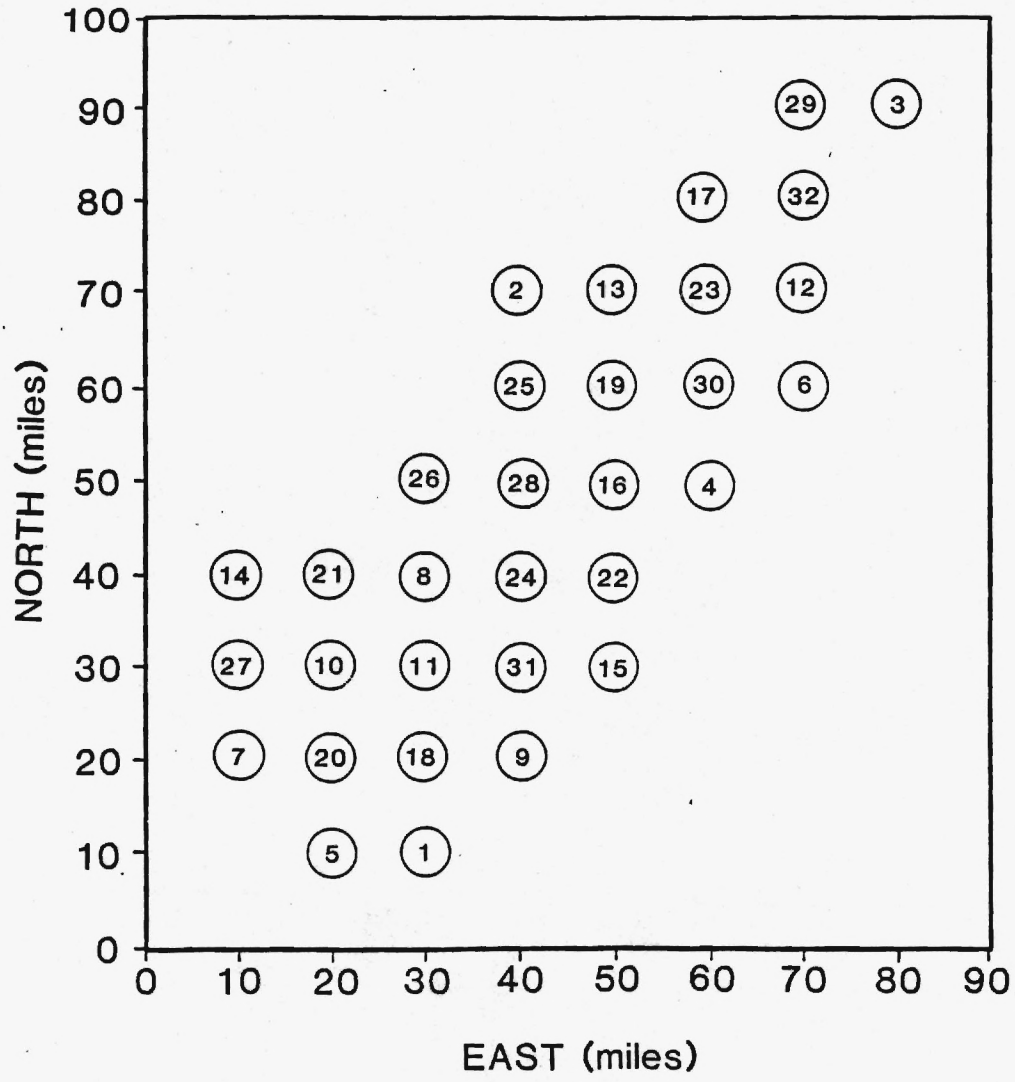


Figure 3.1.1.2 Variance Reduction Ranking for Leakance.

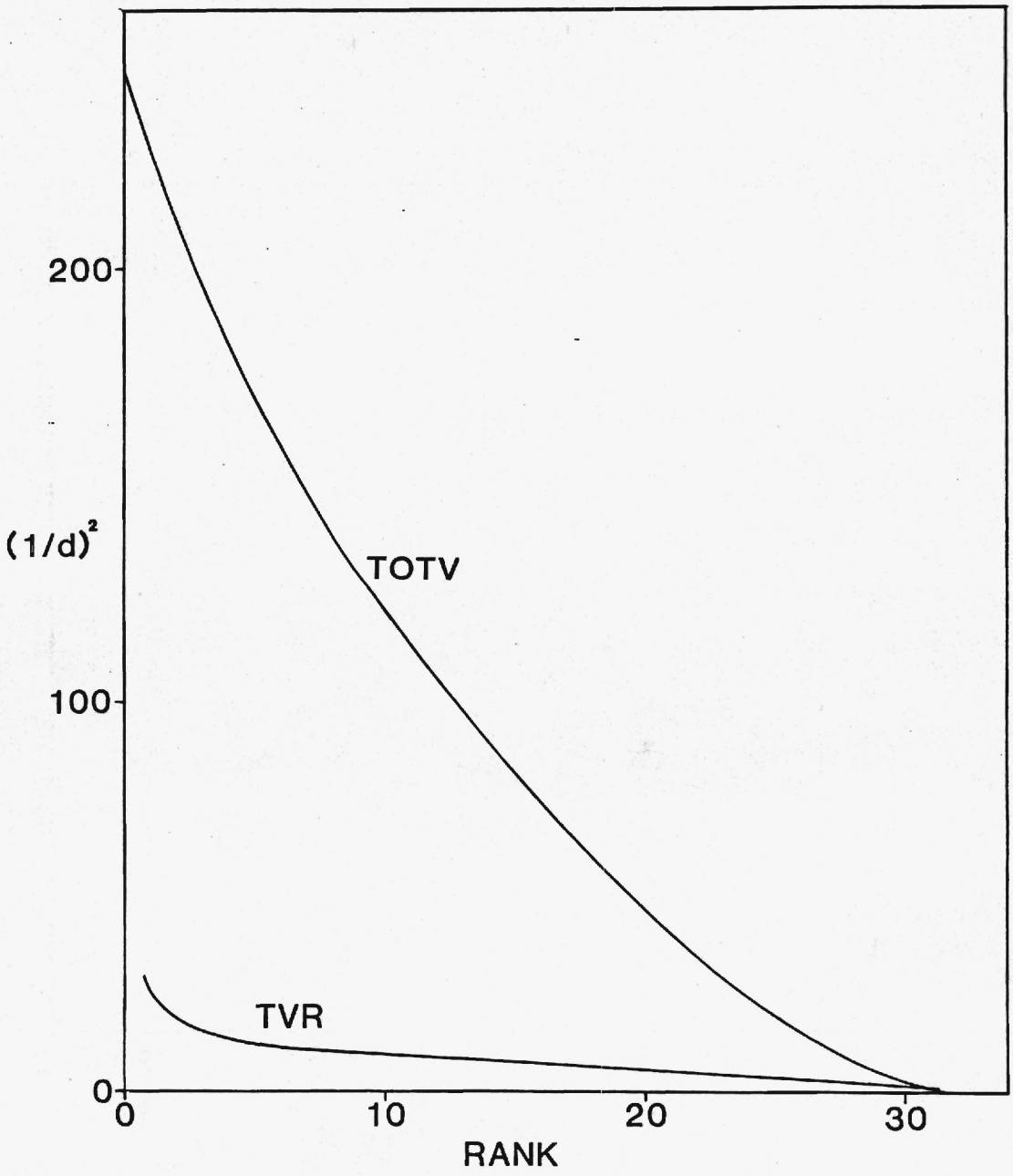


Figure 3.1.1.3 Total Variance Reductions of Added Sites.

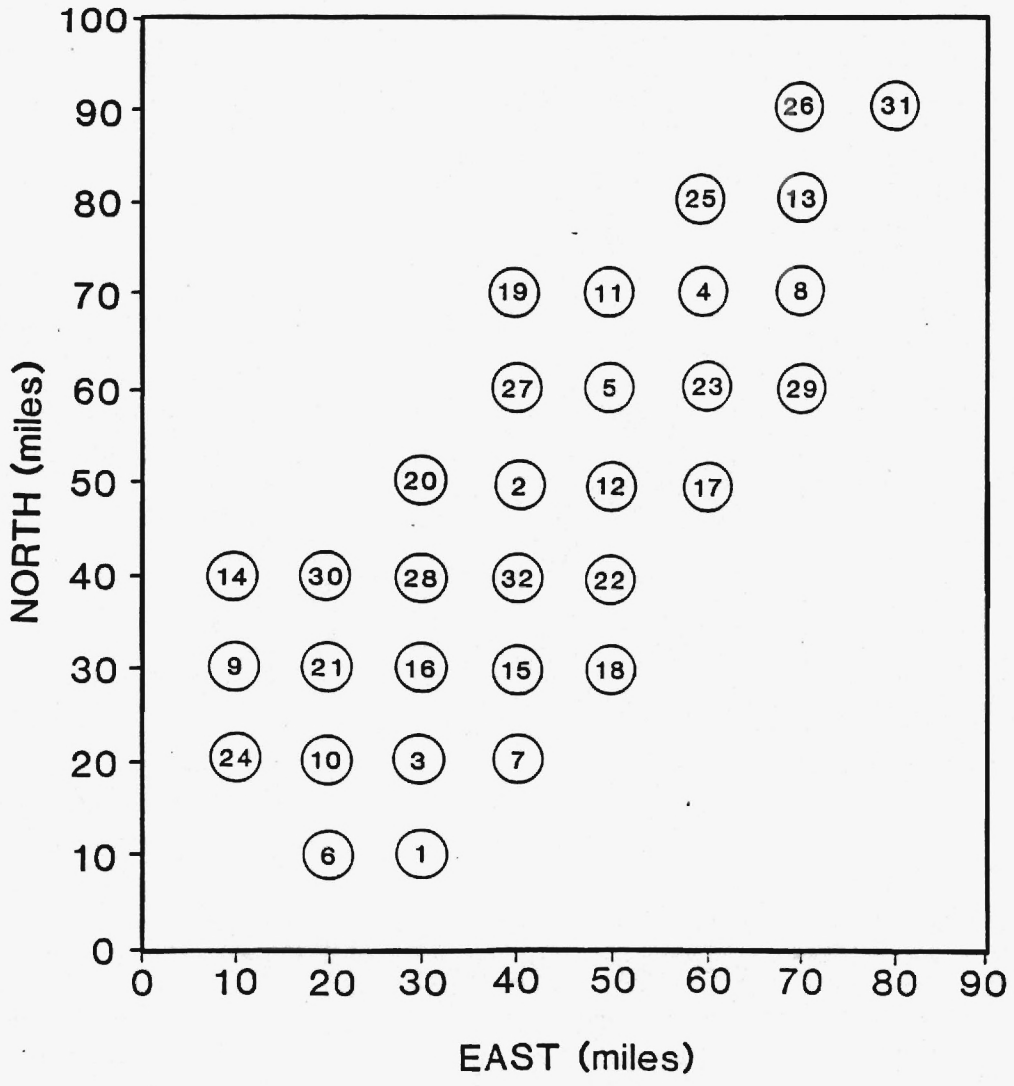


Figure 3.1.1.4 Median Ranking for Leakance.

bias.

The sampling activities are conducted in a non-sequential manner , which means that the estiamted covariance function remains unchanged. This assumption can be violated if the measured values at the new sampling sites turn out to be significantly different from their predicted values. In a non-sequential sampling we implicitly assume that the measured value at the new site belongs to the same predicted population. The rankings are conducted similar to the procedure outlined in Section 3. We, however, use three selection criteria, as discussed above. For the sake of brevity we only present a summary of results. For a complete description of results readers are referred to Rouhani and Hall (1987).

Figure 3.1.1.2 shows the sequence of selected points based on variance reduction analysis. As expected, the boundary nodes located in the eastern and southern sections of the palin have higher ranks. The centeral nodes and the western boundary, on the other hand, have lower ranks. Figure 3.1.1.3 displays the relative information gain by each ranked site, in terms of its TVR_*. Sampling at the top five points yield the highest amount of gain. Additional sampling appear to worth only marginally. One could assume that there must be a finite number of sampling sites, such that, any sampling beyond these points results in small information gains that cannot be economically justified.

The above criterion, despite its versitality, ignores the magnitude of estiamted values. So, in our second ranking we use a criterion which only depends on the estiamted magnitudes. Figure 3.1.1.4 shows the result of median ranking. As seen, in contarst to the previous ranking, the upper centeral portion of the plain, as well as, its southern tip have gained the highest ranks.

In our third criteria, we used risk values as the basis of our selections. Figure 3.1.1.5 shows such ranking, using the 10% risk values (i.e., such values whose probability of exceedence is only 10%). This corresponds to (2.11), where z_α is equal to 1.28. In this ranking, the southern tip and the centeral zone (nodes ranked fourth, fifth, and sixth) have the highest rankings. The centeral region was generally ignored by the first criterion. This is due to fact that this region despite of its high leakance is relatively well sampled, and thus, has low estiamtion variance. This illustrates one of the advantages of the risk ranking which, in addition to the accuracy of estiamted points, considers their magnitude, as well.

Equation (2.11) shows that the risk value is basically a weighted sum of the expected value and its estimation variance. So, as we decrease the probability of exceedence of our risk values (i.e., making them more extreme), we are giving more weights to the variances, and conversely. This indicates that the risk value has practically two extremes. If α is very small, the risk ranking appraoches the variance reduction ranking. On the other hand, as α nears 50%, or as z_α approaches 0, the risk ranking becomes closer to the median ranking. These tendencies are cleraly demonstrated in Figures 3.1.1.6 and 3.1.1.7, which represent the 1% and 4% risk rankings, respectively. This characteristic is another advantage of risk ranking that provides a flexible weight for users to adjust their sampling plans according to their specific objectives.

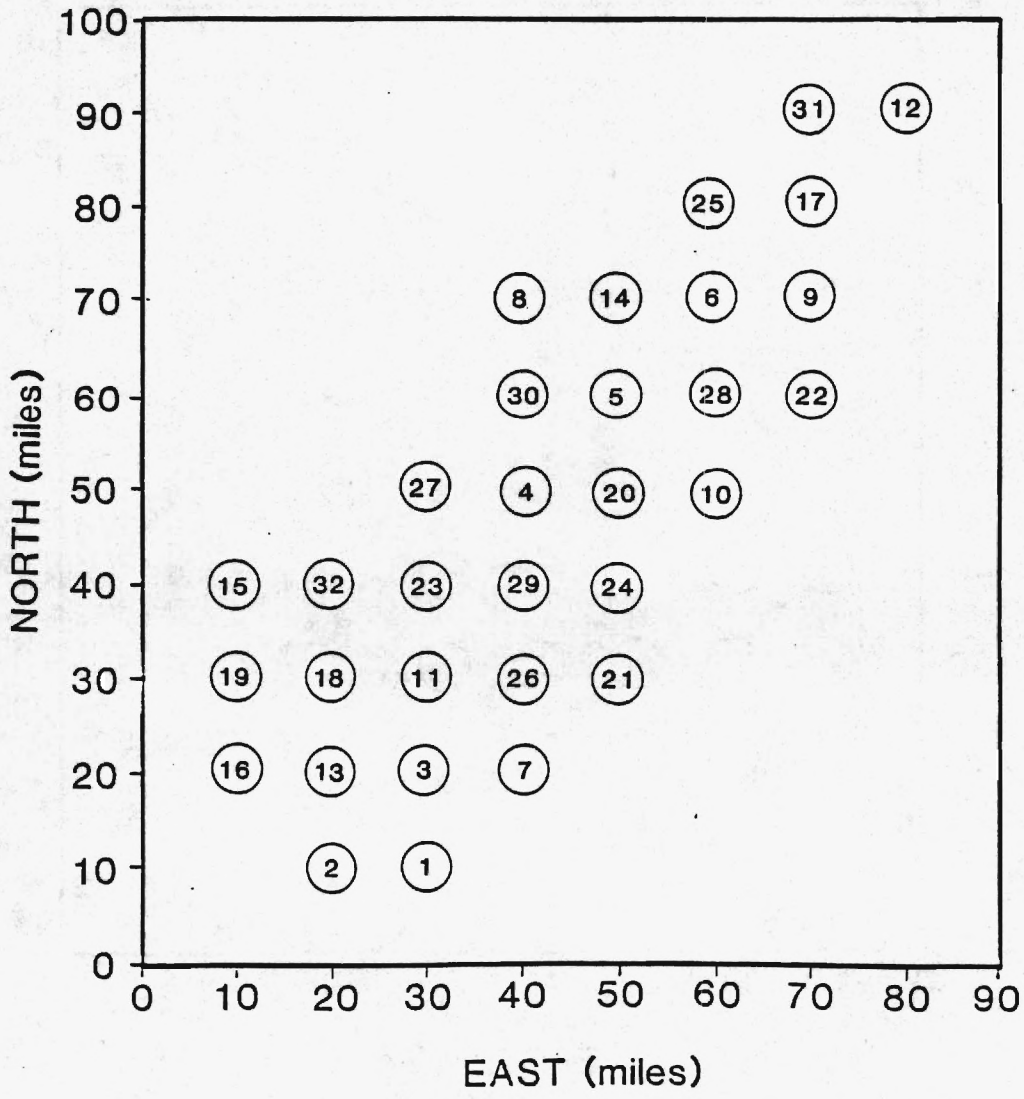


Figure 3.1.1.5 10% Risk Ranking for Leakance.

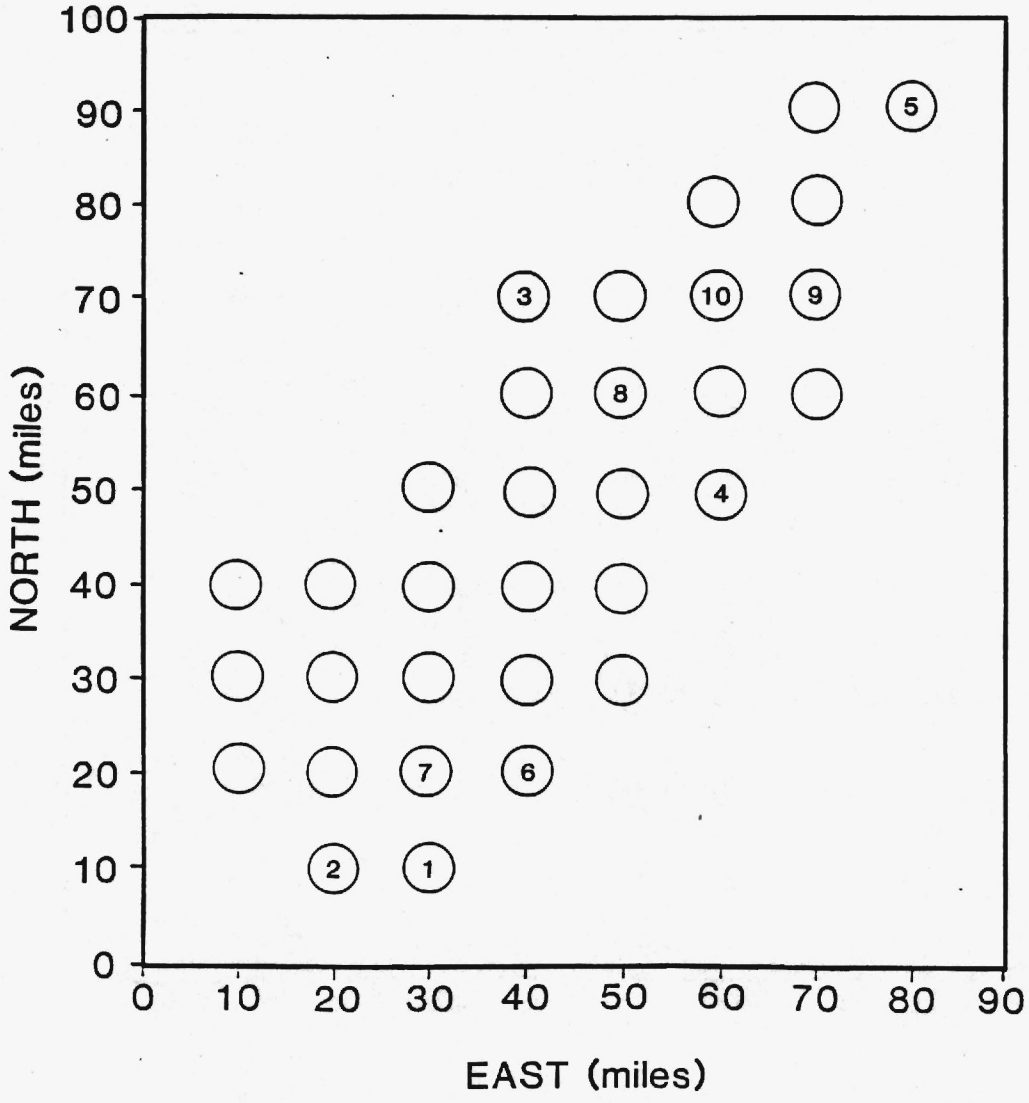


Figure 3.1.1.6 1% Risk Ranking for Leakance.

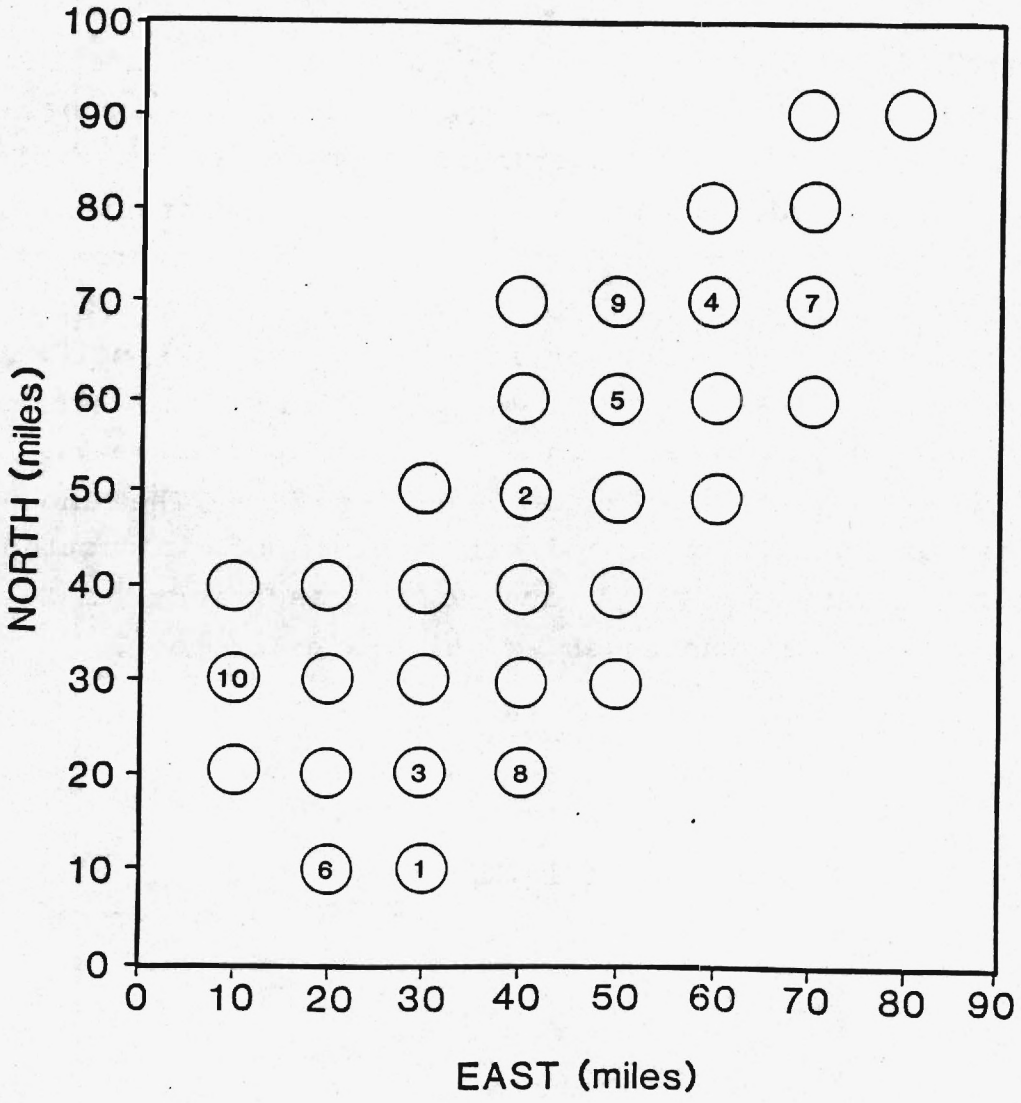


Figure 3.1.1.7 40% Risk Ranking for Leakance.

3.2 SPACE-TIME VARIANCE REDUCTION ANALYSIS

In this section we attempt to expand variance reduction analysis into space time domain. Our survey indicates that this study is the first attempt to expand variance reduction analysis to be applicable to the case of spatiotemporal variables.

The basis of our approach is space-time kriging, as described in Section 2.3. In particular, it can be shown that the addition of a measurement point at an arbitrary site, X_*, with coordinates (x_*, y_*, t_*), transforms kriging matrix in (2.17) into a bordered matrix. The inverse of this matrix, and hence, the solution to this kriging system, can be determined by utiizing the inverse of the original matrix, see Rouhani (1985). Thus, we can derive the estimation varaince at an estimated point, X_o, with coordinates (x_o, y_o, t_o), if a sampling point is added at X_*. The amount of information gain, measured in terms of reduction in estimation variance at X_o, due to a measurement at X_*, can then be calculated, using an equation similar to (3.1), as:

$$VR_{o*} = (V_*(N))^{-1}\left[K_{*o} - \sum_{i=1}^{N}\lambda_{i*}K_{io} - \sum_{p=1}^{l(k_s)}\mu_{p*}f_p(x_o,y_o) - \sum_{p=l(k_s)+1}^{l(k_s)+l'(k_t)}\mu_{p*}f_p(t_o)\right]^2 \qquad (3.6)$$

where l and l' are the numbers of monomials in the spatial and temporal drifts, which depend on the order of these drifts, k_s and k_t, respectively. The rest of definitions are indicated in (3.1) and (2.17). For the case of a variable with a two dimensional space coordinates and a one dimesional time coordinate, the number of monomials are given as

follows:

$$l = (k_s + 1)(k_s + 2)/2$$

and

$$l' = k_t \qquad (3.7)$$

Please note that the constant monomial is not included in the temporal drift. This is due to the fact that it is already contained in the spatial drift. It implies that the universality condition for the constant monomial is included: $\Sigma\ \lambda_{io}=1$, for i=1,...,N. As an example, we can look at the case of $k_s=k_t=2$, where l=6 and l'=2. The monomials for the above case are: 1, x_o, y_o, $x_o y_o$, x_o^2, and y_o^2 (for space), and two additional temporal monomials t_o and t_o^2; a total of 8 monomials.

In the case that X_o and X_* have two different neighboring set of data, we have to make an additional assumption. For this purpose, we use the argument stated by Rouhani (1985). He states that since we would like to predict the impact of X_* on X_o, we should use the former's neighboring data as the basis in (3.6). This allows us to measure the impact of additional information which is presently contained in the estimate at X_*.

The neighborhood of each estimated point is determined by using the space-time covariance function, as defined in Section 2.3. To accomplish this task, the number of neighboring points, N, has to be specified. The program identifies the N data point with the highest covariances with the estimated points, as its neighboring sites. This is a specially efficeint scheme for spatiotemporal variables.

Similar to the space variance reduction analysis, we can

expand (3.6) over the entier domain, in order to measure the information efficiency of an added sampling at X_*, in terms of its total variance reduction (TVR_*):

$$TVR_* = \int_{S,T} VR_{O*} dX_O$$

or in the dicrete case,

$$TVR_* = \sum_j VR_{j*} \qquad (3.8)$$

where, X_O sweeps the spatiotemporal domain, and j is the set of estimated points in space and time. At each round of sampling we can identify the point with the highest TVR as our next added measurement. This process continues until we have either exhausted our budget, or have satisfied an accuracy criterion, such as maximum allowable uncertainty.

For sampling, the program allowes a general flexible schemes, suitable for a variety of different hydrological problems. The user can select n sites at each time period as one time measurements or measurements that will be collected for the next m time intervals. For instance, an oceanographic vessel allows only a one time sampling at each location along it path. In contrast, a stationary device, such as a piezometer inside a well, provides discrete or continious measurements at only one lcation. Thus, the above scheme can be easily adapted to a variety of monitoring devices.

3.2.1. Groundwater Sampling in Space and Time in Southern Georgia

In this project we initiated the applications of space-time variance reduction by using it for the design of a sapmling scheme for groundwater monitoring in southern Georgia. The data set is already discussed in Section 2.3.1. It consists of average monthly piezometric heads in 8 wells

for a 12 month duration.

For the purpose of sampling the centeral region of this area, where 10 miles $\leq x \leq$ 100 miles, and 40 miles $\leq y \leq$ 100 miles, is selected as the area of potential new measurements. We then establish a 6 x 4 grid, with 15 mile increments in both directions. The nodes of this grid are defined as the potential sampling sites, as shown in Figure 3.2.1.1.

Our measurement devices are assumed to be stationary piezometers, Which are to be installed one month after the end of our available measurements, i.e., 13 months. This task would be impossible, if we only had the spatial variance reduction analysis.

We then used a neighborhood size of 10. The result for five additional sampling sites are given in Figure 3.2.1.2. It is interesting to note that the estimation variance of the first point is less than the one for the second point. This condition also exist for the fourth and the fifth point. It indicates that at the first and the fourth round of ranking, despite the fact there are points with higher varainces, the program selects other sites that have more effective impacts on the accuracy of the entier field. Such a procedure results into higher information gains.

Figure 3.2.1.3 reflects the level of information gain at each sampling. As expected this level drops quickly to an assymptotic level after the first few added points.

In our next trial, we study the sensitivity of space-time variance reduction analysis with respect to the size of the neighborhood, N, which is a rather arbitrary measure. For this purpose we used an N=6. The results of analysis are given in Figure 3.2.1.4. Which are almost identical to the previous case. The only diference is in the

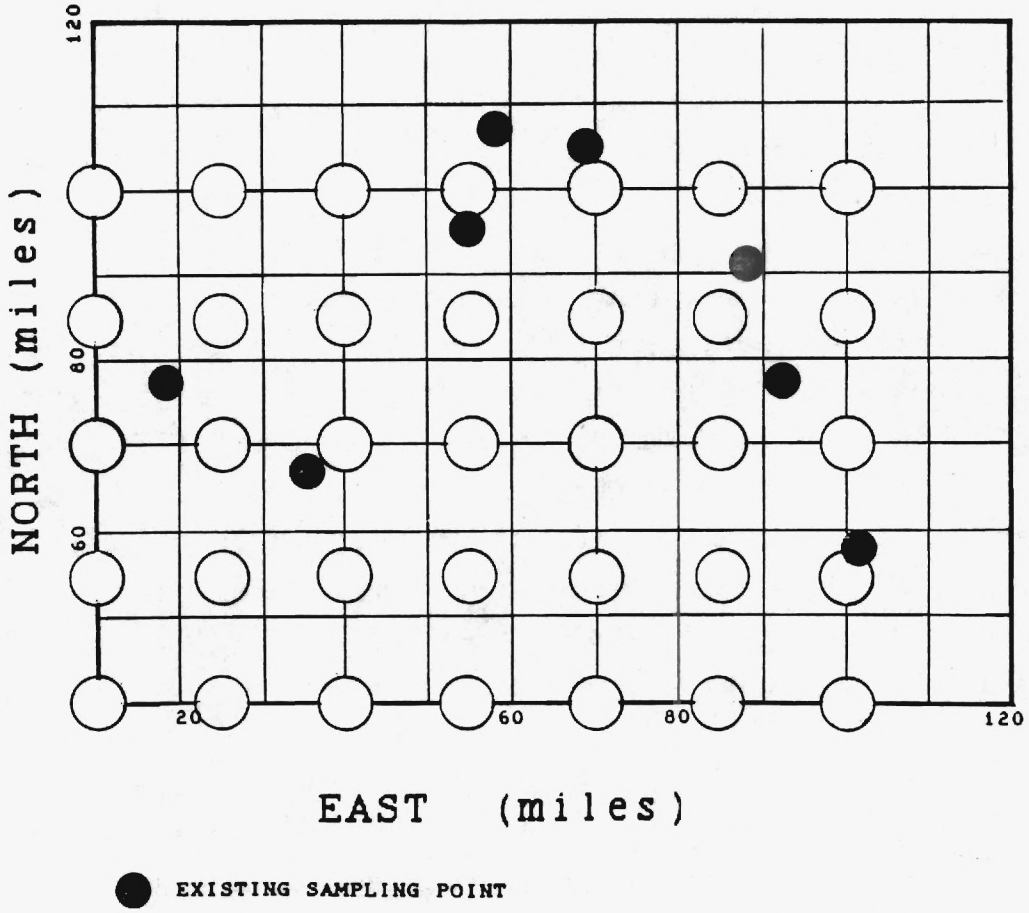


Figure 3.2.1.1 Existing and Potential Sampling Sites, Southern Georgia (Potential sites are given as blank circles).

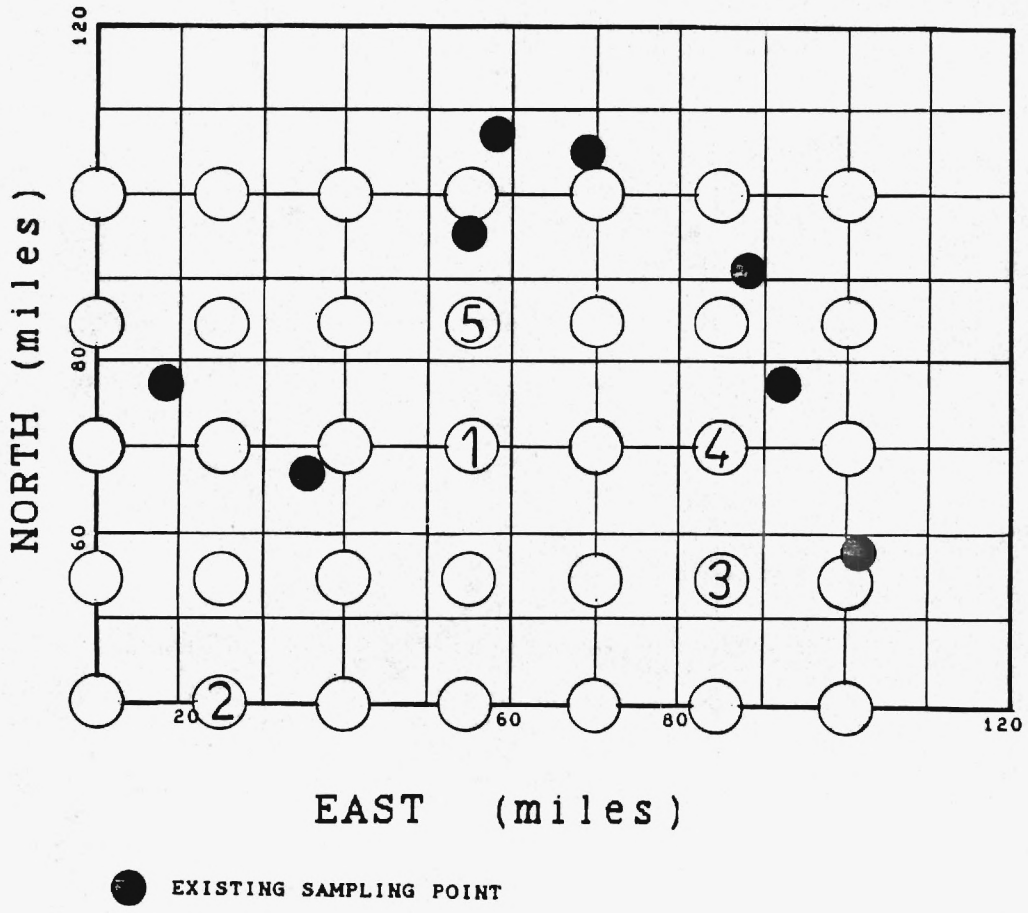


Figure 3.2.1.2 Variance Reduction Ranking (N=10).

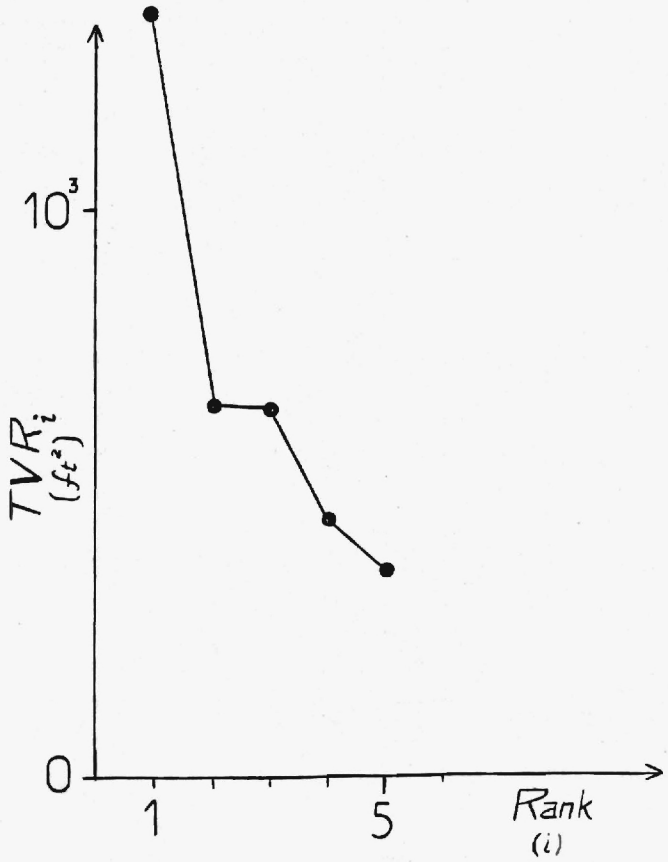


Figure 3.2.1.3 Total Variance Reductions of Added Sites. (N=10).

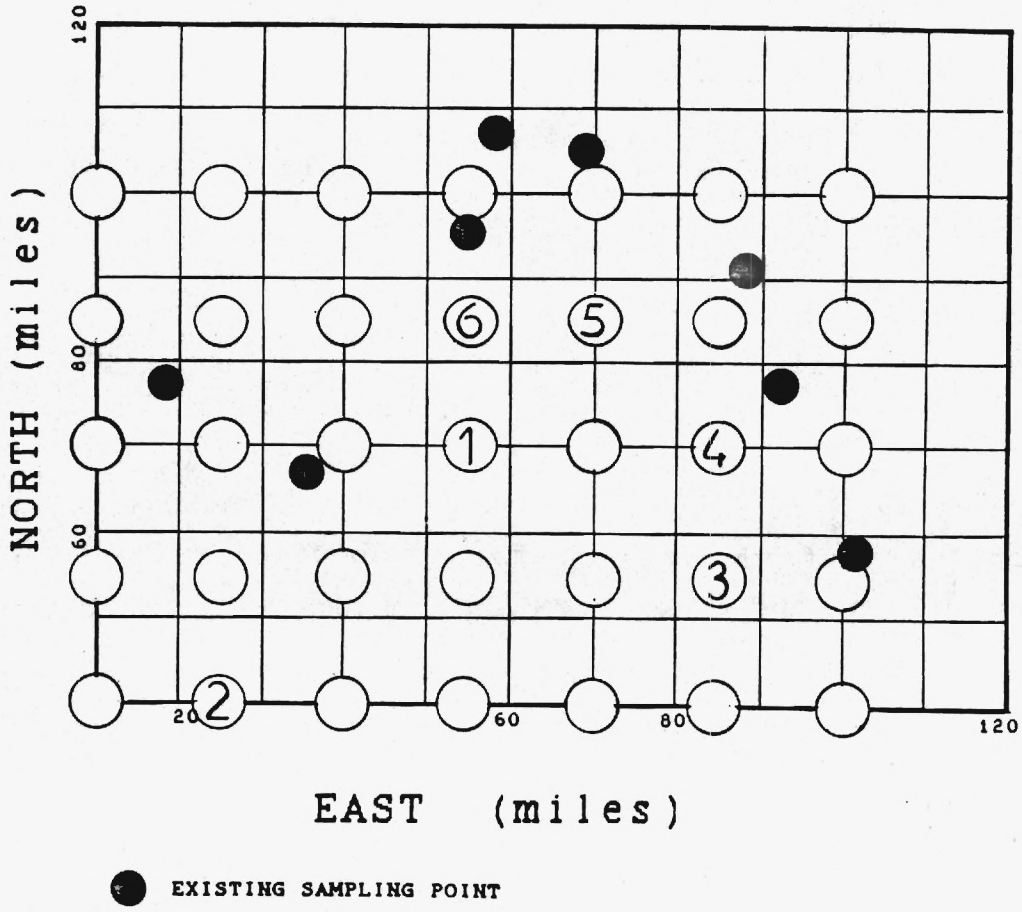


Figure 3.2.1.4 Variance Reduction Ranking (N=6).

choice of the fifth point, which has a low information gain, anyway. Figure 3.2.1.5 displays the estiamted information efficiencies of the selecd point that are also very similar to the previous case. Thus, it can be concluded that variance reductiona analysis is relatively robust with repect to the size of neighborhood. This is true, as long as, the first few nearby data points adequately describe the process in the vicinity of the estimated point.

The major problem that we encountered here was the problem of ill-conditioned kriging matrices, which occured more frequently as we added more sampling points. This problem manifests itself by yielding unrealistic λ_{io}, which result into varainces of estimation with very high, and sometimes, unbounded absolute values. This in turn prohibits the use of varaiance reduction analysis, for TVR may show an unacceptable upward trends. This condition in fact occured for the sixth points and on, we thus ignored their results. Presently, we are studying alternative approaches to resolve this problem. For instance, in such cases we can use approximate solutions to the kriging system in (2.17). This, however, may reduce the accuracy of our scheme. It is anticipated that the search for a solution for the above problem will be one of our next objectives.

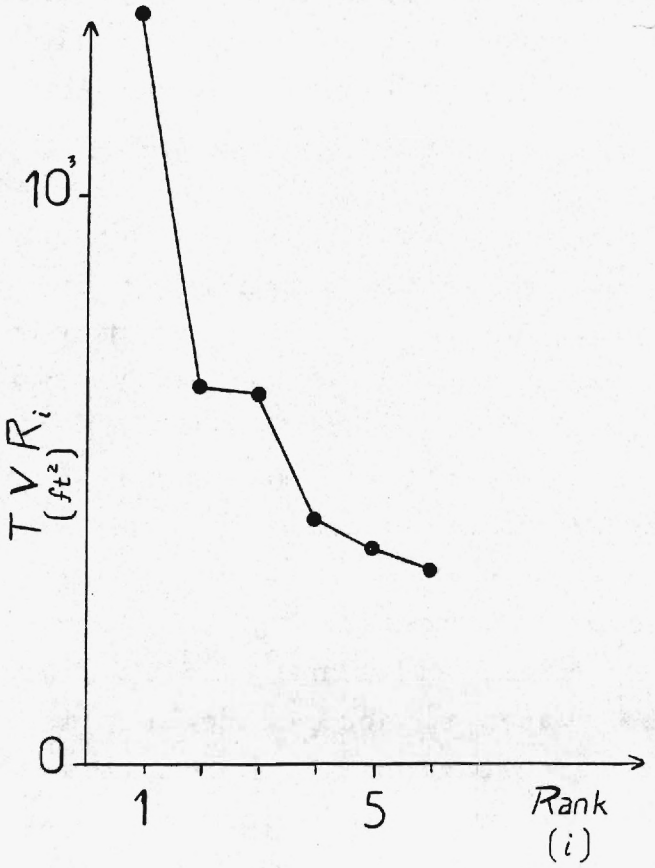


Figure 3.2.1.5 Total Variance Reductions of Added Sites (N=6).

4. RESILIENCE OF VARIANCE REDUCTION ANALYSIS

As noted in Section 1, one of our goals has been to test the resilience of variance reduction analysis, in cases where some of its basic assumptions are violated. In previous sections we have already study the sensitivity of resilience vis-à-vis the variability in some of the involved parameters, such the size of the neighborhood. In this section, however, we are more interested to study the dynamic behavior of variance reduction analysis. The following paragraphs provide a brief report of this study. For a more detailed presentation, readers are referred to Rouhani and Fiering (1986).

Our study is designed to answer such questions, as: What is the effect on the sampling scheme if the predicted values are significantly under- or over-estimated? How does the covariance function respond to the newly sampled values? Are the decisions of the above analysis stable under such situations? To answer these questions we propose to study the resilience of variance reduction analysis.

The concept of resilience in water resources is a relatively new topic (Fiering, 1982). Resilience is the ability of the system to accomodate surprises and to survive unanticipated perturbations. It implies that even if an unlikely event occurs, the decision has an acceptably high probability of being either correct or good enough. In other words, a tolerance ("good enough") and a confidence ("acceptably high") are required.

Resilience is a more general concept than the robustness. Fiering (1982) gives an example to illustrate the differences between robustness and resilience of a system: "The sensitivity of the system response with respect

to a decision variable x_i is given by the partial derivative $\partial f/\partial x_i$. If the partial derivative is small, the system is "robust" with repect to such changes. If the partial derivative is not small, the system need not suffer important shifts in its reponse because changes in other decision variables might be made to accommodate an unfortunate choice of x_i." Therefore, robustness alone does refect the behavior of the entier system. The total derivative $df/dx_i = \Sigma_j(\partial f/\partial x_j)(\partial x_j/\partial x_i)$ measures the system's ability to adjust to changes in x_i. A linear combination of all total derivatives df/dx_i might suggest a measure of resilience of the given system.

In this section we consider variance reduction analysis as a system, composed of an input space (set of measured values), a parameter space (covariance models and their estimated parameters), and an action space (selected sampling sites).

To test the resilience of the above system we first assume that the measured values in the data space are under- or over-estimated to such degree that we have to reject the hypothesis that they belong to the population with a mean equal to their predicetd values. In the second step the parameters of the covariance are re-estimated, affecting the parameter space. Finally, we proceed with the selection of next added sites, thus, studying the impact of mis-estimation on the action space.

To generate the "measured" values at the sampling site , we use the risk values of the variable:

$$z^{i+1}(X_*) = \hat{z}^i(X_*) \pm z_\alpha(v^i(X_*))^{1/2} \qquad (4.1)$$

where,

$Z^{i+1}(X_*)$ = $(i+1)^{th}$ added measured value at X_*;
$\hat{Z}^i(X_*)$ = estimated value at X_* based on K^i;
$V^i(X_*)$ = varaince at X_* based on K^i;
K^i = estimated covariance at the i^{th} round of sampling;
z_α = standardized normaly disributed random variable with a probability of exceedence of α percent; and
α = level of deviation.

We then define a number of schemes on the basis of type of mis-estimation (O for over-estimated, U for under-estimated, and S for alternating under- and over-estimation), and their level of deviations in percent. So, the U-90 refers to the case that all measured values are assumed to be under-estimated, using z_{90}=1.280 in (4.1).

The data set used are 84 piezometric data in northwstern Kansas, scatterd over an area equal to 2,560. square miles, as described in Rouhani (1985). The study area is divided into a 4 x 5 grid with increments of 8 and 16 miles in x and y directions, respectively. The nodes are defined as potential sampling sites, as indicated in Figure 4.1.

Nine studies of sampling planning are conducted, using different mis-estiamtion schemes and levels of deviations. Some of the generated data with large perturbations might be unrealistic. For example, in study U-99, large additions to Z might yield a water table significantly higher than the ground level. These values are included in this study to test the reliability of the proposed algorithm under some extreme, unexpected or counter-expected events (Fiering and Kindler, 1981).

The results of sampling studies are included in Figure

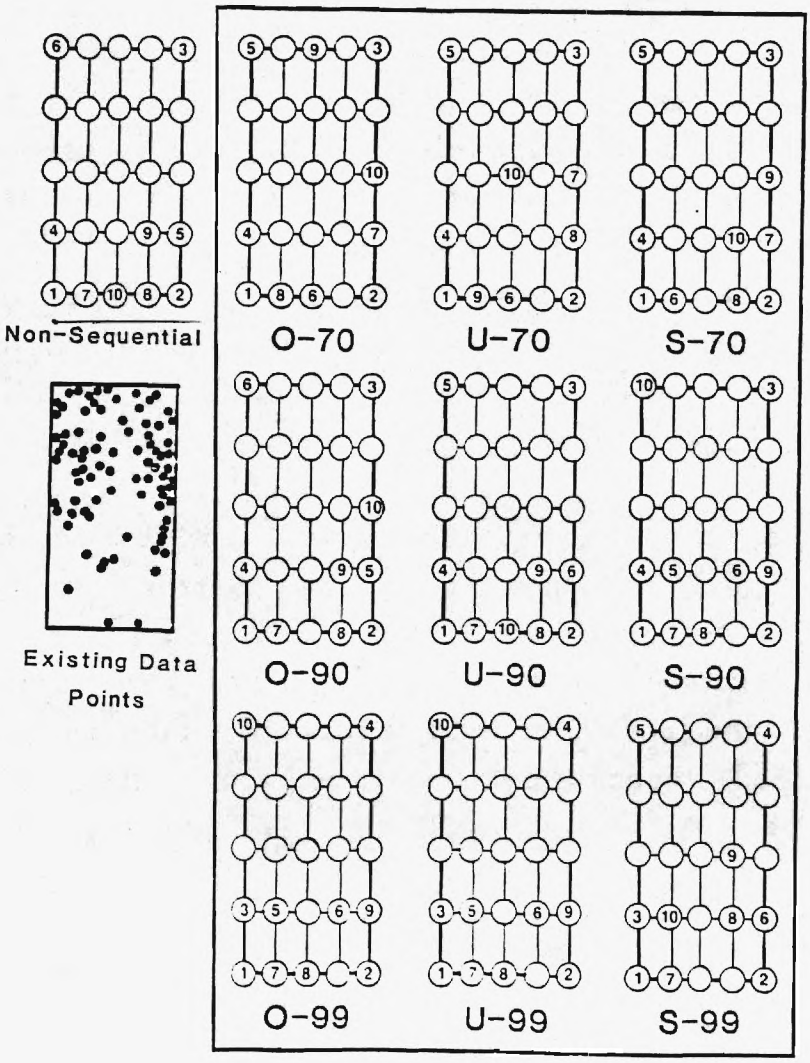


Figure 4.1 Original and Sequential Sampling Schemes along with the Existing Data Points (Northwest Kansas).

4.1. Comparison between these results with the original case with no generated noise, reveals a case of an unstable parameter space, but a resilient action space. In general, when the level of noise is low, universal kriging treats it primarily as measurement error. Cosequently, the structural analysis produces covariance functions with larger nugget effects (C in Eqaution (2.14)). In such instances the priorities are furhter shifted toward border nodes.

When the level of noise is high, universal kriging considers it as indication of error caused by an under-estimated covariance function. Consequently, the parameters of the structured part (coefficients a_{2p+1} in Equation (2.14)) increase. This in turn causes an increase in the influence of the internal nodes on their neighboring points, which makes the internal locations more advantageous as sampling sites.

Despite the large amount of simulated noise, all selected sequences show a great degree of geometrical similarity. A regret analysis also shows a case of near-optimality among all selected sets. This study clearly illustrated the resilience of variance redution analysis. A more detailed description is given in Rouhani and Fiering (1986).

5. GENERAL CONCLUSIONS AND FUTURE PLANS

As described, universal kriging and variance reduction analysis appear to be effective tools for performance of hydrological estiamtion and sampling. Universal kriging has proven to be easily applicable to cases of non-negative log-normally distributed random variables. Its theoretical expansion to the time-space domain is also conducted without any difficulty. This expansion in turn allows the development of variance reduction analysis for spatiotemporal processes. In the following paragraphs we try to highlight the important conclusions that lead us to our future plans.

In the first section, the theoretical part of the expansion of universal kriging appear to be rather straightforward. This is done by assuming that time is $(n+1)^{th}$ dimension for the variable of interest. Presently, we are exploring other theoretical arrangements to deal with space-time dimensions. For instance, we can study data at each location as a time series, characterized by a temporal covariance or variogram. Then, we consider these time series as a set of seperate correlated random variables. This will allow us to estimate them using co-kriging or principal componenet analysis. The PI is pursuing this objective in his present project: "Advanced geostatistical studies at Centre de Geostatitique, Ecole des Mines de Paris, Fontainebleau, France", funded by NSF (Project Number INT-8702264).

The next step is the selection of an appropriate family of covariance functions. As discussed in Section 2.3, we decided to utilize polynomial covariance functions. This, however, does not mean that this choice is the definitive answer. In fact, variogram models are far easier to be interpreted and compared to the physical processe under

study. So researchers may find the use of variograms more appealing.

The problem of structural analysis is then studied. It should be noted that the results of structural analysis have to be checked thouroughly, in order to identify the best set of parameters. The procedure proposed by Delfiner (1975) has a tendency to yield covariance functions with large nugget effects. In the same time, it delets many valid forms of the covariance function that contain both the nugget effect and the structured part. This is caused by the fact that the above procedure idetermines parameter values, which do not satisfy the positive definite criterion.

The other problem which we encounter in the course of our study is the ill-conditioned matrices in the kriging system. As noted, this problem is partially due to the fact that some of rows associated with the same data location at different time intervals are rather similar.

Another reason for the above behavior is the limited number of neighboring data points. This condition gives a dominanat role to the drift block in the kriging matrix, as defined in (2.17). Consequently, the elements of the covariance block make little difference among these rows. It appears that in our next step we should first increase the number of neighboring points in order to give an advantage to the covariance block in the kriging process. We will also explore the possibilty of using approximate matrix inversion procedures, if the ill-conditioned matrices persist.

At the next phase of our study we focuse on the expansion of variance reduction analysis. Its application to non-negative log-normally distributed variables is rather straightforward. We also show that result produced for log-transformed values are equally valid for the original

data. This makes variance reduction analysis an effective procedure for identifying sequences of the best sampling points.

The above step is followed by the developement of space-time variance reduction analysis. Its theoretical formulation is presented in Section 4.3. Except the ill-conditioned matrices, we did not encounter any problem in the application of this procedure. Presently, we are focusing on applying this algorithm to identify the best route for a moving sampling device, such as an oceanographic vessel. In this approach we define the route as the best sequence of points, where its initial and final point, as well as, its duration is given. It can be easily observed that each point in this sequence is located at the vicinity of its proceeding point. This characteristic significantly reduces the needed number of computations. We hope that this will be the topic of one of our next projects.

In the last phase, we presented the result of a study on the resilience of variance reduction analysis. It shows while the parameters of the covariance function are highly unstable, the decisions made by varaince reduction analysis display a remarkable degree of resilience. Thus, it may be used as an effective tool for planning of sampling activities. Considering the high cost of sampling in hydrology, the use of variance reduction analysis may yield significant savings in time and money.

REFERENCES

Current List:

The current list is followed by the main list.

Ahmed, S., and G. DeMarsily, Comparison of geostatistical methods for estimating transmissivity using data on transmissivity and specific capacity, Water Resou. Res., Vol 23(9), pp. 1717-1737, 1987.

Baffi, E.Y., Y.C. Kim, and F. Szidarovsky, On nonnegative weights of linear kriging estimation, Min. Engin., Vol. 38(6), pp. 437-442, 1986.

Benjamin, J.R., and C.A. Cornell, Proability, Statistics, and Decision for Civil Engineers, pp. 269, McGraw-Hill, New York, 1970.

Bilonick, R.A., The space-time distribution of sulfate deposition in the northeastern United States, Atmospheric Environment, Vol. 19(11), pp. 1829-1845, 1985.

Bilonick, R.A., Monthly hydrogen ion deposition maps for the northeatern U.S. from July 1982 to September 1984, Consolidation Coal Co., Pittsburgh, 1987.

Cargile, K.A., A geostatistical method in drought management, M.S. Special Research Problem, School of Civil Engineering, Georgia Institute of Technology, Atlanta, 1987.

Clarke, J.S.,S.A. Longsworth, K.W. McFadden, and M.F. Peck, Ground-water data for Georgia, U.S.G.S., Open-File Report 85-331, Doraville, GA, 1985.

Delhomme, J.P., Modéles de simulation et de gestion des

ressources en eau des basins de l'Orne, la Dives, et la Seulles, Centre d'Informatique Géologique, ENSMP, Fontainebleau, 1977.

DeMarsily, G., and S. Ahmed, Application of kriging techniques in groundwater hydrology, Journal of the Geological Society of India, Vol 29(1), pp. 4-28, 1987.

Fiering, M. B, and J. Kindler, Surprise in water resources design, IIASA, Laxenburg, Austria, 1981.

Hall, T.J., Space-time kriging analysis of groundwater data, M.S. Special Research Problem, School of Civil Engineering, Georgia Institute of Technology, Atlanta, 1987.

Hayes, L.R., M.L. Maslia, and W.C. Meeks, Hydrology and model evaluation of the principal artesian aquifer, Dougherty Plain, Southwest Georgia, Bulletin 97, Georgia Geologic Survey, Atlanta, 1983.

Journel, A.G., and C. Huijbregts, Mining Geostatistics, Academic Press, New York, 1977.

Matheron, G. Eléments pour une théorie des milieux poreux, Masson, Paris, 1967.

Meyer, S.T., Data Analysis for Scientists and Engineers, pp. 285, John Wiley & Sons, Inc., New York, 1975.

Rouhani, S., Variance reduction analysis, Water Resources Research, Vol. 21, No. 6, pp. 837-846, June 1985.

Rouhani, S., Comparative study of ground water mapping techniques, Journal of Ground Water, Vol. 24, No. 2, pp. 207-216, March-April 1986.

Rouhani, S., and M. B Fiering, Resilience of a statistical sampling scheme, Journal of Hydrology, Vol. 89, pp. 1-11, December 1986.
Also presented in a brief form at the Americal Geophysical Union Fall Meeting, San Francisco, December, 1985. (Abstarct published in EOS, Vol. 66, No. 46, pg. 897, November, 1985).

Rouhani, S., Optimal sampling of stochastis processes, Annual Progress Report, School of Civil Engineering, Georgia Institute of Technology, Atlanta, 1987.

Rouhani, S., Water resources monitoring: A combined information-economic approach, Journal of Resouces Policy, under review, 1987.

Rouhani, S., and T.J. Hall, Optimal schemes for ground water monitoring in the shallow aquifer, Dougherty Plain, Southwestern Georgia, Technical Completion Report, USDI/USGS Project G-1219 (05), School of Civil Enginnering in cooperation with Environmental Resources Center, Georgia Institute of Technology, Atlanta, GA, March 1987.

Rouhani, S., and T.J. Hall, Geostatitical schemes for regional ground water monitoring, Proceedings of the Third National Groundwater Technology Conference, City University of New York, NY, September 1987.

Rouhani, S., and T.J. Hall, Geostatistical schemes for groundwater sampling, Journal of Hydrology, under review, 1987.

Rouhani, S., and T.J. Hall, Space-time kriging analysis of groundwater data, Proceedings of the Third International Geostatistics Congress, Avignon, France, September, 1988.

Szidarovsky, F., F.Y. Baffi, and Y.C. Kim, Kriging without

negative weights, Math. Geology, Vol. 38(6), pp. 437-442, 1987.

Yates, S.R., Warrick, D. Meyers, Disjunctive kriging, Part 1 and 2, Water Resour. Res., Vol. 22(5), pp. 615-630, 1986a and 1986b.

MAIN LIST

Bakr, A.A., L.W. Gelhar, A.L.Gutjahr, and J.R. MacMillan, Stochastic analysis of spatial variability in subsurface flows, 1. Comparison of one- and three-dimensional flows, Water Resources Res., 14(2), 263-271, 1978.

Bear, J., Hydraulics of Groundwater, McGraw-Hill, New York, 1979.

Ben-Zvi, M. and Y. Bachmat, Management of a water resource under uncertainty, Hydrological Service, Jerusalem, Israel, 1979.

Box, G.E.P. and G.M. Jenkins, Time Series Analysis, Forecasting and Control, Holden-Day, San Francisco, 1970.

Bras, R.L. and I. Rodriguez-Iturbe, Network design for the estimation of areal mean of rainfall events, Water Resources Res., 12(6), 1185-1208, 1976.

Bryant, G.T., Stochastic theory of queues applied to design of impounding reservoirs, Ph.D. Thesis, Division of Applied Sciences, Harvard University, Cambridge, 1961.

Carter, E.F. and A.R. Robinson, Synoptic maps of the main thermocline from polymode XBTS: A time series via space-time objective analysis, Center for Earth and Planetary Physics, Harvard University, Cambridge, October, 1982.

Chirlin, G.R. and G. Dagan, Theoretical head variogram for steady flow in statistically homogeneous aquifers, Water Resources Res., 16(6), 1001-1015, 1980.

Chirlin, G.R. and E.F. Wood, One the relationship between kriging and state estimation, Water Resources Res., 18(2), 432-438, 1982.

Chua, S.H. and R.L. Bras, Estimation of stationary and non-stationary random fields: Kriging in the analysis of orographic precipitation, Report No. 255, Ralph M. Parsons Lab., M.I.T., Cambridge, April 1980.

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

Cooley, R.L., A method of estimating parameters and assessing reliability for models of steady state groundwater flow, 1. Theory and numerical properties, Water Resources Res., 13(2), 318-324, 1977.

David, M., Geostatistical ore reserve estimation, Elsevier Scientific Publishing Co., Amsterdam, 1977.

Delfiner, P., Linear estimation of non-stationary spatial phenomena, in Guarascio et al., Advanced Geostatistics in the Mining Industry, Nato Advanced Study Institute, D. Reidel Publishing Co., Boston, 1975.

Delfiner, P. and J.P. Delhomme, Optimum interpolation by kriging, in Davis et al., Display and Analysis of Spatial Data, Nato Advanced Study Institute, John Wiley and Sons, New York, 1975.

Delhomme, J.P., Spatial variability and uncertainty in groundwater flow parameters: A geostatistical approach, Water Resources Res., 15(2), 269-280, 1979.

DeMarsily, G., The estimation of parameters (e.g., transmissivity) in an aquifer: A stochastic approach, presented at Ralph M. Parsons Lab., M.I.T., Cambridge, 1979.

Dettinger, M.D. and J.L. Wilson, Numerical modeling of aquifer systems under uncertainty: A second moment analysis, M.I.T./Cairo University Technological Planning Program, Department of Civil Engineering, M.I.T., Cambridge, May 1979.

Dunlap, L.E. and J.M. Spinazola, Interpolating water-table altitudes in West-Central Kansas, Open-File Report 81-1062, U.S.G.C. 1981.

Fiering, M.B., An optimization scheme for gaging, Water Resources Res., 1(4), 463-470, 1965.

Freeze, R.A., A stochastic-conceptual analysis of one-dimensional groundwater flow in non-uniform homogeneous media, Water Resources Res., 11(5), 725-741, 1975.

Gambolati, G. and G. Volpi, Groundwater contour mapping in Venic by stochastic interpolators, 1. Theory, Water Resources Res., 15(2), 281-197, 1979.

Gandin, L.S., Objective Analysis of Meteorological Fields, Israel Program for Scientific Translations, Jerusalem, 1965.

Hughes, J.P. and D.L. Lettenmaier, Data requirement for kriging: Estimation and network design, Water Resources Res., 17(6), 1641-1650, 1981.

Kafritsas, J. and R.L. Bras, The practice of kriging, Report No. 263, Ralph M. Parsons Lab., M.I.T., Cambridge, January 1981.

Kitanidis, P.K. and E.G. Vomvoris, A geostatistical approach to the inverse problem in groundwater modeling (steady state), 1, Institute of Hydraulic Research, The University of Iowa, Iowa City, August 1982.

Maddock, T. III, Management model as a tool for studying the worth of data, Water Resources Res., 9(2), 270-280, 1973.

Maddock, T. III, Network design on a formal sensitivity analysis, Presented at Socorro Workshop on stochastic methods in subsurface hydrology, Socorro, New Mexico, April 26-27, 1979.

Mantoglou, A. and J.L. Wilson, Simulation of random fields with the turning bands method, Report No. 264, Ralph M. Parsons Lab., M.I.T., Cambridge, July 1981.

Marin, C., Parameter uncertainty in water resource planning, Ph.D. Thesis, Division of Applied Sciences, Harvard University, Cambridge, 1983.

Matalas, N.C., Optimum gaging station location, Proceedings IBM Scientific Computing Symposium, Water and Air Resources Management, 85-94, White Plains, New York, 1968.

Matalas, N.C., Notes on time series, private communication to M.B. Fiering (Division of Applied Sciences, Harvard University), 1969.

Matalas, N.C. and M.B. Fiering, Water-resources systems planning, Climate, Climatic Changes, and Water Supply, N.A.S., Washington, D.C., 1977.

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

Meyer, S.L., Data Analysis for Scientific and Engineers, p. 390, John Wiley and Sons, Inc., New York, 1975.

Neuman, S.P. and S. Yakowitz, A statistical approach to the inverse problem of aquifer hydrology, 1. Theory, Water Resources Res., 15(4), 845-860, 1979.

Noble, B. and J.W. Daniel, Applied Linear Algebra, Second ed., p. 28, Prentice Hall, Englewood Cliff, New Jersey, 1977.

Pimental, K.D., Asymptotic estimation error growth applied to monitoring, Chapman Conference on Applications on Kalman Filter to Hydrology, Hydraulics and Water Resources, A.G.U., Pittsburgh, PA, May 1978.

Rodriguez-Iturbe, I. and J.M. Mejia, The design of rainfall networks in time and space, Water Resources Res., 10(4), 713-728, 1974.

Rouhani, S., Optimal interpolation of regionalized variables: An application of kriging to groundwater mapping, Division of Applied Sciences, Harvard University, Cambridge, June 1981.

Rouhani, S., Optimal data collection in random fields, Ph.D. Thesis, Division of Applied Sciences, Harvard University, Cambridge, 1983.

Rouhani, S., Optimal groundwater data collection, waterlogging and salinity considerations, Proceedings International Seminar on Water Resources Management, Lahore, Pakistan, 1983a.

Schweppe, F.C., Uncertain Dynamic Systems, Prentice-Hall, Englewood Cliffs, New Jersey, 1973.

Smith, L. and R.A. Freeze, Stochastic analysis of steady state groundwater flow in a bounded domain, 1. One-dimensional simulations, Water Resources Res., 15(3), 521, 1979.

Sophocleous, M., J.E. Paschetto, and R.A. Olea, Groundwater network design for Northwest Kansas, using the theory of regionalized variables, submitted to Water Resources Bull., Kansas Geological Survey, Lawrence, Kansas, March 1981.

Tu, K., A combined statistical and dynamical approach to regional forecast modeling of open ocean currents, Ph.D. Thesis, Division of Applied Sciences, Harvard University, Cambridge, February 1981.

Yeh, W.W.G., Y.S. Yoon, and K.S. Lee, Aquifer parameter identification with kriging and optimum parametrization, Water Resources Res., 19(1), 225, 1983.

APPENDIX 1. EQUIPMENT

As noted in the proposal, we aim at applying the above procedure to actual field activities. This appear to be accomplished by adapting the program for micro-computers. To achieve this an IBM PC AT with color display and an HP Thinkjet printer was purchased at a dicounted price, half of which was paid by Georgia Institute of Technology.

The above equipment played a major role on our efferts. First, all administrative and word processing tasks associated with the project, were conducted on the above PC. They include a number of papers, reports, and realted proposals. Secondly, we use it as the main terminal when using the mainframe for our computations and mapping activities.

The application of space-time variance reduction analysis to micro-computers appear to be straightforward. However, judging from the required time of computations on the main frame, we decided to perform all our runs on the CYBER computer that is available at Georgia Tech. We believe that in very near future the speed of micros will be sufficiently high to be able to run variance reduction analysis for large data bases on PCs.

APPENDIX 2. STUDENT TRAINING AND ABSTRACTS OF THESES

In the course of this project, three graduate students were involved. The first one was Mr. M. Zakikhani, a Ph.D. student in geohydrology, who was supported for a period of three months. During this period he was trained in using spatial universal kriging and variance reduction analysis.

The second student was Mr. T. J. Hall who was supported throughout his M.S. studies. For a period of three months he studied spatial kriging and variance reduction analysis. He then assisted the PI on the extension of these algorithms to non-negative phenomena. At this stage, he was jointly supported by another project of the PI, funded by U.S. Geological Survey for applying these methods to the problem of groundwater quality monitoring network in Dougherty Plain, southwestern Georgia. For his special M.S. project, he selected the topic of space-time universal kriging, whose results are given in Section 2.4.1. He also assisted the PI in developing the time kriging computer program. Mr. Hall graduated in September, 1987, and presently works as a geohyrologist in an environmental consulting firm in Massachusetts.

The third student was Mr. K. A. Cargile who studied under the direct supervision of the PI for a period of 9 months. During this period he worked on the application of time kriging in drought management schemes, as his special M.S. project. A brief summary of his work is given in Section 2.3.1. He received his M.S. degree in September, 1987. He currently works in an enginnering consulting firm in Georgia.

The abstarcts of the above M.S. special problems are given in this appendix.

SPACE-TIME KRIGING ANALYSIS OF GROUNDWATER DATA

A Special Research Problem

Presented to

The Faculty of the School of Civil Engineering
Georgia Institute of Technology

by

Timothy J. Hall

In Partial Fulfillment
of the Requirements for the Degree of
Master of Science in Civil Engineering

August 1987

ABSTRACT

There are many naturally occurring processes and parameters which can be described as stochastic processes. These processes can be interpolated by using a Gauss-Markov estimator such as kriging. Presently most kriging packages are designed for estimation of spatially random variables. It is shown that with certain modifications, kriging can be expanded to the space-time domain to be applicable to a more general class of stochastic processes. This is analogous to combining spatial analysis with time series analysis. In this study a series of hydrologic data from Georgia is simultaneously analyzed in time and space using kriging, and the results are presented in a series of spatial maps for different time periods. In this way valuable new information has been gained by utilizing both the spatial and the temporal data. Space-time kriging also yields more accurate results by allowing the addition of all the available space-time data. Finally, it allows hindcasting and forecasting for periods when no sampling is conducted.

A GEOSTATISTICAL METHOD IN DROUGHT MANAGEMENT

A Special Research Problem
Presented to
The Faculty of the School of Civil Engineering

by
Kenneth Alvin Cargile

In Partial Fulfillment
of the Requirements for the Degree
Master of Science
in the School of Civil Engineering

Georgia Institute of Technology

September, 1987

ABSTRACT

A Geostatistical Method in Drought Management

Kenneth Alvin Cargile

Directed by Dr. S. Rouhani

This report presents research findings and design analyses for a water resources engineering project. The project involves the planning for the conditions of a potential water management deficit. Planning requires the definition of a drought and the implementation of design methods to prepare water management planners and operators for the drought condition.

Planning for extreme hydrological events requires the analysis of statistical data. These events can be characterized as stochastic processes, and the geohydrological variables such as low streamflows into a reservoir can be viewed as random fields. Analysis of historic hydrological data allows the planner to derive the means to predict the outcome of extreme events in the future.

A proposed site for a reservoir is analyzed for its potential water usage to demonstrate the capability of the reservoir to perform under the drought condition. A geostatistical method is presented for applications in water management to predict the impact of drought on a proposed reservoir. This method can also be used to assess the reliability of an existing reservoir for its performance during a drought.

APPENDIX 3. RELATED RESEARCH AND SCIENTIFIC COLLABORATORS

During this project two related projects were proposed and funded. The first one was titled: "Optimal schemes for ground water qulaity monitoring in the shallow aquifer, Dougherty Plain, southwestern Georgia." This project was funded by a grant from U.S. Departemnt of Interior, Geological Survey; USDI/USGC Project G-1219 (05), for the period of April 1986 to March 1987.

The above USGS project was jointly conducted with the PI's initiation project, as discussed in Sections 2.4, 2.4.1, 3.1, and 3.1.1. The abstract of the final report of this project is given in this appendix.

The second realted project that was proposed by the PI and was later funded, is titled: "Advanced geostatistical studies at the Centre de Geostatistique, Ecole des Mines de Paris." This project was funded by Natioanl Science Foundation under the US-Industrialized Countries Program for the Exchange of Scientists and Engineers; Project No. INT-8702264, for the period of September 1987 to October 1988.

This project is in fact a continuation of the present project which allows the PI to conduct advanced research at the Centre de Geostatistique, Ecole des Mines de Paris for furhter studies on space-time kriging and varaince reduction analysis. The technical abstract of the proposal is given at the end of this appendix.

In the following page a list of scientific collaborators of the PI in his initiation project is given.

List of Scientific Collaborators
NSF Grant No. ECE-8503897

1. Prof. Myron B Fiering
Gordon McKay Professor of Environmental Engineering
Division of Applied Sciences
Harvard University
Cambridge, MA 02138

2. Mr. Timothy J. Hall
Graduate Research Assistant
School of Civil Engineering
Georgia Institute of Technology
Atlanta, GA 30332

3. Mr. Kenneth A. Cargile
Graduate Student
School of Civil Engineering
Georgia Institute of Technology
Atlanta, GA 30332

4. Mr. Mansour Zakikhani
Graduate Research Assistant
School of Civil Engineering
Georgia Institute of Technology
Atlanta, GA 30332

ERC 03-87 March 1987

OPTIMAL SCHEMES FOR GROUND WATER QUALITY MONITORING IN THE SHALLOW AQUIFER, DOUGHERTY PLAIN, SOUTHWESTERN GEORGIA

by

SHAHROKH ROUHANI, Ph.D.
Assistant Professor
of Hydrology and Water Resources

and

TIMOTHY J. HALL
Graduate Research Assistant

Technical Completion Report
USDI/USGS Project G-1219 (05)

The research on which this report is based was financed in part by the U.S. Department of the Interior, Geological Survey as authorized by the Water Resources Research Act of 1984 (P.L. 98-242)

Contents of this publication do not necessarily reflect the views and policies of the U.S. Department of the Interior, nor does mention of trade names or commercial products constitute their endorsement or recommendation for use by the U.S. Government.

School of Civil Engineering
in cooperation with
Environmental Resources Center
Georgia Institute of Technology
Atlanta, Georgia 30332

ABSTRACT

Geostatistical schemes for ground water quality monitoring in the shallow aquifer of Dougherty Plain, Georgia are presented. This aquifer is not generally used for water supply purposes. However, it is the main recharge route to the principal artesian aquifer which is the primary source of water supply in this rapidly growing agricultural region. The desired monitoring network acts as an early warning system for ground water pollution in deeper layers. We have utilized the available data on hydraulic properties of the shallow aquifer to identify the zones which should be the primary locations for our sampling activities. The one variable which appears to be most suitable for our study is leakance. Statistical analyses indicate that leakance has a log-normal distribution with a constant trend and a linear covariance function. Ranking criteria for the selection of the best sampling points are: the variance reductions, the medians, and the risk values. Due to the nature of our monitoring network we suggest to use mainly risk ranking as the basis of our sampling activities. The results of our risk rankings demonstrate that the southern tip of the Dougherty Plain and its upper central zone should be the prime targets of our monitoring activities.

Keywords: Network Design, Statistical Methods, Regional Analysis, Water Quality, Water Management (Applied), Georgia.

PROPOSAL TO THE NATIONAL SCIENCE FOUNDATION
Cover Page

FOR CONSIDERATION BY NSF ORGANIZATIONAL UNIT (Indicate the most specific unit known, i.e. program, division, etc.)	IS THIS PROPOSAL BEING SUBMITTED TO ANOTHER FEDERAL AGENCY? Yes____ No__X__ IF YES, LIST ACRONYM(S):
U.S.-Industrialized Countries Program for the Exchange of Scientists and Engineers (INT)	

PROGRAM ANNOUNCEMENT/SOLICITATION NO:	CLOSING DATE (IF ANY) Nov. 1, 1986

NAME OF SUBMITTING ORGANIZATION TO WHICH AWARD SHOULD BE MADE (INCLUDE BRANCH/CAMPUS/OTHER COMPONENTS)

GEORGIA TECH RESEARCH CORPORATION

ADDRESS OF ORGANIZATION (INCLUDE ZIP CODE)

GEORGIA INSTITUTE OF TECHNOLOGY
ATLANTA, GA 30332-0420

TITLE OF PROPOSED PROJECT

Advanced Geostatistical Studies at the Centre de Geostatistique, Ecole des Mines de Paris

REQUESTED AMOUNT	PROPOSED DURATION	DESIRED STARTING DATE
$ 72,066.	1 year	Sept. 1, 1987

PI/PD DEPARTMENT	PI/PD ORGANIZATION	PI/PD PHONE NO.
School of Civil Engineering	GEORGIA INSTITUTE OF TECHNOLOGY	(404)894-2201

PI/PD NAME	SOCIAL SECURITY NO.*	SIGNATURE	MALE*	FEMALE*
Dr. Shahrokh Rouhani	026-52-0706		X	

Technical Abstract

A significant number of natural and physical variables in hydrology, hydrogeology, and oceanography can often be viewed as stochastic processes. Geostatistics provides a means for the statistical study of such processes. This branch of applied statistics was first developed at the Geostatistics Center, School of Mines of Paris, under the direction of Prof. G. Matheron. Since its inception in 1968, this center has been one of the main research groups in the world in the field of geostatistics- both on the theoretical and the applied sides.

In this proposal I am seeking support for a one year sabbatical-type visit to conduct joint research with members of the above group and its associated centers in Fontainebleau, France. My first objective is to incorporate the disjunctive kriging into the variance reduction analysis (a geostatistical sampling scheme developed by the PI.) This expansion enables the program to identify the optimal sequence of sampling points for random variables with any distribution. It will be attempted to expand this algorithm to the time-space domain to be applicable to most physical processes in water sciences. My attention will then be focused on the application of the expanded variance reduction analysis to actual field cases, including the water quality sampling in Lake Geneva, ground water observation networks in Northern France, and rainfall-gage networks in the Aquitaine Basin.

As stated in the official letters from the above centers, they also believe that this cooperative project can lead to some very fruitful research in the study of spatiotemporal variables, and in the applied fields of hydrogeology and oceanography. At the same time this project provides an opportunity to develop and further stimulate scientific, engineering, and technical cooperation between the United States and France.

APPENDIX 4. PUBLICATION CITATIONS

Since the initiation of this project a number of publications are produced by the team lead by the PI. The following lists and corresponding abstarcts include only the completed works. Two papers on space-time kriging and varaiance reduction analysis, as well as, another one on drought management are under preparations. We intend to submit them for publication to Water Resources Research and Water International. The completed works are as follows:

Rouhani, S., Variance reduction analysis, Water Resources Research, Vol. 21, No. 6, pp. 837-846, June 1985.

Rouhani, S., Comparative study of ground water mapping techniques, Journal of Ground Water, Vol. 24, No. 2, pp. 207-216, March-April 1986.

Rouhani, S., and M. B Fiering, Resilience of a statistical sampling scheme, Journal of Hydrology, Vol. 89, pp. 1-11, December 1986.
Also presented in a brief form at the Americal Geophysical Union Fall Meeting, San Francisco, December, 1985. (Abstarct published in EOS, Vol. 66, No. 46, pg. 897, November, 1985).

Rouhani, S., Water resources monitoring: A combined information-economic approach, Journal of Resouces Policy, under review, 1987.

Rouhani, S., and T.J. Hall, Optimal schemes for ground water monitoring in the shallow aquifer, Dougherty Plain, Southwestern Georgia, Technical Completion Report, USDI/USGS Project G-1219 (05), School of Civil Enginnering in cooperation with Environmental Resources Center, Georgia Institute of Technology, Atlanta, GA, March 1987.

Rouhani, S., and T.J. Hall, Geostatitical schemes for regional ground water monitoring, Proceedings of the Third National Groundwater Technology Conference, City University of New York, NY, September 1987.

Rouhani, S., and T.J. Hall, Geostatistical schemes for groundwater sampling, Journal of Hydrology, under review, 1987.

Rouhani, S., and T.J. Hall, Space-time kriging analysis of groundwater data, Proceedings of the Third International Geostatistics Congress, Avignon, France, September, 1988.

The abstarcts of these publication are given in the following pages. Reprints of the published articles are also included.

WATER RESOURCES RESEARCH, VOL. 21, NO. 6, PAGES 837–846, JUNE 1985

Variance Reduction Analysis

SHAHROKH ROUHANI

School of Civil Engineering, Georgia Institute of Technology, Atlanta

This paper presents an algorithm for optimal data collection in random fields, the so-called variance reduction analysis, which is an extension of kriging. The basis of variance reduction analysis is an information response function (i.e., the amount of information gain at an arbitrary point due to a measurement at another site). The ranking of potential sites is conducted using an information ranking function. The optimal number of new points is then identified by an economic gain function. The selected sequence of sites for further sampling shows a high degree of stability with respect to noisy inputs.

1. INTRODUCTION

Many physical variables involved in hydrological phenomena may be viewed as random fields, also known as regionalized variables [e.g., *David*, 1977]. The geohydrological variables such as transmissivity, storativity, and steady state piezometric heads are of this type. Examples of the stochastic analysis of these variables can be found in the work of such authors as *Freeze* [1975], *Smith and Freeze* [1979], *Bakr et al.* [1978], *Dettinger and Wilson* [1979], *Delhomme* [1979], *Gambolati and Volpi* [1979], *Chirlin and Dagan* [1980], *Clifton and Neuman* [1982], and *Yeh et al.* [1983].

The data management of these spatially distributed variables can be studied in the framework of random fields. For such fields, the location and rates of sampling depend upon the objectives of the planning approach. Often very little data are available. Furthermore, the measured values may be clustered together and therefore not provide information about the whole field. For example, the study of water table data in northwestern Kansas [*Rouhani*, 1983] revealed that most measured values were clustered around major towns and farm communities. Consequently, a significant portion of the whole region was sporadically sampled. In such situations, planners may wish to design a data collection scheme in order to better define the variable of interest.

The following questions then arise.

1. Where are the optimal locations for further sampling?
2. What is the optimal size of the sample set?

In order to answer the first question there is an initial need to quantify the uncertainty in the estimated field variable at any one point. Kriging provides such an indicator: the estimation variance. For instance, to minimize the regional variance of estimation one may add a measurement point at the site with maximum estimation variance [see *Matalas*, 1968; *DeMarsily*, 1979]. However, the estimation variance alone is not sufficient. One needs an indicator of the relative influence of the added sample on the reliability of the whole field in order to select a point providing maximum information gain. In response to this problem a new algorithm called variance reduction analysis is developed, yielding a method for the selection of sequences of sites for further sampling in random fields.

Two optimality criteria are utilized for the ranking of potential sampling sites. The first one reflects the amount of information gain (i.e., the variance reduction) due to a new measurement. The second function is proportional to the expected economic gains (i.e., the loss reduction) due to further sampling.

Finally, one is faced with the more sophisticated question:

3. How reliable are these decisions?

To answer this question it is necessary to study the resilience of the prescribed decisions of the variance reduction analysis.

This paper is divided into three parts. The first part is devoted to a brief review of kriging and variance reduction analysis. In the second part the author describes several data collection management approaches and discusses the advantages of the proposed algorithm, which leads to a process for ranking of prospective sampling sites. In the last section the author applies this algorithm to water table level observations in northwestern Kansas and briefly discusses the reliability and resilience of the variance reduction analysis.

2. DEVELOPMENT OF VARIANCE REDUCTION ANALYSIS

2.1. *Kriging Method*

Kriging has been applied to groundwater hydrology by such authors as *Delhomme* [1979], *DeMarsily* [1979], *Gambolati and Volpi* [1979], *Chirlin and Dagan* [1980], *Dunlap and Spinazola* [1981], *Sophocleous et al.* [1982], *Clifton and Neuman* [1982], *Kitanidis and Vomvoris* [1982], *Aboufirassi and Marino* [1983], and *Yeh et al.* [1983]. In these papers kriging was used mainly as a tool for the interpolation of either transmissivities or piezometric heads.

In point kriging one estimates the value of the random field at an arbitrary point X_0 based on the given measured values in a linear form of

$$\hat{Z}(X_0) = \sum_{i=1}^{N} \lambda_{i0} Z(X_i) \tag{1}$$

where

$\hat{Z}(X_0)$ kriging estimate at X_0;
$Z(X_i)$ measured value at X_i, $i = 1, \cdots, N$;
λ_{i0} kriging weight for $Z(X_i)$ to estimate $\hat{Z}(X_0)$.

The λ_{i0} are defined by two criteria: (1) unbiasedness: $E[\hat{Z}(X_0) - Z(X_0)] = 0$, where $Z(X_0)$ is the true value of the field at X_0, and (2) minimum squared error of estimation: this requires $E[\hat{Z}(X_0) - Z(X_0)]^2$ to be minimum. These conditions can be written as

$$E[\hat{Z}(X_0) - Z(X_0)] = 0$$
$$\text{Var}\,[\hat{Z}(X_0) - Z(X_0)] = \min \tag{2}$$

Paper number 5W0116.
0043-1397/85/005W-0116$05.00

where Var $[\hat{Z}(X_0) - Z(X_0)]$ is known as the estimation or kriging variance.

In kriging one may view the process $Z(X)$ as a spatial random function with the following structure:

$$Z(X) = M(X) + R(X) \tag{3}$$

where $M(X)$ is a slowly varying deterministic function known as the "drift," which is equal to the expected value of Z at point $X \in \mathbb{R}^n$. It may be further assumed that $M(X)$ admits a local representation in the form of a polynomial of order k as follows:

$$M(X) = \sum_{p=1}^{l(k)} b_p f_p(X) \tag{4}$$

where b_p are fixed unknown coefficients, and $f_p(X)$ are basic monomials of the polynomial

$$f_p(X) = x_1^{p_1} x_2^{p_2} \cdots x_n^{p_n} \qquad p_1 + p_2 + \cdots + p_n \le k \tag{5}$$

$l(k)$ is the number of such monomials in $M(X)$. $R(X)$ is a spatially fluctuating random function with zero expectation.

Matheron [1973] proposes a new method in which the process $Z(X)$ is viewed as an intrinsic random function (*IRF*), which could be made stationary by a process known as "incrementing." A kth-order intrinsic random function (*IRF-k*) is defined as a random process that requires a kth order filtering to achieve stationarity. The linear combination $\sum_{i=0}^{N} \lambda_{i0} Z(X_i)$ is a generalized increment of order k, if and only if

$$\sum_{i=0}^{N} \lambda_{i0} (x_{1i}^{p_1})(x_{2i}^{p_2}) \cdots (x_{ni}^{p_n}) = 0 \tag{6}$$

for all integers $p_1, \cdots, p_n \ge 0$ such that $p_1 + p_2 + \cdots + p_n \le k$, where X_i stands for the point $(x_{1i}, \cdots, x_{ni})$ in n space $(\lambda_{00} = -1)$.

For the case of an *IRF* of order 0, 1, or 2 in $\mathbb{R}^2$ with Cartesian coordinates (x_i, y_i), (6) can be written as

$$\begin{aligned}
&k = 0 \quad && \sum_{i=0}^{N} \lambda_{i0} = 0 && \\
&k = 1 \quad && \sum_{i=0}^{N} \lambda_{i0} x_i = 0 && \sum_{i=0}^{N} \lambda_{i0} y_i = 0 \\
&k = 2 \quad && \sum_{i=0}^{N} \lambda_{i0} x_i y_i = 0 && \sum_{i=0}^{N} \lambda_{i0} x_i^2 = 0 \\
& && \sum_{i=0}^{N} \lambda_{i0} y_i^2 = 0 &&
\end{aligned} \tag{7}$$

The above constraints constitute the unbiasedness criterion of the original kriging (2).

Variance of estimation (2) can be written as

$$\text{Var}\left[\sum_{i=0}^{N} \lambda_{i0} Z(X_i)\right] = \sum_{i=0}^{N} \sum_{j=0}^{N} \lambda_{i0} \lambda_{j0} K(|X_i - X_j|) \tag{8}$$

where $K(|X_i - X_j|)$ = covariance function of $Z(X_i)$ and $Z(X_j)$.

Matheron [1973] proposes a polynomial function of $2k + 1$ order as the generalized covariance (GC) for an *IRF-k*, as follows

$$K(h) = C_0 \delta(h) + \sum_{p=0}^{k} a_{2p+1} h^{2p+1} \tag{9}$$

where

- h length of vector distance between two points;
- C nugget effect;
- $\delta(\ )$ Dirac's delta function.

Now in order to calculate λ_{i0}, it is necessary to minimize (8) subject to constraints (6). By using the Lagrange multiplier μ_{p0}, (8) can be minimized with respect to λ_{i0} and μ_{p0} if

$$\begin{aligned}
&\sum_{j=1}^{N} \lambda_{j0} K(|X_i - X_j|) + \sum_{p=1}^{l(k)} \mu_{p0} f_p(X_i) = K(|X_0 - X_i|) \qquad i = 1, \cdots, N \\
&\sum_{j=1}^{N} \lambda_{j0} f_p(X_j) = f_p(X_0) \qquad p = 1, \cdots, l(k)
\end{aligned} \tag{10}$$

The above set of equations is the so-called "kriging system." At its minimum the estimation variance (8) takes the value of

$$\text{Var}\,[\hat{Z}(X_0) - Z(X_0)] = K(|X_0 - X_0|) - \sum_{j=1}^{N} \lambda_{j0} K(|X_0 - X_j|) - \sum_{p=1}^{l(k)} \mu_{p0} f_p(X_0) \tag{11}$$

2.2. *Variance Reduction Analysis*

The kriging variance (11) can be utilized as a guideline for optimal sampling [see *DeMarsily*, 1979]. For instance, the area with the highest level of estimation uncertainty can be targeted for further monitoring. However, such an approach ignores the overall effect of a new measurement on the level of accuracy of the estimated field as a whole. In particular, it overlooks the influence of added data on the estimation variances of other interpolated values. The author proposes an algorithm to establish a measure for such an influence.

As the first step, a relationship is established between the reduction in kriging variance at an arbitrary point with respect to the sampling at another location. This relationship resembles a common "response" function. It gives the level of improvement in the accuracy of $\hat{Z}(X_0)$ due to a new measurement at X_*. This level of improvement is measured in terms of reductions in the kriging variances. Furthermore, this measure of variance reduction can be expanded to cover the whole field. This enables the planner to rank the prospective locations for further data collections.

In order to obtain this response function the kriging system (10) may be written in matrix form. This system is composed of $N + l$ equations, where N is the number of data points used in kriging, and l is the number of monomials in the drift function, which is a function of the order of *IRF*.

In the case of an *IRF*-2 in $\mathbb{R}^2$, kriging system (10) can be written as

$$\begin{bmatrix}
0 & 0 & \cdots & 0 & 0 & 1 & \cdot & \cdots & 1 \\
0 & \cdot & & \cdot & 0 & x_1 & & & x_N \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & & \vdots \\
0 & 0 & & 0 & 0 & y_1^2 & & & y_N^2 \\
1 & x_1 & \cdots & \cdot & y_1^2 & K_{11} & K_{12} & \cdots & K_{1N} \\
1 & x_2 & \cdots & \cdot & y_2^2 & K_{21} & K_{22} & \cdots & K_{2N} \\
\vdots & & & & & \vdots & & & \\
1 & x_N & \cdots & \cdot & y_N^2 & K_{N1} & \vdots & \cdots & K_{NN}
\end{bmatrix}
\begin{bmatrix} \mu_{10} \\ \mu_{20} \\ \vdots \\ \mu_{60} \\ \lambda_{10} \\ \lambda_{20} \\ \vdots \\ \lambda_{N0} \end{bmatrix}
=
\begin{bmatrix} 1 \\ x_0 \\ \vdots \\ y_0^2 \\ K_{10} \\ K_{20} \\ \vdots \\ K_{N0} \end{bmatrix} \tag{12}$$

where K_{ij} is the covariance between X_i and X_j, λ_{i0} is the kriging weight of $Z(X_i)$ to interpolate $\hat{Z}(X_0)$, and μ_{p0} is the pth Lagrange multiplier in the kriging system (10). The above equation can be written as

$$Aw_0 = a_0 \tag{13}$$

with the obvious notations.

The kriging variance at X_0 estimated by N existing data points can then be denoted by $V_0(N)$ in the following form, when $k = 2$

$$V_0(N) = K_{00} - [\mu_{10} \cdots \mu_{60}\ \lambda_{10} \cdots \lambda_{N0}] \begin{bmatrix} 1 \\ x_0 \\ \vdots \\ y_0^2 \\ K_{10} \\ \vdots \\ K_{N0} \end{bmatrix} \tag{14}$$

From the definitions of w_0 and a_0

$$V_0(N) = K_{00} - w_0{}^T a_0 \tag{15}$$

and from (13) and the symmetry of A,

$$V_0(N) = K_{00} - a_0{}^T A^{-1} a_0 \tag{16}$$

Superscript T defines the transpose of a vector.

Equation (16) can be expanded to include the effect of a new added measurement. If one adds a new measurement point at X_*, the A matrix in (13) acquires the form of a bordered matrix A_*. A_* is A with a new bottom row and a new right-hand side column. Consequently, $V_0(N+1)$ can be written as follows:

$$V_0(N+1) = K_{00} - [a_0{}^T\ K_{0*}] A_*{}^{-1} \begin{bmatrix} a_0 \\ K_{0*} \end{bmatrix} \tag{17}$$

where

$$A_* = \begin{bmatrix} A & a_* \\ a_*{}^T & K_{**} \end{bmatrix}$$

$$a_*{}^T = [1\ x_*\ \cdots\ y_*{}^2 K_{1*} K_{2*} \cdots K_{N*}]$$

$$K_{0*} = K(|X_0 - X_*|)$$

$$K_{**} = K(|X_* - X_*|)$$

$$K_{i*} = K(|X_i - X_*|)$$

Nobel and Daniel [1977] introduce a theorem concerning the bordered matrices which says that

$$A_*{}^{-1} = \begin{bmatrix} A & a_* \\ a_*{}^T & K_{**} \end{bmatrix}^{-1} = \begin{bmatrix} F & p \\ p^T & \alpha \end{bmatrix} \tag{18}$$

where

$$F = A^{-1} + \alpha A^{-1} a_* a_*{}^T A^{-1}$$

$$\alpha = [K_{**} - a_*{}^T A^{-1} a_*]^{-1} = [V_*(N)]^{-1}$$

$$p = -\alpha A^{-1} a_*$$

and A is invertible. Substituting (18) into (17) yields

$$V_0(N+1) = K_{00} - a_0{}^T A^{-1} a_0 - \alpha a_0{}^T w_* w_*{}^T a_0 + 2\alpha K_{*0} a_0{}^T w_* - \alpha K_{*0}{}^2 \tag{19}$$

Considering (16) one can write the variance reduction as

$$V_0(N) - V_0(N+1) = \frac{1}{V_*(N)} [K_{*0} - a_0{}^T w_*]^2 \tag{20}$$

where

$w_*{}^T = [\mu_{1*} \mu_{2*} \cdots \mu_{6*} \lambda_{1*} \cdots \lambda_{N*}]$;

$V_*(N)$ variance of estimation at X_* prior to any sampling at that point;

λ_{i*} kriging weight of $Z(X_i)$ to estimate $\hat{Z}(X_*)$ prior to sampling at X_*;

μ_{p*} pth Lagrange multiplier in the kriging system for estimation of $\hat{Z}(X_*)$ prior to sampling at X_*.

Finally, substituting the elements of w_* and a_0 into (20) yields

$$V_0(N) - V_0(N+1) = \frac{1}{V_*(N)} \left[K_{*0} - \sum_{i=1}^{N} \lambda_{i*} K_{i0} - \sum_{p=1}^{l(k)} \mu_{p*} f_p(X_0) \right]^2 \tag{21}$$

Equation (21) can be defined as

$$VR_{0*} = V_0(N) - V_0(N+1) \tag{22a}$$

as the "variance reduction VR" at X_0 due to a measurement at X_*. VR_{0*} can be utilized as a direct measure of the improvement in the reliability of kriging estimates due to sampling at a new location. Equation (22) can be expanded to other versions of kriging. For instance, in universal kriging one can write (21) as

$$VR_{0*} = \frac{1}{V_*(N)} \left[\gamma_{*0} - \sum_{i=1}^{N} \lambda_{i*} \gamma_{i0} - \sum_{p=1}^{l} \mu_{p*} f_p(X_0) \right]^2 \tag{22b}$$

where γ_{ij} is the semivariogram of $Z(X_i)$ and $Z(X_j)$.

If simple covariance functions are used, then (21) can be written as

$$VR_{0*} = \frac{1}{V_*(N)} \left[C_{*0} - \sum_{i=1}^{N} \lambda_{i*} C_{i0} - \sum_{p=1}^{l} \mu_{p*} f_p(X_0) \right]^2 \tag{23}$$

where C_{ij} is the covariance of $Z(X_i)$ and $Z(X_j)$.

VR_{0*} (21) is dependent only on the covariance function and the geometry of the points. Thus it is a suitable tool for the design and planning of data collection schemes. The other advantage of VR_{0*} is due to its computational efficiency. For calculation of (21) there is no need to solve another kriging system or invert another matrix A_* for each possible additional sampling site.

It is also easy to see that VR_{0*} (21) is always positive. It implies that any new sampling would cause only reductions in kriging variances. In other words, any additional sampling would improve the reliability of the estimated field. This optimistic conclusion may sound logical; however, it implicitly assumes that the additional data would not significantly affect the assumed covariance function. In the real world, the new measured values are sometimes so different from the estimated values that drastic changes in the covariance function may result. In such cases, there might be a need for reevaluating the assumed covariance function which could force the kriging variances to go up (i.e., the actual VR becomes negative)! It can be concluded that as long as the assumed covariance function remains intact, VR_{0*} (21) is a valid measure of the improvement of the reliability of the estimated field.

3. Data Management Approaches

Prior to any sampling design, one should establish the objective of the study in order to deal with the question of data collection. There are two major approaches which are commonly used in groundwater data management studies. In the

first method, sampling procedures are designed based on maximization of the accuracy of the estimated field with budget constraints or by minimizing the sampling cost subject to a criterion of minimal acceptable accuracy. Such programs are suitable for regional studies, where the errors in data cannot be easily related to any monetary measure besides measurement costs. The level of accuracy of variables has to be substituted for more common economic criteria such as economic benefits. These methods usually lack a meaningful interpretation of the optimal level of accuracy of the data. They will not tell the planner how much is gained by adding a new data point.

On the other hand, in the second approach the accuracy of parameters is interpreted in economic terms. This approach is easily applicable to problems dealing with specific planning and management activities. Such programs yield more meaningful measures for optimal data management plans.

3.1. *Previous Work*

Many authors have advocated the use of the accuracy of the estimates as the criterion function for their proposed sampling schemes. For instance, *Fiering* [1965] and *Matalas* [1968] use the total variance of estimates as the objective function of their schemes for gaging. By using a nonlinear integer programming they identify the best locations for sampling among a set of potential sites that yield minimum total variance. This approach is operationally slow and inefficient. *Bastin et al.* [1984] compute all possible combinations of n sampling sites out of m potential locations in order to identify the subset that produces minimum normalized kriging variance. This method becomes costly as the number of combinations increases. *Hughes and Lettenmaier* [1981] and *Chou and Scheck* [1984] use iterative algorithms to adjust the location of sampling sites in order to minimize regional or areal kriging variances. In these works there is no need to specify the potential sites. However, the efficiency of the iterative algorithm depends on the assumed initial locations. In the case of *Chou and Scheck* [1984], the minimization of the regional kriging variance is a nonlinear programming problem subject to constraints that may become operationally inefficient as the number of sites increases. Other such as *Pimental* [1978] include the accuracy of their results as a constraint in the form of a maximum allowable variance of estimation. The objective functions in such cases are sampling rates or costs.

Examples of the second type of approach can be found in the work by *Maddock* [1973], where the accuracy of the data is associated with the mean expected loss in total farm income. These models usually identify the optimal sampling rates or sites by comparing marginal benefits of additional data to measurement costs. *Maddock* [1973] also proposes a method to rank different types of data based on their relative influence on the risk function.

In many cases of the second approach, the expected loss value remains almost unaffected by variations in hydrological parameters. *Maddock* [1973] concludes that the value of the risk is practically insensitive to changes in the value of transmissivities and storativities, yet it is highly dependent on crop prices and pumping costs. Similar results are also reported by *Ben-Zvi and Bachmat* [1979]. Therefore in spite of the fact that the second approach gives a meaningful interpretation to the accuracy of data, it fails to give a significant role of hydrological parameters.

One reason for such behavior lies in the fact that in the second type of models the geohydrological parameters are linked to the economic functions through the groundwater level. The depth of the water table always plays a major role in the total (farm) income functions; for example, it is the main factor in the pumping cost functions. In turn, the piezometric head is calculated through the flow equation in aquifers, as a function of transmissivity and storativity values. Such head estimates show significant level of robustness with respect to variations in transmissivity and storativity values. As *Fogg et al.* [1979] notice, radical changes in transmissivity values are reflected by only scarcely perceptible changes in head. This problem usually leads to identification instability in inverse problems [see *Neuman and Yakowitz*, 1979]. Moreover, *Bakr et al.* [1979] showed that spatially varying transmissivities in a three-dimensional space results only in small head variances. Consequently, the economic risk or loss functions which are dependent on piezometric head values also show little or no sensitivity to the variations in the values of transmissivity and storativity. So it seems appropriate that in a combined hydroeconomic approach to data management problems, groundwater levels should be considered as an independent variable rather than a function of other hydrological parameters. This leads us to the study of the third approach.

3.2. *A Third Approach*

Each of the above methods has deficiencies. The first approach puts heavy emphasis on the accuracy of results but fails to interpret them in a meaningful manner. On the other hand, the second approach provides an economic interpretation for the accuracy levels but appears to ignore the hydrological data. In order to solve this problem, *Bras and Rodriguez-Iturbe* [1976] propose the use of a weighted sum of the accuracy and the cost of observation as the objective function of their data management program. However, the relative weight of these two factors, the so-called "trade-off" coefficient, remains a subjective measure.

The solution to the optimal data management can be inferred to lie in a proper link between the economic risks and the accuracy of the hydrological parameters, particularly the groundwater levels. One can build this link by defining the monetary losses associated with uncertainties in water levels. The kriging variance can be utilized as a measure of accuracy of the estimates. Our objective is to define the expected losses in terms of Var $[\hat{Z} - Z]$. Ultimately, by using the variance reduction analysis, one can estimate the reduction in expected losses due to the addition of a new data point.

Such a loss function can be defined in terms of over or underestimation of $\hat{Z}$ (e.g., piezometric head estimates). For example, whenever $\hat{Z} - Z$ is positive (i.e., the estimated piezometric head is higher than the actual one), the operators are faced with a penalty. These losses may be in the form of higher costs of pumping. However, if the estimation results turn out to be underestimating the water table, the operators may have to pay other forms of penalties such as higher drainage costs. It can be argued that these marginal losses may not be equal, and thus the loss function is asymmetric. Moreover, this function may have a shape similar to the pumping cost functions. Concerning the overall cost of pumping, *Bredehoeft and Young* [1972] and *Maddock* [1973] both assumed that the cost of water production is a linear function of the depth of the water table.

Considering all the above factors, one can define a loss function as follows:

$$\begin{aligned} L &= C_U(\hat{Z} - Z) \qquad \hat{Z} - Z \le 0 \\ L &= C_O(\hat{Z} - Z) \qquad \hat{Z} - Z \ge 0 \end{aligned} \tag{24}$$

where

L loss function (dollars);
$\hat{Z}$ estimated piezometric head (ft);
Z actual piezometric head (ft);
C_U loss per foot of underestimation (dollars/ft);
C_O loss per foot of overestimation (dollars/ft).

It must be mentioned that the role of this loss function is simply to interpret the level of accuracy of $\hat{Z}$ in monetary terms. Our estimation criteria remains to be unbiasedness and minimum squared error as defined by (2). One, however, can utilize a similar function to (24) as a basis for estimation of $\hat{Z}$. The corresponding estimate is no longer given by kriging or more generally by a conditional expectation type estimator, but by a conditional quantile estimator [see *Journel*, 1984].

In order to evaluate the expected losses one must make some assumptions about the statistical nature of the estimation errors. *Freeze* [1975] ran an extensive Monte Carlo simulation of water heads in a one-dimensional flow based on uncorrelated lognormally distributed transmissivities. He concluded that the steady state system with low estimation variance tends toward a normal frequency distribution for piezometric heads over a greater portion of the field. It seems reasonable to assume that estimation fluctuations (i.e., $\hat{Z} - Z$) are normally distributed, with a zero mean and a variance equal to the so-called kriging variance. The expected loss can be written as follows:

$$E(L) = \int_{-\infty}^{\infty} L(u) f(u)\, du \tag{25}$$

where

$L(u)$ loss function;
$f(u)$ frequency distribution of the estimation error;
$u = \hat{Z} - Z$.

Following *Bryant* [1961], substituting a normal frequency in (25) and using the loss function described in (24) we can write the expected loss as

$$\begin{aligned} E(L) &= \int_{-\infty}^{0} C_U u (2\pi V)^{-1/2} \exp(-u^2/2V)\, du \\ &\quad + \int_{0}^{+\infty} C_O u (2\pi V)^{-1/2} \exp(-u^2/2V)\, du \\ &= \frac{C_U + C_O}{(2\pi)^{1/2}} V^{1/2} = c V^{1/2} \end{aligned} \tag{26}$$

where V is the kriging variance $E[\hat{Z} - Z]^2$ (ft^2), and c is the net loss coefficient $(C_O + C_U)/(2\pi)^{1/2}$ (dollars/ft).

Equation (26) shows the expected losses at each estimated point as a function of the kriging variance at that site. Expected loss (26) is based on a rather simple distribution function. One may estimate the expected losses based on more sophisticated conditional probability of $\hat{Z} - Z$ fluctuations, given the N existing data points, which is likely to be much more complex than (26) [see *Journel*, 1984]. Now, using (26) the total expected losses (TEL) prior to any new sampling can be defined as

$$TEL = \sum_j E(L_j) = c_j \sum V_j^{1/2} \tag{27}$$

where $E(L_j)$ is the expected loss associated with V_j, which is the kriging variance at X_j. Adding a new data point reduces variances of estimations. As defined in (21), VR_{ji} is the amount of such reductions in V_j due to a new measurement at X_i. The TEL after a new sampling at X_i can be written as

$$TEL_i = c \sum_j (V_j - VR_{ji})^{1/2} \tag{28}$$

Thus the total loss reduction (TLR) due to an additional measurement at X_i is

$$\begin{aligned} TLR_i &= TEL - TEL_i \\ &= c\left[\sum_j V_j^{1/2} - \sum_j (V_j - VR_{ji})^{1/2}\right] \end{aligned} \tag{29}$$

These loss reductions aside, adding a new measurement requires more investments. The net expected benefit (NEB) of a new data point is defined as

$$NEB_i = TLR_i - MC_i \tag{30}$$

where

NEB_i net expected benefit of sampling at X_i;
TLR_i total loss reduction due to sampling at X_i;
MC_i measurement cost at X_i.

The above results can be shown in a different way. The following can be defined as

$$\begin{aligned} TOTV &= \sum_j V_j \\ TOTSD &= \sum_j V_j^{1/2} \end{aligned} \tag{31}$$

where $TOTV$ is the total sum of kriging variances, and $TOTSD$ is the total sum of kriging standard deviations. $TOTV_i$ and $TOTSD_i$ are defined as $TOTV$ and $TOTSD$ after the addition of the new data point at X_i. TVR_i (total variance reduction due to sampling at X_i) can be written as

$$\begin{aligned} TVR_i &= TOTV - TOTV_i \\ &= \sum_j V_j - \sum_j (V_j - VR_{ji}) = \sum_j VR_{ji} \end{aligned} \tag{32}$$

TVR_i represents the total gain in accuracy or the information gain due to the measurement at X_i [*Matalas*, 1968].

Similarly, $TSDR_i$ (total standard deviation reduction due to sampling at X_i) is defined as

$$\begin{aligned} TSDR_i &= TOTSD - TOTSD_i \\ &= \sum_j V_j^{1/2} - \sum_j (V_j - VR_{ji})^{1/2} \end{aligned} \tag{33}$$

Substituting the above into (27), one gets

$$TLR_i = c(TSDR_i) \tag{34}$$

In other words, $TSDR_i$ reflects the economic gain due to a new measurement at X_i, while TLR_i represents the monetary value of added information. NEB_i (30) can also be shown as

$$NEB_i = cTSDR_i - MC_i \tag{35}$$

If the cost of measurement exceeds the economic gain of the added information, the result is as follows:

$$TSDR_i < MC_i/c \tag{36}$$

In the process of variance reduction analysis, all points where (36) holds should be eliminated as potential new measurement sites.

3.3. *Ranking of Prospective Data Points*

Equation (35) can be utilized in two ways. First, all points that show negative NEB_i can be eliminated as potential data

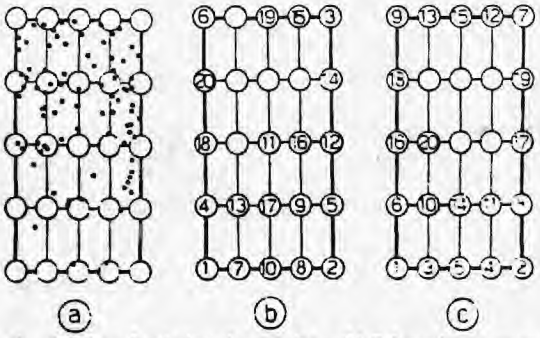


Fig. 1. (*a*) Set of existing data points. (*b*) Selected sequence based on the variance reduction analysis. (*c*) Selected sequence based on the criterion of maximum distance. (Numbers in circles correspond to the rank of the selected sites.)

locations. Second, the sites with positive NEB_i can be ranked as a sequence of points for further sampling. A set of weights may be assigned to potential sites to reflect their relative importance. It makes this ranking procedure more flexible for different cases of data management.

The above ranking is valid as long as only one additional data point is involved. This means that the reduction in V_j caused by a set of new data points is not equal to the sum of the corresponding VR_{ji}. For example, if two points are closely located, measurement at one will reduce the effectiveness of the other as a new sampling point. For an efficient data management scheme, the best feasible algorithm seems to be the following.

1. Perform the kriging and calculate VR_{ji}.
2. Pick X_i with maximum TVR. If NEB_i was negative, then stop.
3. If not, assume that X_i is a data point. Then go to step 1.

In this process X_i is selected based on maximum information gain. The number of added points, however, depends on the economic gain function. In fact, when the net benefit of the added point becomes negative, the planner should stop sampling.

For the purposes of kriging a computer package named AKRIP (an acronym for a kriging program) is utilized. AKRIP is a kriging algorithm for *IRF* of order 0, 1, and 2 developed by *Kafritsas and Bras* [1981]. This program is based on the proposed algorithm by *Delfiner* [1975]. It includes a step-by-step structural analysis which is the core of any kriging procedure. It also provides options for point or block kriging.

In the variance reduction (VR) equation (21) it is assumed that the set of neighboring data points for both X_0 and X_* are identical. However, in AKRIP, each point may be interpolated by different sets of "nearest" measured values. X_* may be located within the radius of nearest data points to X_0 but has a different set of neighboring points for itself. In such a case, X_* represents the added information about its surrounding area. Thus it seems appropriate to use its neighboring data points in the VR analysis. This way the calculated VR_{0*} shows the impact of the addition of X_* which is currently estimated by its neighboring measured values.

4. Application of Variance Reduction Analysis

4.1. *Data Description*

The available data are groundwater level observations made in January 1979 in Groundwater Management District no. 4 of Kansas, an area of nearly 5000 square miles in northwestern Kansas, including Sherman, Thomas, and Sheridan counties and parts of Cheyenne, Rawlins, Decatur, Graham, Logan, and Gove counties. The data set consists of 327 measurements made in water wells scattered at irregular locations within the district and outside but close to its boundaries. Average spacing between wells is about 3.6 miles (5.8 km). The measurements define a water surface that forms an undulating plane dipping to the east and northeast. For further study of the geohydrology of this region, readers are referred to *Pearl et al.* [1972].

An area of 2048 square miles (5302 square km) is selected, as indicated in Figure 1*a*. This subregion lies between latitudes 38°48′ and 39°48′ north and longitudes 101° and 101°36′ west. There are 84 measurement points in this area (see Table 1). Northeastern and northwestern corners of this zone are rather densely measured, while central and southern parts of this subregion have relatively scattered data points.

4.2. *Summary of the Numerical Results*

Based on the variance reduction analysis, a ranking of the prospective new measurement sites has been conducted. For

TABLE 1. Existing Data Locations and Values

Point	*y*, miles Down	*x*, miles Across	$Z(x, y)$, ft	Point	*y*, miles Down	*x*, miles Across	$Z(x, y)$, ft
(1)	1.18	4.16	3239.00	(43)	32.02	7.35	3356.00
(2)	1.93	7.08	3196.00	(44)	32.27	7.84	3349.00
(3)	3.31	9.95	3175.00	(45)	33.16	6.81	3372.00
(4)	6.06	8.43	3205.00	(46)	31.04	0.70	3445.00
(5)	4.18	0.33	3295.00	(47)	31.16	5.08	3391.00
(6)	5.94	2.11	3292.00	(48)	32.54	1.84	3443.00
(7)	7.44	0.81	3308.00	(49)	36.55	4.33	3433.00
(8)	12.58	7.62	3244.00	(50)	43.17	9.84	3228.00
(9)	12.20	3.46	3312.00	(51)	17.10	29.52	3039.00
(10)	13.33	1.57	3339.00	(52)	21.61	29.79	3050.00
(11)	14.83	2.16	3341.00	(53)	17.85	27.41	3060.00
(12)	48.94	26.87	3099.00	(54)	17.98	27.90	3045.00
(13)	46.55	13.52	3219.00	(55)	18.74	24.54	3101.00
(14)	46.67	13.14	3216.00	(56)	16.97	21.90	3112.00
(15)	61.59	22.49	2912.00	(57)	20.36	22.00	3146.00
(16)	2.18	26.65	2886.00	(58)	16.47	11.52	3219.00
(17)	3.43	24.38	2926.00	(59)	17.46	15.95	3189.00
(18)	1.55	21.35	2973.00	(60)	20.86	11.89	3263.00
(19)	0.53	14.27	3128.00	(61)	21.85	15.19	3229.00
(20)	0.92	11.25	3161.00	(62)	23.37	29.90	3041.00
(21)	6.82	29.30	2981.00	(63)	23.13	24.49	3124.00
(22)	8.82	31.03	2972.00	(64)	26.12	29.41	3055.00
(23)	9.70	24.33	3039.00	(65)	26.75	28.11	3065.00
(24)	4.92	16.92	3092.00	(66)	24.00	20.49	3169.00
(25)	6.69	22.22	2986.00	(67)	24.99	12.38	3273.00
(26)	8.94	18.16	3068.00	(68)	27.63	15.41	3252.00
(27)	4.06	10.22	3176.00	(69)	29.13	30.28	3041.00
(28)	5.93	12.60	3172.00	(70)	28.26	28.76	3056.00
(29)	14.34	29.73	3023.00	(71)	28.77	23.52	3124.00
(30)	10.08	27.36	3021.00	(72)	31.77	27.63	3062.00
(31)	14.47	25.09	3072.00	(73)	30.13	12.00	3302.00
(32)	11.95	19.68	3099.00	(74)	38.65	29.84	3017.00
(33)	13.46	12.33	3200.00	(75)	35.39	29.36	3043.00
(34)	16.47	8.38	3261.00	(76)	43.79	29.57	3057.00
(35)	17.09	5.51	3294.00	(77)	40.15	29.09	3015.00
(36)	18.97	7.79	3288.00	(78)	41.54	28.60	3038.00
(37)	17.22	2.00	3354.00	(79)	42.55	27.46	3076.00
(38)	19.48	0.87	3391.00	(80)	45.59	16.87	3189.00
(39)	22.24	7.41	3318.00	(81)	52.10	4.60	3229.00
(40)	22.86	6.38	3336.00	(82)	61.94	13.46	3011.00
(41)	23.36	9.89	3290.00	(83)	39.69	19.30	3173.00
(42)	24.37	6.76	3337.00	(84)	25.75	32.06	3026.00

One mile = 1.609 km; 1 foot = 0.3048 m.

the purpose of *VR* analysis the field is divided into a 5 × 5 grid with $\Delta x = 8$ miles (12.9 km) and $\Delta y = 16$ miles (25.7 km). The nodes are described as the set of potential sampling sites. At each round of kriging, the point with maximum *TVR* is selected as the new added data point. The basis of this selection is the maximization of the added information. It is further assumed that the new measurements do not cause any change in the parameters of the selected covariance function. So, in the process of data collection, no further structural analysis is conducted.

The structural analysis of the initial data set indicated an *IRF* of the first order with the following covariance function:

$$K(h) = 145.68\delta(h) + 0.89914h^3 \qquad (37)$$

where δ is Dirac's delta function, and h is the length of vector distance between two points in miles. It must be mentioned that the above covariance function (37) has a tendency toward pure nugget effect. In cases of pure nugget effect, kriging is reduced to a moving average process.

Using the newly added *VR* option to *AKRIP*, the top 20 points have been ranked as the sequence of best locations for further measurements. The actual results and the effects of each added data point on the overall reliability of the estimated field are shown in Table 2.

Figure 1*b* illustrates the spatial distribution of the ranked sites. As is expected, most of the added points are in the lower section of the field which has few existing sampling sites. For example, eight of the top 10 points are in the southern part of the region. In contrast, the central region of the upper section which was already densely measured does not gain any new data point among the top 20.

Another look at Figure 1*b* shows that almost all nodes on the border lines are selected as sites for further sampling. One hundred percent of the top 5 and 90% of the top 10 points are boundary nodes. Among the top 20, 15 points are located on the edge of the field; this is 94% of all possible boundary nodes. Meanwhile, the internal grid points get five sites, which is only 56% of the total available internal nodes. In statistical terms, the boundary nodes are extrapolated, while the interior points are usually interpolated. The extrapolated nodes are less reliable than the interpolated ones. In other words, given equal weights to each point, the boundary nodes are predominant choices for further measurements. This conclusion can also be verified by comparing it to the case of stochastic steady state flow in aquifers. In such instances, the variations of the boundary values of piezometric heads are the most influential factors on the variances of the estimated water tables [see *Dettinger and Wilson*, 1979]. It must be noted that in this example all data points outside the area of study are ignored. The addition of outside sampling sites near to the boundary may reduce the priority of border nodes in the ranking process.

TABLE 2. Results of Sampling Based on Variance Reduction Analysis

Rank	x, miles Across	y, miles Down	Z, ft	V, ft²	TVR, ft²	TOTV, ft²	TOTSD, ft
0	...	...	...	...	...	27,019	669.4
1	0	64	3119	8311	9644	17,445	559.3
2	32	64	2775	6415	7104	10,400	467.4
3	32	0	2825	1600	1634	8,770	426.5
4	0	48	3347	1120	1163	7,643	393.1
5	32	48	3020	714	736	6,915	366.0
6	0	0	3281	711	718	6,197	339.2
7	8	64	3028	640	670	5,521	313.0
8	24	64	2850	452	496	5,025	290.6
9	24	48	3123	344	406	4,642	270.9
10	16	64	2940	403	403	4,265	251.6
11	16	32	3244	365	380	3,895	232.4
12	32	32	3015	368	379	3,532	213.4
13	8	48	3235	351	353	3,178	194.5
14	32	16	3004	348	349	2,830	175.9
15	24	0	2906	317	325	2,507	157.9
16	24	32	3114	306	306	2,201	140.4
17	16	48	3189	288	288	1,912	123.4
18	0	32	3463	280	280	1,632	106.7
19	16	0	3087	272	273	1,360	90.2
20	0	16	3379	267	267	1,093	73.8

One mile = 1.609 km, 1 ft = 0.3048 m, and 1 ft² = 0.0929 m².

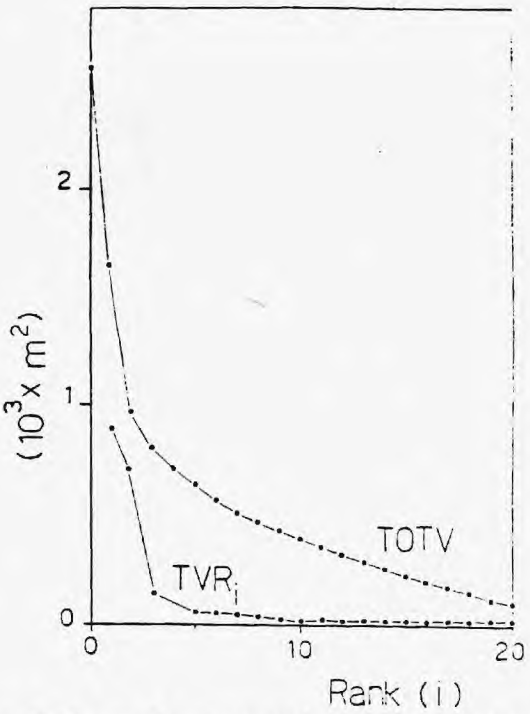


Fig. 2*a*. Total sum of variances and the corresponding marginal information gains due to additional sampling.

Figures 2*a* and 2*b* represent total variances and standard deviations at each round of kriging. They also show the corresponding marginal improvements in the accuracy of the estimated field due to the addition of each new point. As ex-

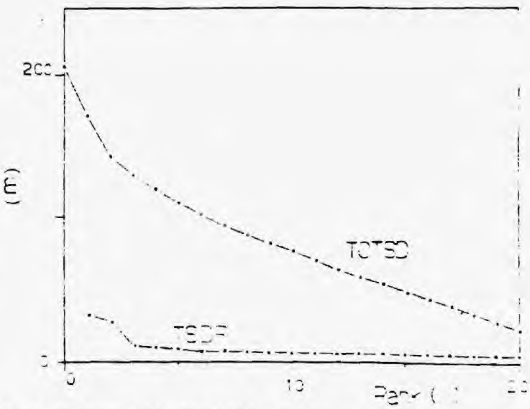


Fig. 2*b*. Total sum of standard deviations and the corresponding marginal economic gains due to additional sampling.

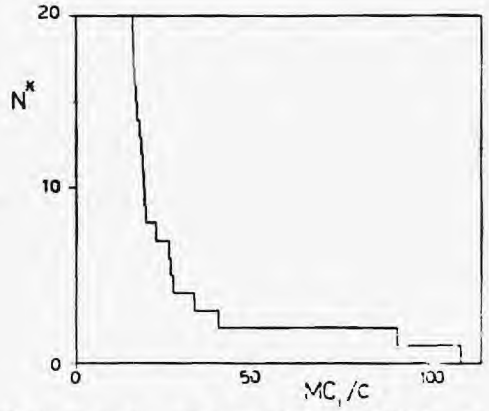


Fig. 3. Optimal number of new data points (N^*) versus cost-loss ratio (MC_i/c).

pected, both *TOTV* and *TOTSD* decrease as the number of new data points increase. These optimistic results are valid as long as the estimated generalized covariance function remains unchanged.

TVR and *TSDR* show the level of reduction in *TOTV* and *TOTSD* (i.e., improvement in the reliability of results) at each round of kriging. In initial rounds *TVR* and *TDSR* are quite high, but after few rounds they both approach almost asymptotic levels. This monotonic decrease in the values of *TVR* and *TSDR* is similar to the concept of "diminishing rate of return" in economics. As the number of new sites increases, the marginal improvement caused by additional measurements decreases. So there should be a finite optimal number of new measurements (N^*).

As is mentioned in (35), the net expected benefit of each sampling is a linear function of *TSDR* and the corresponding measurement costs (*MC*). As long as marginal benefits exceed measurement costs, new data points should be added, given no budget constraint. N^* reaches its optimal value when equilibrium is established. In neoclassical economics, equilibrium is referred to the state of equality between marginal benefits and costs [see *Hirshleifer*, 1976]. Naturally, no further measurement should be conducted beyond the state of equilibrium. In terms of (35), N^* is equal to i if and only if

$$\begin{aligned} TSDR_j &< MC_j/c \qquad j > i \\ TSDR_j &\geq MC_j/c \qquad j = i \\ TSDR_j &> MC_j/c \qquad j < i \end{aligned} \tag{38}$$

The locus of points where equilibrium (38) hold is shown in Figure 3 (N^* as a function of MC_i/c). The shape of the graph indicates that N^* is extremely sensitive with respect to small values of MC_i/c (e.g., less than 20). However, as MC_i/c increases, N^* becomes significantly less sensitive to the value of MC_i/c. The following are some approximate estimates of the sensitivity of N^* with respect to MC_i/c (i.e., $\hat{c}N^*/\hat{\partial}(MC_i/c)$):

	$\left[\frac{\hat{c}\hat{N}^*}{\partial MC_i/c}\right]$
$MC_i < 20c$	-2.09
$20c < MC_i < 50c$	-0.26
$50c < MC_i$	-0.05

The above results indicate that as MC_i/c decreases, the sensitivity of N^* to its values goes up as much as 40 times. It can be inferred that for small MC_i/c, the high reliability of its estimates is essential to identify a robust estimate of N^*. In other words, as MC_i/c goes down, the relative importance of economic data (i.e., MC_i and c) increases dramatically. In contrast, even an approximate estimate of large MC_i/c produces a robust N^*.

4.3. *Resilience of Variance Reduction Analysis*

Variance reduction analysis depends only on the covariance function and the geometry of points. This property in turn implies that the new added values should comply with the estimated covariance function. The assumption of the constancy of the covariance function is the basis of the optimality of the variance reduction decisions (i.e., the selected sequence).

The questions that immediately arise are, What is the effect of the predicted values being significantly under or overestimated? How does the generalized covariance function respond to fluctuations in the newly sampled values? How does the decisions based on the variance reduction analysis behave under such situations? In order to answer these questions it seems necessary to test the robustness and the resilience of the variance reduction analysis.

For the purpose of answering these questions the *VR* analysis was divided into three spaces. The data set was denoted as the input space. The results of the structural analysis (i.e., the estimated parameters of the covariance function) were defined as the parameter space, and the selected sets were represented as the action space.

At each round of kriging the value of Z at the selected site, X_*, was defined as

$$Z(X_*) = \hat{Z}(X_*) \pm \varepsilon$$

where

$Z(X_*)$ simulated measured value at X_*;
$\hat{Z}(X_*)$ estimated value at X_*;
ε simulated normally distributed noise; $N(0, \text{Var}[\hat{Z} - Z])$.

$Z(X_*)$ was then added to the input space, which was followed by a structural analysis. The parameter space showed a significant degree of instability with respect to the noisy input space. Similar results are also reported by *Kitanidis* [1983]. Even small levels of fluctuations in the added data caused large changes in parameters of the estimated generalized covariance function. In contrast, the instability of the parameter space had a negligible effect on the action space. Selected sets under noisy input space showed strong similarities with the original selected sequence. For actual results, readers are referred to *Rouhani* [1983].

4.4. *Comparison of Sampling Based on Simple Criteria to Variance Reduction Analysis*

Figure 1*c* shows the top 20 selected nodes based on a simple criterion of maximum distance from data set. At each round, the point among the potential sites with maximum average distance from existing data points is selected as the new added measurement site. Marginal economic and information gains (*TSDR* and *TVR*) due to sampling based on this simple criteria is then calculated. The comparison of this set to the selected sequence based on *VR* analysis reveals a significant difference between them. The above simple criterion ignores the fact that points with maximum distances from existing data set are not necessarily the most uncertain nodes. For example, in some cases, the node with maximum distance is

located very close to a single data point. So, even though it is the farthest point, it may not be the most uncertain site.

Following such a simple criterion may lead to information and economic regrets. These regrets are defined as reductions in marginal information and economic gains caused by using the maximum distance criterion as compared to gains of the selected sequence based on *VR* analysis.

Figures 4*a* and 4*b* show information and economic regrets due to sampling based on the maximum distance criterion. As seen from these figures, the regrets are significant. They also show an interesting pattern. At the first stage of ranking of the top ten points, all points are bordered nodes in sparsely sampled areas. At this stage the simple method leads the planner to nonoptimal points with large regrets, such as the fifth-ranked node. After some lags it identifies the optimal points and the level of regret goes down. In the second stage (i.e., the second top ten points) selected nodes are located in relatively densely sampled areas; almost half of them are internal points. Both regret functions at this stage show a second rise in regret values; however, the level of regrets are smaller. From the above it can be concluded that simple criteria can lead the planners to nonoptimal points in both sparsely and densely sampled regions with significant regrets.

5. Summary and Conclusions

This paper is an attempt to develop a data collection algorithm, known as the variance reduction analysis. The proposed method is based on an information response function (i.e., the amount of information gain at an arbitrary point due to a measurement at another site). This method was later applied to groundwater data management problems. Total variance reduction (a measure for the information gain) which is independent of measured values was used as a tool for the design and planning of such data collection schemes. However, this algorithm still required an additional measure for the monetary or economical interpretation of the gained information. It was suggested that by utilizing a loss function the planners can estimate monetary values of their added data. By assuming that measurement fluctuations are normally distributed, a two-piece linear loss function yielded an expected loss that was directly proportional to the square root of the kriging variance. This measure of economic gain provided a monetary interpretation for the value of the added information.

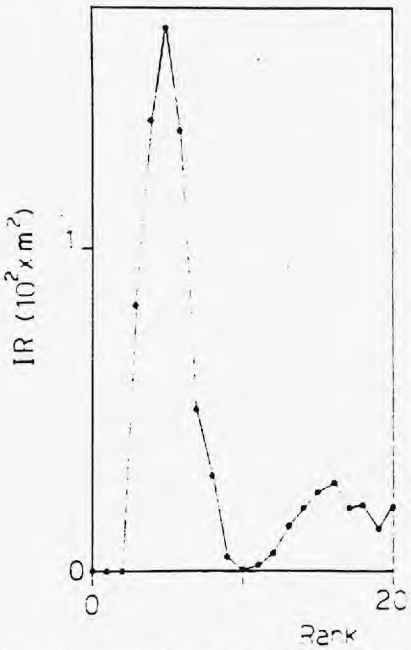


Fig. 4*a*. Information regrets due to sampling based on the criterion of maximum distance.

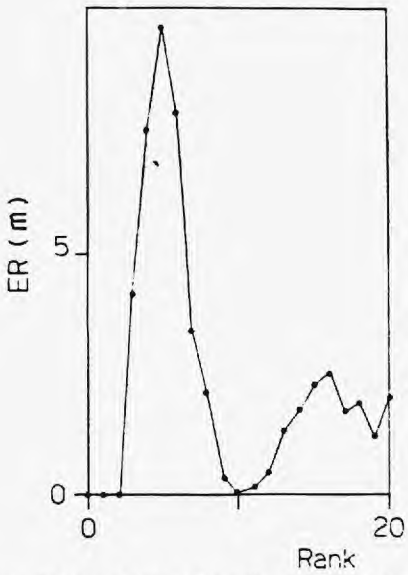


Fig. 4*b*. Economic regrets due to sampling based on the criterion of maximum distance.

The two indicators of information and economic gains led to an optimal sampling scheme. Based on the information gain function, selected points were ranked as prospective new sampling sites. Then, using the economic gain function, the optimal number of added points was calculated as a function of cost-loss ratio (i.e., MC_i/c). Studying the pattern of selected points produced the following conclusions:

1. Given equal weights to all nodes, border (extrapolated) nodes have a higher priority over internal (interpolated) points.
2. Areas with low sampling density get a clear priority for further measurements.
3. Marginal information and economic gains diminish to almost asymptotic values as the number of added points increases.
4. As the measurement cost increases relative to net loss coefficient (i.e., MC_i/c goes up), the number of optimal new points (N^*) decreases.
5. When MC_i/c is small the sensitivity of N^* with respect to the cost/loss ratio is far greater than the case of larger MC_i/c. So at low MC_i/c more accurate economic data is needed in order to produce equally robust estimates of N^*.
6. Adding noisy input shows that the proposed structural analysis yields estimates of covariance function parameters that lack robustness. However, selected sets show significant stability under noisy inputs.
7. Simple criteria for the selection of sampling sites such as "the maximum distance from data set," tend to ignore the influence of added data on their neighboring points. As a result, they lead the planners to sampling at nonoptimal sites.

In general, it can be concluded that the variance reduction analysis is an effective algorithm for the planning and design of data collection schemes in random fields.

Acknowledgments. This research is based on my Ph.D. thesis work at the Division of Applied Sciences, Harvard University, which was supported in part by a grant from the Ford Foundation to Harvard University. Many thanks go to my thesis advisors P. P. Rogers and

M. B. Fiering for their support, guidance, and friendship. Thanks also to L. LaSalata for an excellent job typing the manuscript.

REFERENCES

Aboufirassi, M., and M. A. Marino, Kriging of water levels in the Sauss Aquifer, Morocco, *Math Geol.*, *15*(4), 537–551, 1983.

Bakr, A. A., L. W. Gelhar, A. L. Gutjahr, and J. R. MacMillan, Stochastic analysis of spatial variability in subsurface flows, 1, Comparison of one- and three-dimensional flows, *Water Resour. Res.*, *14*(2), 263–271, 1978.

Bastin, G., B. Lorent, C. Duque, and M. Gevers, Optimal estimation of the average rainfall and optimal selection of rain gauge locations, *Water Resour. Res.*, *20*(4), 463–470, 1984.

Ben-Zvi, M., and Y. Bachmat, Management of a water resource under uncertainty, technical report, Hydrol. Serv., Jerusalem, Israel, 1979.

Bras, R. L., and I. Rodriguez-Iturbe, Network design for the estimation of areal mean of rainfall events, *Water Resour. Res.*, *12*(6), 1185–1208, 1976.

Bredehoeft, J. D., and R. A. Young, The temporal allocation of groundwater—A simulation approach, *Reprint 103*, Resour. For the Future, Washington, D. C., 1972.

Bryant, G. T., Stochastic theory of queues applied to design of impounding reservoirs, Ph.D. thesis, Div. of Appl. Sci., Harvard Univ., Cambridge, Mass., 1961.

Chirlin, G. R., and G. Dagan, Theoretical head variogram for steady flow in statistically homogeneous aquifers, *Water Resour. Res.*, *16*(6), 1001–1015, 1980.

Chou, D., and D. E. Scheck, Selecting optimum drilling locations by groups, paper presented at the SME-AIME Annual Meeting, Soc. of Mech. Eng. and Am. Inst. of Min. and Eng., Los Angeles, Calif., February 26–March 1, 1984.

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

David, M., *Geostatistical Ore Reserve Estimation*, Elsevier Scientific, New York, 1977.

Delfiner, P., Linear estimation of non-stationary spatial phenomena, *Advanced Geostatistics in the Mining Industry*, edited by M. Guarascio, M. David, and C. Huijbregts, Nato Advanced Study Institute, D. Reidel, Hingham, Mass., 1975.

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

DeMarsily, G., The estimation of parameters (e.g., transmissivity) in an aquifer: A stochastic approach, paper presented at Ralph M. Parsons Laboratory, Mass. Inst. Technol., Cambridge, Mass., 1979.

Dettinger, M. D., and J. L. Wilson, Numerical modeling of aquifer systems under uncertainty: A second moment analysis, technical report, M.I.T./Cairo Univ. Technol. Plann. Program, Dep. of Civ. Eng., Mass. Inst. Technol., Mass., May 1979.

Dunlap, L. E., and J. M. Spinazola, Interpolating water-table altitudes in West-Central Kansas, *U.S. Geol. Surv., Open File Rep., 81-1062*, 1981.

Fiering, M. B., An optimization scheme for gaging, *Water Resour. Res.*, *1*(4), 463–470, 1965.

Fogg, G., E. Simpson, and S. P. Neuman, Aquifer modeling applied to an Arizona groundwater basin, *Tech. Rep. 32*, Dep. of Hydrol. and Water Res., Univ. of Ariz., Tucson, 1979.

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

Gambolati, G., and G. Volpi, Groundwater contour mapping in Venice by stochastic interpolators, 1, Theory, *Water Resour. Res.*, *15*(2), 281–297, 1979.

Hirshleifer, J., *Price Theory and Applications*, Prentice-Hall, Englewood Cliffs, N. J., 1976.

Hughes, J. P., and D. L. Lettenmaier, Data requirements for kriging: Estimation and network design, *Water Resour. Res.*, *17*(6), 1641–1650, 1981.

Journel, A. G., mAD and conditional quantile estimators, in *Geostatistics for Natural Resources Characterization*, edited by G. Verly, M. David, A. G. Journel, and A. Marechal, D. Reidel, Hingham, Mass., 1984.

Kafritsas, J., and R. L. Bras, The practice of kriging, *Rep. 263*, Ralph M. Parsons Lab., Mass. Inst. Technol., Cambridge, Mass., January 1981.

Kitanidis, P. K., Statistical estimation of polynomial generalized covariance functions and hydrologic applications, *Water Resour. Res.*, *19*(4), 909–921, 1983.

Kitanidis, P. K., and E. G. Vomvoris, A geostatistical approach to the inverse problem in groundwater modeling (steady state), 1, technical report, Inst. of Hydraul. Res., Univ. of Iowa, Iowa City, August 1982.

Maddock, T., III, Management model as a tool for studying the worth of data, *Water Resources Res.*, *9*(2), 270–280, 1973.

Matalas, N. C., Optimum gaging station location, paper presented at Proceedings at IBM Scientific Computing Symposium, Water and Air Resources Management, IBM, White Plains, N. Y., 1968.

Matheron, G., The intrinsic random function and their applications, *Adv. Appl. Prob.*, *5*, 439–468, 1973.

Neuman, S. P., and S. Yakowitz, A statistical approach to the inverse problem of aquifer hydrology, 1, Theory, *Water Resour. Res.*, *15*(4), 845–860, 1979.

Noble, B., and J. W. Daniel, *Applied Linear Algebra*, 2nd ed., p. 28, Prentice-Hall, Englewood Cliffs, N. J., 1977.

Pearl, R. H., R. S. Roberts, K. M. Keene, and T. J. McClain, Water Resources of Northwestern Kansas, *Hydrologic Investigations*, Atlas HA-429, Washington, D. C., 1972.

Pimental, K. D., Asymptotic estimation error growth applied to monitoring, paper presented at Chapman Conference on Applications of Kalman Filter to Hydrology, Hydraulics and Water Resources, AGU, Pittsburgh, Pa., May 1978.

Rouhani, S., Optimal data collection in random fields, Ph.D. thesis, Div. of Appl. Sci., Harvard Univ., Cambridge, Mass., 1983.

Smith, L., and R. A. Freeze, Stochastic analysis of steady state groundwater flow in a bounded domain, 1, One-dimensional simulations, *Water Resources Res.*, *15*(3), 521–528, 1979.

Sophocleous, M., J. E. Paschetto, and R. A. Olea, Groundwater network design for Northwest Kansas, using the theory of regionalized variables, *Ground Water*, *20*(1), 48–58, 1982.

Yeh, W. W. G., Y. S. Yoon, and K. S. Lee, Aquifer parameter identification with kriging and optimum parametrization, *Water Resources Res.*, *19*(1), 225–233, 1983.

S. Rouhani, School of Civil Engineering, Georgia Institute of Technology, Atlanta, GA 90332.

(Received October 4, 1984;
revised January 28, 1985;
accepted February 1, 1985.)

Comparative Study of Ground-Water Mapping Techniques

by Shahrokh Rouhani[a]

ABSTRACT

Mapping of ground-water spatial data is an important part of any geohydrologic investigation. There are three main classes of interpolators used for such mappings. The first group include simple estimators which are commonly used in practice. The second group are least-squares estimators which are basically fitting processes. The last category are Gauss-Markov estimators, such as kriging, which beside being exact interpolators, produce measures for the accuracy of the estimated field. These estimators are compared theoretically and numerically. These studies show that kriging yields relatively robust estimates. However, its suggested statistical inference method may not always produce robust estimates of the covariance function parameters. Simple estimators produce unstable results, while least-squares methods ignore local variations by fitting a single polynomial function over the whole field. For this study, water-table data from northwest Kansas are used.

INTRODUCTION

In most ground-water investigation studies, the initial geohydrologic data are in the form of scattered point values. It is one of the tasks of the investigators to interpolate these values in order to get a more complete picture of the spatial characteristics of the field of interest. To accomplish this, automatic mapping and interpolating techniques may be utilized. These methods vary in their levels of complexity and operational efficiency.

Most common interpolation procedures are based on linear combinations of existing data. These methods are linear in the following sense:

$$\hat{Z}(X_o) = \sum_i \lambda_{io} Z(X_i) \qquad (1)$$

where

$\hat{Z}(X_o)$ = estimated value at X_o (e.g., estimated ground-water levels at X_o);

$Z(X_i)$ = measured value at X_i, $i = 1, \ldots, N$ (e.g., measured ground-water levels);

λ_{io} = coefficient or weight of $Z(X_i)$ to estimate $Z(X_o)$.

These algorithms can be divided into three classes: (1) simple estimators, (2) least-squares estimators, and (3) Gauss-Markov estimators. In the first group, interpolation is done by using an assumed function for λ_{io}. In least-squares estimators, λ_{io} are estimated by fitting a function to the data. In the third group, $Z(X)$ is assumed to be a spatially distributed random variable with a specific correlation function, also known as a random field.

1. The Simple Methods

These algorithms can be readily applied to any spatial data. In this group are the following:

a. The Nearest Neighbor Method

According to this technique the estimated value at any given point is taken as the measured value at the nearest data point. This method represents the simplest approach to interpolation. It does not make any explicit assumption about the underlying field, and consequently, does not require any statistical information about the structure of $Z(X)$.

[a]Assistant Professor, School of Civil Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332.

Received October 1984, revised August 1985, accepted September 1985.

Discussion open until September 1, 1986.

b. The Arithmetic Mean

In this method $Z(X_o)$ is estimated by the average of its surrounding data values as follows:

$$\hat{Z}(X_o) = \sum_i \lambda_{io} Z(X_i) \qquad (2)$$

where $\lambda_{io} = 1/N$. N is the number of data points in the surrounding area. The above values for λ_{io} are identical to estimation weights used in the case of a random field with a constant expected value and uncorrelated residuals. The similarity shows that the seemingly simple model for λ_{io} in equation (2) may imply an implicit statistical structure for Z(X), if studied as a random field.

c. The Distance Weighting Function

In this approach the data values are weighted according to their distances from the estimated point. For example in inverse squared distance weighting (ISDW), λ_{io} is defined as:

$$\lambda_{io} = | X_i - X_o |^{-2} \qquad (3)$$

where X_i and X_o correspond to the locations of the measured and the estimated points, respectively. The above model indicates that the closer the two points, the higher their correlation. However, similar to the previous method, the relative location of data points with respect to each other has no significance in the interpolation process.

2. Least-Squares Estimators

Least-squares estimation also can be categorized as a linear method (Delfiner, 1975). Given basic functions $f_p(X)$ (e.g., monomial functions), Z(X) is estimated as follows:

$$\hat{Z}(X_o) = \sum_p b_p f_p(X_o) \qquad (4)$$

which is the regression equation with unknown coefficients b_p. The b_p are estimated by minimizing the following equation with respect to b_p:

$$S = \sum_i [Z(X_i) - \sum_p b_p f_p(X_i)]^2 \qquad (5)$$

where S is known as the sum of squares of errors (SSE). X_i denotes the measurement points. In this class are:

a. The Ordinary Least-Squares Estimator (LSE)

LSE views Z(X) as a field with the following structure:

$$Z(X) = M(X) + R(X) \qquad (6)$$

where

$$M(X) = \sum_p b_p f_p(X)$$

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

$$E[R(X_i)R(X_j)] = C\delta(h)$$

where $\delta(h)$ is the Dirac delta function. M(X) is a deterministic function, and R(X) is an uncorrelated zero-mean error term with a variance of C.

LSE also produces a measure of the goodness-of-fit, $\hat{\sigma}^2$, or the mean square error which is defined as

$$\hat{\sigma}^2 = \frac{1}{N} \sum_i [Z(X_i) - \sum_p b_p f_p(X_i)]^2 \qquad (7)$$

Unfortunately, $\hat{\sigma}^2$ gives only an over-all measure of the goodness-of-fit at data points, and does not reflect the accuracy of individual point estimates.

b. The Generalized Least-Squares Estimator (GLSE)

This method is identical to the LSE except that a correlated error with K( ) as its correlation function is included:

$$E[R(X_i)R(X_j)] = K(X_i, X_j) \qquad (8)$$

As a result, the criterion SSE in equation (5) becomes more complicated:

$$S = \sum_i \sum_j [Z(X_i) - M(X_i)]\, K(X_i, X_j) [Z(X_j) - M(X_j)] \qquad .....(9)$$

where S in equation (9) is called the "weighted" sum of squares of errors (WSSE).

Due to the more complicated form of WSSE, simple optimization methods such as Lagrange multipliers are not usually applicable. So, quadratic programming algorithms have to be used. These nonlinear optimization schemes are computationally less efficient than Lagrange multipliers. Furthermore, GLSE requires prior knowledge of $K(X_i, X_j)$ (i.e., the correlation matrix). Here, one can easily see that as the level of the sophistication of the method increases, the background information requirements along with the operational complexity of the method also increase.

3. The Gauss-Markov Estimators

The following techniques yield unbiased minimum variance estimates, which are the core of the Gauss-Markov theorem (Meyer, 1975). In practice these algorithms substitute the minimization of SSE by the minimization of the estimation variance. This substitution becomes very useful when dealing with fields with correlated residuals. In such cases the Gauss-Markov estimators can

adapt simple optimization techniques such as Lagrange multipliers, and thus they are operationally more efficient than GLSE. These methods are based on the following criteria:

(1) Unbiasedness: $E[\hat{Z}(X_o) - Z(X_o)] = 0$, where $Z(X_o)$ is the unknown underlying value of Z at X_o, and

(2) Minimum squared error: this requires $E[\hat{Z}(X_o) - Z(X_o)]^2$ to be minimum. These conditions can be written as

$$E[\hat{Z}(X_o) - Z(X_o)] = 0$$

$$\text{Var}[\hat{Z}(X_o) - Z(X_o)] = \text{minimum} \qquad (10)$$

where $\text{Var}[\hat{Z}(X_o) - Z(X_o)]$ is defined as the estimation variance.

From the point of view of modeling, a Gauss-Markov estimator views the field Z(X) as a spatial random function with the following model:

$$Z(X) = M(X) + R(X) \qquad (11)$$

where:

(i) M(X) is a slowly varying deterministic function known as the "drift" which is equal to the expected value of Z at point X ($x_{1i}, x_{2i}, \ldots, x_{ni}$) in R^n. It may be further assumed that M(X) admits a local representation in the form of a polynomial of order k as follows:

$$M(X) = \sum_{p=1}^{\ell(k)} b_p f_p(X) \qquad (12)$$

where b_p are fixed unknown coefficients, and $f_p(X)$ are basic functions of the polynomial, i.e.:

$$f_p(X) = x_1^{p_1} x_2^{p_2} \ldots x_n^{p_n} \qquad (13)$$

on the condition that $p_1 + p_2 + \ldots + p_n \leq k$. $\ell(k)$ is the number of monomials M(X).

(ii) R(X) is a spatially fluctuating random component with zero expectation. In this group are:

a. Objective Analysis

This method was developed by Gandin (1965) and has been widely applied for mapping of random fields in meteorology and oceanography (e.g., Tu, 1981). For this technique, as for all Gauss-Markov estimators, the random field is viewed as a Bayes model (Schweppe, 1973) as defined by equation (11), with:

$$M(X) = E[Z(X)] = M$$

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

$$E[R(X_i)R(X_j)] = K(|X_i - X_j|) \qquad (14)$$

According to this technique, M(X) is constant throughout the field and Z(X) is second-order stationary. Based on equation (14), $Z(X_o)$ is estimated as:

$$\hat{Z}(X_o) = \sum_i \lambda_{io} Z(X_i)$$

where λ_{io} are calculated by the minimization of the $\text{Var}[\hat{Z}(X_o) - Z(X_o)]$. Naturally, one has to know the following statistical properties prior to the interpolation process: (1) the constant drift, M, and (2) the correlation structure of the random field, K( ). Gandin (1965) slightly varied the model (14). For example, he assumed M(X) to be a correlated random function with zero mean and R(X) was assumed to be uncorrelated with a nonzero mean, known as the measurement bias.

b. Simple Kriging

This method has been applied to many geological and hydrological estimation problems; for instance, see David (1977), Delhomme (1979), Sophocleous *et al.* (1982), Aboufirassi and Marino (1983), and Bastin *et al.* (1984). Kriging is essentially similar to objective analysis; however, there is a difference between these two methods. For the objective analysis one assumes weak stationarity for Z(X). For kriging it is only assumed that the first-order increments of Z(X) are weakly stationary. The assumption for simple kriging can be written:

$$E[Z(X_i) - Z(X_j)] = 0$$

$$\text{Var}[Z(X_i) - Z(X_j)] = 2\gamma(|X_i - X_j|) \qquad (15)$$

where $\gamma(h)$ is defined as the "semivariogram." Assumptions (15) eliminate the need for prior estimation of M which is required by the objective analysis.

c. Intrinsic Random Functions (IRF)

If M(X) is nonstationary, then the drift has to be estimated prior to the mapping. In order to avoid this, Matheron (1973) proposes a new method. In this method the process Z(X) is viewed as an intrinsic random function (IRF) which could be made stationary by a process known as "incrementing." A k^{th} order intrinsic random function (IRF) is defined as a random process which requires a k^{th} order filtering to achieve stationarity. In other words in IRF, estimation weights, λ_{io}, are defined in such a manner that the linear combination $\sum_{i=0}^{N} \lambda_{io} Z(X_i)$ filters out the

mean, which is assumed to be a polynomial of order k. Consequently, this approach leaves out the important step of estimating a drift.

For the case of an IRF of order 0, 1, or 2 in a two-dimensional space with Cartesian coordinates (x_i, y_i), the above incrementing constraints can be written as:

$$
\begin{aligned}
k = 0 \quad & \sum_{i=0}^{N} \lambda_{io} = 0 \\
k = 1 \quad & \sum_{i=0}^{N} \lambda_{io} x_i = 0 \qquad \sum_{i=0}^{N} \lambda_{io} y_i = 0 \\
k = 2 \quad & \sum_{i=0}^{N} \lambda_{io} x_i y_i = 0 \qquad \sum_{i=0}^{N} \lambda_{io} x_i^2 = 0 \\
& \sum_{i=0}^{N} \lambda_{io} y_i^2 = 0 \qquad (16)
\end{aligned}
$$

The above constraints (16) constitute the unbiasedness criterion of the original kriging [equation (10)].

The criterion of minimum squared error in equation (10) is defined as:

$$
\mathrm{Var}\left[\sum_{i=0}^{N} \lambda_{io} Z(X_i)\right] = \sum_{i=0}^{N} \sum_{j=0}^{N} \lambda_{io} \lambda_{io} K(|X_i - X_j|) \quad \ldots . \quad (17)
$$

where $K(\ )$ is the covariance function, and $\lambda_{oo} = -1$.

Now, in order to calculate λ_{io}, it is necessary to minimize (17) subject to constraints (16). This minimization can be done by Lagrange multipliers, μ_{po}. At its minimum the estimation variance (17) takes the value of:

$$
\mathrm{Var}[\hat{Z}(X_o) - Z(X_o)] = K(|X_o - X_o|) - \sum_{j=1}^{N} \lambda_{io} K(|X_o - X_j|) - \sum_{p=1}^{\ell(k)} \mu_{po} f_p(X_o) \qquad (18)
$$

Equation (18) is a measure for the accuracy of $\hat{Z}(X_o)$. The knowledge of covariance function and the order k are the only prerequisites for the minimum variance estimation of $\hat{Z}(X_o)$. Kafritsas and Bras (1981) note that in the process of estimation, kriging preserves the observed values. This quality makes this method an "exact interpolator."

As a step toward more computational efficiency, Matheron (1973) suggests a family of functions as the general form of admissible polynomial isotropic covariance functions for IRF-k in n dimensional space, as shown in Table 1. Variograms and simple covariance functions may also be utilized.

SUMMARY OF THE THEORETICAL ANALYSIS

From the above brief description of these methods, one can infer that for an over-all theoretical comparison of interpolation techniques, three main characteristics should be considered. These are: (1) the required information prior to the interpolation, (2) the efficiency of the interpolation algorithm, and (3) the type of measure of the accuracy of the outputs. These factors are all interdependent. For instance, given a good estimate of K(h), kriging yields robust estimates with a measure of their accuracy. However, if the data are scarce, the choice of K(h) becomes rather subjective which in turn adds more uncertainty to the results.

Keeping the above characteristics in mind, one can find instances in which one method has clear advantages over the other. For example, the IRF is computationally more efficient than the GLSE, the objective analysis, and the simple kriging, because it requires similar or less information than others while yielding better results with a measure for their accuracy. However, when one compares the IRF to the LSE or simple methods,

Table 1. Selected Models for Generalized Covariances (Delfiner, 1975)

Drift	*k*	f_p *in* R^2	f_p *in* R^3	*Models of GC*
Constant	0	1	1	$K(h) = C\delta(h) + a_1 h$
Linear	1	1, x, y	1, x, y, z	$K(h) = C\delta(h) + a_1 h + a_3 h^3$
Quadratic	2	1, x, y, xy, x^2, y^2	1, x, y, z, xy, xz, yz, x^2, y^2, z^2	$K(h) = C\delta(h) + a_1 h + a_3 h^3 + a_5 h^5$
Constraints on the coefficients		$a_1 \leq 0 \quad a_5 \leq 0$		
		in R^2: $a_3 \geq -\frac{10}{3}(a_1 a_5)^{1/2}$	in R^3: $a_3 \geq -(10\, a_1 a_5)^{1/2}$	

such clear advantages do not exist because unlike the IRF, the LSE and simple methods do not require detailed statistical information. In fact, the simple methods, and to some degree the LSE, have very simple interpolation algorithms. This operational simplicity is achieved by ignoring the statistical structure of Z(X). So the choice of the best appropriate method depends on the amount of available information, the computational resources available to the user, and the significance or the desired level of accuracy of the mapping results.

THE NUMERICAL COMPARISON

Based on the above conclusions, those methods which show no distinct advantage or disadvantage over the others are compared. The IRF, the LSE, and the inverse squared distance weighting (ISDW) methods are selected for numerical comparison. For the following case study, a set of water-table data has been used.

Data Description

The available data are ground-water level observations made in January 1979 in a section of Groundwater Management District No. 4 of Kansas, an area of 2048 square miles in northwestern Kansas, including parts of Sherman, Thomas, Cheyenne, Rawlins, Logan, and Wallace counties (Figure 1). This subregion lies between latitudes 38°48′ and 39°48′ North and longitudes 101° and 101°36′ West. There are 84 measurement points in this area (see Table 2). Northeastern and northwestern corners of this zone are rather densely measured, while central and southern parts of this subregion have relatively scattered data points. In general there is no area of excessive pumping or recharging that can be shown as a major sink or source for the aquifer. For further study of the geohydrology of this region, readers are referred to Pearl *et al.* (1972).

Computer Programs

Two major programs have been used in this study. For the purpose of the structural analysis and kriging, a versatile program named AKRIP (Kafritsas and Bras, 1981) has been utilized. This program is based on the suggested algorithm by Delfiner (1975). The other program used was the SYMAP (SYnographic MAPping System) developed by the Harvard Laboratory for Computer Graphics and Spatial Analysis (Dougenik, 1975). SYMAP is basically employed for the purpose of mapping the kriging results, least-squares trend fitting, and distance weighting interpolations.

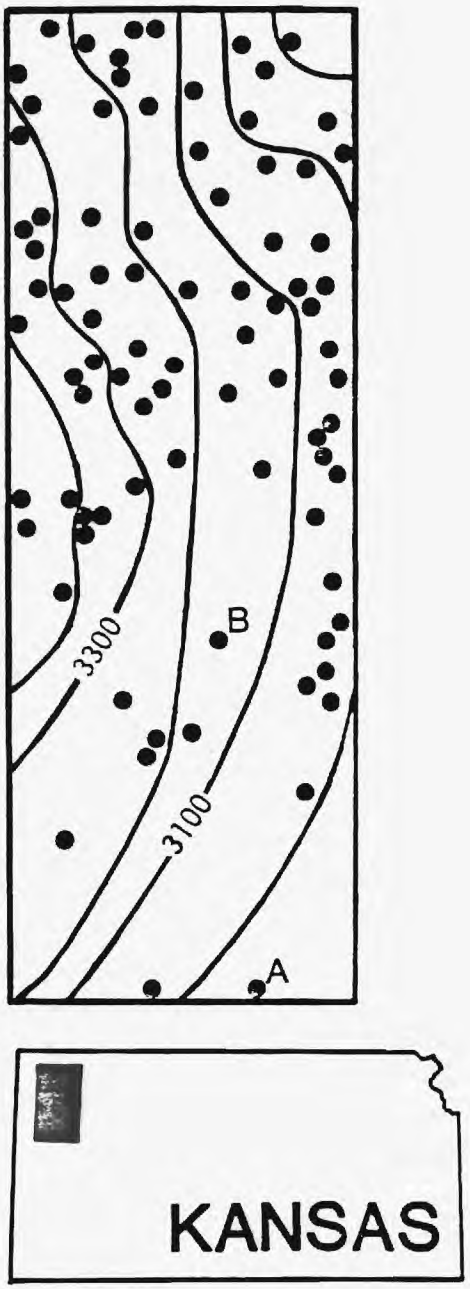


Fig. 1. Hand-drawn contour map of ground-water levels in northwest Kansas. (Map scale: 1 inch = 11.77 miles for X = 7.75 miles for Y.) Measurements are in feet. Black dots indicate the observation sites.

ANALYSIS OF RESULTS

1. Inverse Squared Distance Weighting (ISDW)

First, the available data are interpolated by ISDW. The point distribution coefficient of the data set is 1.13 which is an indication that the measurements are located irregularly. The randomness of the spatial distribution of the data means that the data points are not clustered so a meaningful interpolation can be done. No other statistical information was required.

Two special data points, A and B, are selected to test the robustness of our estimates in sparsely

Table 2. Data Points Locations and Values

Point	y (mile) Down	x (mile) Across	Z(x,y) (ft)	Point	y (mile) Down	x (mile) Across	Z(x,y) (ft)
(1)	1.18	4.16	3239.00	(43)	32.02	7.35	3356.00
(2)	1.93	7.08	3196.00	(44)	32.27	7.84	3349.00
(3)	3.31	9.95	3175.00	(45)	33.16	6.81	3372.00
(4)	6.06	8.43	3205.00	(46)	31.04	0.70	3445.00
(5)	4.18	0.33	3295.00	(47)	31.16	5.08	3391.00
(6)	5.94	2.11	3292.00	(48)	32.54	1.84	3443.00
(7)	7.44	0.81	3308.00	(49)	36.55	4.33	3433.00
(8)	12.58	7.62	3244.00	(50)	43.17	9.84	3228.00
(9)	12.20	3.46	3312.00	(51)	17.10	29.52	3039.00
(10)	13.33	1.57	3339.00	(52)	21.61	29.79	3050.00
(11)	14.83	2.16	3341.00	(53)	17.85	27.41	3060.00
(12)	48.94	26.87	3099.00	(54)	17.98	27.90	3045.00
(13)	46.55	13.52	3219.00	(55)	18.74	24.54	3101.00
(14)	46.67	13.14	3216.00	(56)	16.97	21.90	3112.00
A (15)	61.59	22.49	2912.00	(57)	20.36	22.00	3146.00
(16)	2.18	26.65	2886.00	(58)	16.47	11.52	3219.00
(17)	3.43	24.38	2926.00	(59)	17.46	15.95	3189.00
(18)	1.55	21.35	2973.00	(60)	20.86	11.89	3263.00
(19)	0.53	14.27	3128.00	(61)	21.85	15.19	3229.00
(20)	0.92	11.25	3161.00	(62)	23.37	29.90	3041.00
(21)	6.82	29.30	2981.00	(63)	23.13	24.49	3124.00
(22)	8.82	31.03	2972.00	(64)	26.12	29.41	3055.00
(23)	9.70	24.33	3039.00	(65)	26.75	28.11	3065.00
(24)	4.92	16.92	3092.00	(66)	24.00	20.49	3169.00
(25)	6.69	22.22	2986.00	(67)	24.99	12.38	3273.00
(26)	8.94	18.16	3068.00	(68)	27.63	15.41	3252.00
(27)	4.06	10.22	3176.00	(69)	29.13	30.28	3041.00
(28)	5.93	12.60	3172.00	(70)	28.26	28.76	3056.00
(29)	14.34	29.73	3023.00	(71)	28.77	23.52	3124.00
(30)	10.08	27.36	3021.00	(72)	31.77	27.63	3062.00
(31)	14.47	25.09	3072.00	(73)	30.13	12.00	3302.00
(32)	11.95	19.68	3099.00	(74)	38.65	29.84	3017.00
(33)	13.46	12.33	3200.00	(75)	35.39	29.36	3043.00
(34)	16.47	8.38	3261.00	(76)	43.79	29.57	3057.00
(35)	17.09	5.51	3294.00	(77)	40.15	29.09	3015.00
(36)	18.97	7.79	3288.00	(78)	41.54	28.60	3038.00
(37)	17.22	2.00	3354.00	(79)	42.55	27.46	3076.00
(38)	19.48	0.87	3391.00	(80)	45.59	16.87	3189.00
(39)	22.24	7.41	3318.00	(81)	52.10	4.60	3229.00
(40)	22.86	6.38	3336.00	(82)	61.94	13.46	3011.00
(41)	23.36	9.89	3290.00	B (83)	39.69	19.30	3173.00
(42)	24.37	6.76	3337.00	(84)	25.75	32.06	3026.00

sampled areas as indicated in Figure 1. Point A is located in the lower part of the map where the number of measurement points is very low, while point B is in an area of sparse data surrounded by areas of high data density.

In the first round of interpolation all points, including A and B, are used to produce the contour map shown in Figure 2a, which shows a lot of small-scale variations throughout the map, especially in its lower part. This area is an area of low data density, so these patterns should be viewed with caution. In order to test the reliability of these estimates, point A is omitted. As a result, the contour lines in the lower part of Figure 2b are changed drastically. This significant change shows one of the weaknesses of ISDW interpolation in sparsely sampled areas. In Figure 2c, point B is also omitted, but in contrast to the previous case, the omission of point B causes only an increase in the hydraulic gradient of the central part of the map. It must be remembered that the choice of the weighting function is still arbitrary, and thus, the validity of the final results has to be checked by other means.

Table 3. Polynomial Trend-Fitting Statistical Results

Order of polynomial	*1*	*x*	*y*	x^2	*xy*	y^2	x^3	x^2y	xy^2	y^3	*Coefficient of correlation*
1	3337.9	−12.30	1.47								.895
2	3252.5	−13.73	12.48	.0181	.0195	−.2133					.983
3	3266.1	−12.37	8.61	−.1171	.0342	−.0354	.0039	−.0022	.0010	−.0024	.986

2. Least Squares Trend Fitting

Three polynomials of the first, second, and third order are fitted to the data. The statistical results are shown in Table 3. They show close correlation between the trend and the measured values. The coefficient of correlation rises from .895 to .986 when it goes from a first- to a third-order polynomial. In spite of a high coefficient of correlation, the least-squares estimator ignores local variations. This forced orderliness is most obvious in the upper right corner and the lower part of this subregion. Furthermore, the coefficient of correlation gives only an over-all measure of the goodness-of-fit with respect to the data points. Thus it does not give any direct measure for the accuracy of estimated values.

3. Kriging Results

The Structural Analysis

As was discussed in the theory of kriging, one has to estimate the order of the IRF and the covariance function parameters prior to the interpolation. This is done by the structural analysis. In this work, the suggested analysis by Delfiner (1975) and polynomial generalized covariance functions (Table 1) are used.

In order to avoid ill-conditioned matrices in the interpolation process, a minimal allowable distance between each pair of two data points, DR, has to be selected. The program discards some of the data so that there are no two measured points with a distance less than or equal to DR between them. It also averages the observed values of each retained point and of its discarded neighbors and assigns the result to the retained point. The choice of the minimal allowable distance is arbitrary and may differ from case to case. One may assume that discarding a few points through the application of DR is merely the smoothing of microscale variations in the data values. The results of these pre-kriging smoothings and the structural analysis for DR = .5, 1, and 2 miles are shown in Table 4.

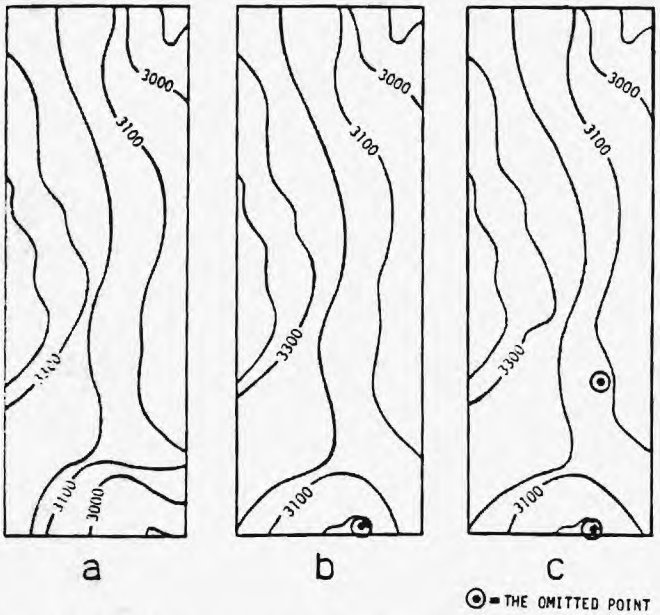


Fig. 2. Water-levels contour map produced by distance weighting function. (Map scale: 1 inch = 11.77 miles for X = 7.75 miles for Y.) a. All points are included; b. Point A is omitted; c. Points A and B are omitted (in feet).

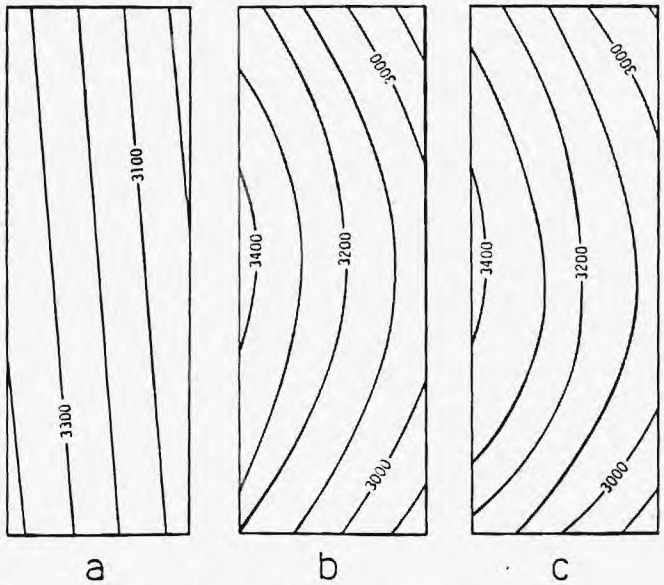


Fig. 3. Water-levels trend-fitted surface. (Map scale: 1 inch = 11.77 miles for X = 7.75 miles for Y.) a. First-order polynomial; b. Second-order polynomial; c. Third-order polynomial (in feet).

Table 4. The Results of Structural Analysis

DR	*n*	*k*	*G.C.*
.5	1	0	$K(h) = 1552.1\,\delta(h)$
1.	4	1	$K(h) = 145.68\,\delta(h) + .89914\,h^3$
2.	14	1	$K(h) = 125.19\,\delta(h) + .98978\,h^3$

DR = Minimum allowable distance between data points (miles).

n = Number of discarded points.

k = Order of the intrinsic function.

K(h) = The generalized covariance (ft^2).

h = The distance vector (miles).

δ = Dirac delta function.

The sensitivity of the estimated covariance function parameters to changes in the data set is quite obvious. However, it can be seen that after some microscale smoothing (i.e., elimination of four points in DR = 1 mile), the structural analysis tends to produce robust estimates of the covariance function. The uncertainty in these pre-interpolation estimates are not formally included in kriging. Consequently, they might add an unmeasured amount of error to the interpolated values.

Mapping Results

For kriging purposes the field is divided into an 8 × 18 grid with $\Delta x = \Delta y = 4$ miles. The results of the contour maps and their corresponding kriging variances for cases of DR = .5, 1, and 2 miles are illustrated in Figures 4 and 5. They provide a basis for a comparative analysis of the covariance functions. As expected in the first case (i.e., the pure nugget effect) the variances are identical for all estimated points (see Figure 5a). The calculated variance seems rather high when compared to the other two cases (see Figures 5b and 5c).

Considering the other two cases, one easily observes the close similarity of the kriging maps (see Figures 4b and 4c). In fact, kriging produces rather similar maps when the covariance functions are of the same order.

The comparison of Figures 5b and 5c also shows close agreement between the estimated accuracies of the kriged values. However, it seems that kriging with DR = 2 miles has produced a slightly more reliable map. This is mainly due to the lower nugget effect (i.e., C) in the case of DR = 2 miles. The above comparisons show how the variabilities in the covariance functions influence the estimation variances. So these variances should be viewed only as a relative measure for the accuracy of estimates.

Two distinct points A and B (see Figure 1) are selected to study the effect of data point omission.

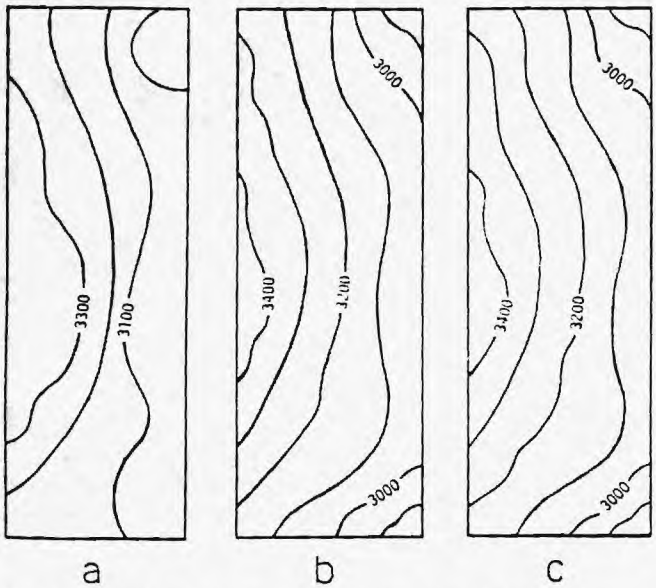


Fig. 4. Water-levels contour map by kriging. (Map scale: 1 inch = 11.77 miles for X = 7.75 miles for Y.)
a. DR = .5 mile; b. DR = 1. mile; c. DR = 2. miles (in feet).

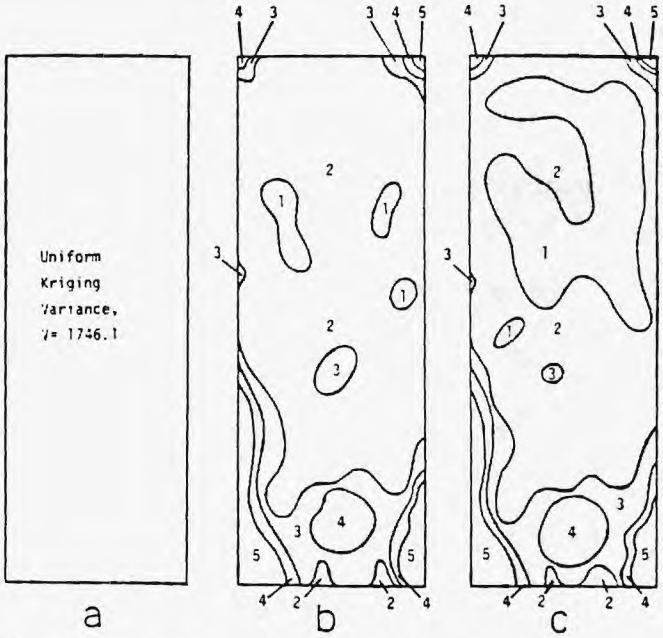


Fig. 5. Contour of estimation variances. (Map scale: 1 inch = 11.77 miles for X = 7.75 miles for Y.)
a. DR = .5 mile; b. DR = 1. mile; c. DR = 2. miles (in feet).

Map symbols:		1	2	3	4	5
Range (sf)	Min	0	200	400	600	800
	Max	199	399	599	799	—

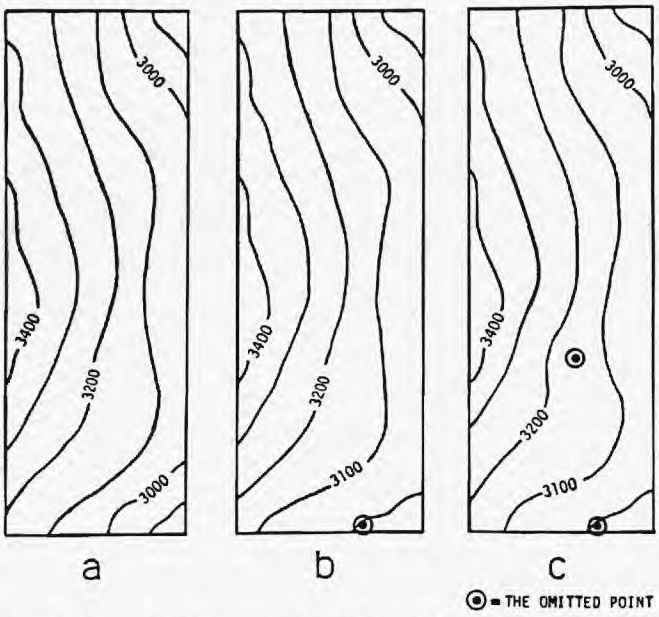


Fig. 6. Water-levels contour map by kriging (DR = 1. mile). (Map scale: 1 inch = 11.77 miles for X = 7.75 miles for Y.) a. All points are included; b. Point A is omitted; c. Points A and B are omitted (in feet).

Both of these two points are located in sparsely sampled areas. Point A is located near the lower boundary of the map, while point B is surrounded by areas of high data point density. With the ISDW procedure, the deletion of point A causes a drastic change in the pattern of contour lines, while skipping point B causes only marginal changes in the hydraulic gradient. Now with kriging, as shown in Figure 6, the removal of points A and B has practically no effect on the map. However, in both cases after omitting points A and B, the level of uncertainty rises in the neighboring areas of these two points, as shown in Figures 7b and 7c. These increases give proper signals to warn the map producer of the high level of uncertainty created by skipping these two points.

Further study of estimation variances indicate that the lower part of the map is basically an uncertain area so that more sampling should be done in that part. The central section, where point B is located, is also an uncertain area, but its level of uncertainty is moderate, and at this point it should not necessarily be a target place for further measurements.

SUMMARY OF THE NUMERICAL ANALYSIS

Generally speaking, the ISDW procedure produces maps with rather low reliability and high sensitivity to measured values especially in sparsely sampled areas. Kriging tends to yield much more robust results and takes the spatial structure of the data points into account. In contrast to LSE, kriging also reflects small-scale variations in its maps. The other important advantage of kriging is the estimation variance which yields a measure of the accuracy of any single interpolated value. This measure can have a dual role. First, it evaluates the reliability of our estimates. Secondly, it can serve as a guideline to identify the most uncertain areas for further measurements. So kriging can be an effective tool both for mapping and planning of data sampling activities (see Rouhani, 1985).

There are several drawbacks to this method. Kriging demands a significant amount of prior statistical information. In addition, the suggested statistical inference algorithm includes some arbitrary choices such as the selection of DR (i.e., minimum allowable distance between data points). Another handicap of this method is the lack of any measure of reliability of estimated covariance functions. In the absence of an extensive data set, kriging might generate significant variations in its covariance function estimates. In such cases a less sophisticated method may be more appropriate for the contouring of a random field. In general, as mentioned earlier, the choice of the best appropriate mapping method depends on the amount of

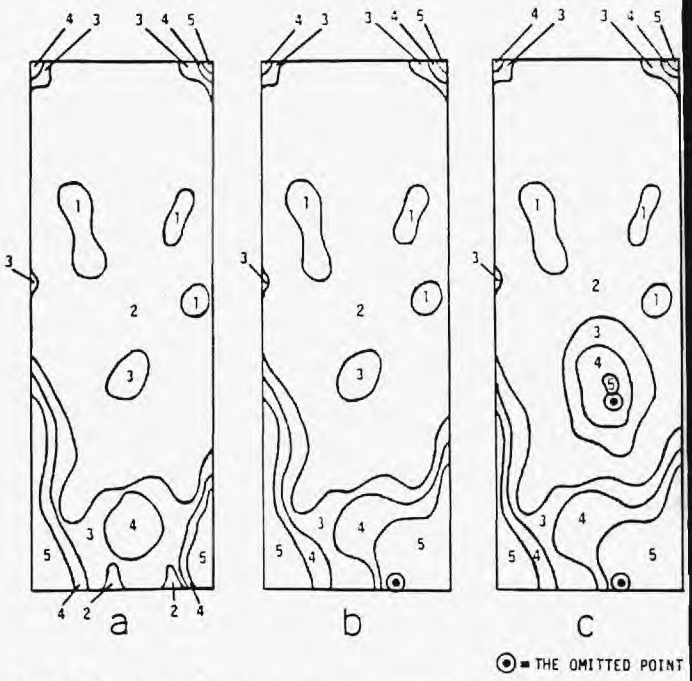


Fig. 7. Contour of estimation variances (DR = 1. mile). (Map scale: 1 inch = 11.77 miles for X = 7.75 miles for Y.) a. All points are included; b. Point A is omitted; c. Points A and B are omitted (in feet).

Map symbols:		1	2	3	4	5
Range (sf)	Min	0	200	400	600	800
	Max	199	399	599	799	—

available data, the technical resources of users, and the desired level of accuracy of interpolated maps.

ACKNOWLEDGMENTS

This research is based on my Ph.D. thesis work at the Division of Applied Sciences, Harvard University, which was supported in part by a grant from the Ford Foundation to Harvard University. Many thanks go to my thesis advisors, Professor Peter P. Rogers and Myron B. Fiering, for their support, guidance, and friendship. Thanks also to Mrs. Linda LaSalata for an excellent job of typing the manuscript.

REFERENCES

Aboufirassi, M. and M. A. Marino. 1983. Kriging of water levels in the Sauss aquifer, Morocco. Math. Geol. v. 15, no. 4, pp. 537-551.

Bastin, G., B. Lorent, C. Duque, and M. Gevers. 1984. Optimal estimation of the average rainfall and optimal selection of rain gauge locations. Water Resour. Res. v. 20, no. 4, pp. 463-470.

David, M. 1977. Geostatistical Ore Reserve Estimation. Elsevier Scientific, New York.

Delfiner, P. 1975. Linear estimation of non-stationary spatial phenomena. Advanced Geostatistics in the Mining Industry, M. Guarascio, M. David, and C. Huijbregts, eds. D. Reidel Publishing Co., Boston. pp. 49-68.

Delhomme, J. P. 1979. Spatial variability and uncertainty in groundwater flow parameters: a geostatistical approach. Water Resour. Res. v. 15, no. 2, pp. 269-280.

Dougenik, J. A. 1975. SYMAP User's Manual. Laboratory of Computer Graphics and Spatial Analysis, Harvard University, Cambridge. 32 pp.

Gandin, L. S. 1965. Objective Analysis of Meteorological Fields. Israel Program of Scientific Translations, Jerusalem. 242 pp.

Kafritsas, J. and R. L. Bras. 1981. The Practice of Kriging. Report No. 263, Ralph M. Parsons Lab., M.I.T., Cambridge. January, 107 pp.

Matheron, G. 1973. The intrinsic random function and their applications. Advanced Applied Probability. v. 5, pp. 439-468.

Meyer, S. L. 1975. Data Analysis for Scientists and Engineers. John Wiley and Sons, Inc., New York. p. 390.

Pearl, R. H., R. S. Roberts, K. M. Keene, and T. J. McClain. 1972. Water resources of northwest Kansas. Hydrologic Investigations Atlas HA-429, USGS, Washington. 2 sheets.

Rouhani, S. 1985. Variance reduction analysis. Water Resour. Res. v. 21, no. 6, pp. 837-846.

Schweppe, F. C. 1973. Uncertain Dynamic Systems. Prentice-Hall, Inc., Englewood Cliffs, NJ. 563 pp.

Sophocleous, M., J. E. Paschetto, and R. A. Olea. 1982. Ground-water network design for northwest Kansas, using the theory of regionalized variables. Ground Water. v. 20, no. 1, pp. 48-58.

Tu, K. 1981. A combined statistical and dynamical approach to regional forecast modeling of open ocean currents. Ph.D. Thesis, Dept. of Applied Sciences, Harvard University, Cambridge. 160 pp.

* * * * *

Shahrokh Rouhani received his B.S. in Civil Engineering from University of California, Berkeley in 1978 and his S.M. and Ph.D. in Engineering Sciences from Harvard University in 1983. He joined the faculty of Civil Engineering at Georgia Institute of Technology in 1983. His research interests are stochastic hydrology and surface and ground-water hydrology.

Journal of Hydrology, 89 (1986) 1–11
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

[4]

RESILIENCE OF A STATISTICAL SAMPLING SCHEME

SHAHROKH ROUHANI and MYRON B FIERING

School of Civil Engineering, Georgia Institute of Technology, Atlanta, GA 30332 (U.S.A.)
Division of Applied Sciences, Harvard University, Cambridge, MA (U.S.A.)

ABSTRACT

Rouhani, S. and Fiering, M.B, 1986. Resilience of a statistical sampling scheme. J. Hydrol., 89: 1–11.

Most statistical sampling algorithms on hydrologic random fields assume that the new measurements will agree reasonably well with their predicted values. This in turn implies the stationarity of the estimated covariance function. In order to test the reliability of one such statistical algorithm (i.e., variance reduction analysis), noisy input data are generated, and results of sampling from these data are compared to the case of sampling with the unperturbed data. These comparisons and a related regret analysis reveal that the effects of the noisy data are primarily accommodated by adjustments to the covariance function parameters, while selected sets show a high degree of resilience. Variance reduction analysis seems to be a reliable method for maximizing information by sampling random fields with an unstable parameter space but a resilient action space.

INTRODUCTION

Many authors have advocated the use of statistical methods for the design of sampling schemes on hydrologic random fields. Commonly these procedures are based on the maximization of incremental information subject to budget constraints. For instance, Fiering (1965) and Matalas (1968) suggest minimization of the total variance of estimates of flow as the objective function for gaging schemes. Using non-linear integer programming, they identify from a set of potential sites those sampling locations which yield minimal total variance. Bastin et al. (1984) compute all possible combinations of n sampling sites out of m potential locations in order to identify the subset which produces minimal kriging variance. Brady (1978), Crawford (1979), Hughes and Lettenmaier (1981) and Chou and Scheck (1984) use iterative algorithms to minimize regional or areal estimation variances. Rouhani (1985) proposes variance reduction analysis to select those sequences of n points so chosen from m potential sites to maximize reduction in the total variance of estimates.

In all the above work it is assumed that the statistical structure of the random field — represented by a covariance function — is known. Furthermore, it is assumed that the availability of new data does not affect the assumed covariance function.

The questions that immediately arise are: What is the effect on the sampling

0022-1694/86/$03.50

scheme if the predicted values of the field are significantly under- or overestimated? How does the covariance function respond to the newly sampled values? Are the decisions, i.e., selection of sampling sites, stable under such situations? To answer these questions we propose to test the robustness and the resilience of these statistical sampling algorithms.

RESILIENCE AND ROBUSTNESS

The concept of resilience is relatively new in the field of water resources. Fiering (1982) describes resilience as analogous to the robustness of statistical estimators. Matalas and Fiering (1977) define robustness as: "the insensitivity of a system design to errors, random or otherwise, in the estimates of those parameters affecting design choice". Resilience is the ability of the system to accommodate surprises and to survive under unanticipated perturbations. It implies that even if an unlikely event occurs, the decision has an acceptably high probability of being either correct or good enough. In other words, a tolerance ("good enough") and a confidence ("acceptably high") are required.

Fiering (1982) gives an example to illustrate the differences between robustness and resilience of a system: "The sensitivity of the system response with respect to a decision variable x_i is given by the partial derivative $\partial f/\partial x_i$. If the partial derivative is small, the system is "robust" with respect to such changes. If the partial derivative is not small, the system need not suffer important shifts in its response because changes in other decision variables might be made to accommodate an unfortunate choice of x_i." Therefore robustness alone does not reflect the behavior of the entire system. The total derivative $df/dx_i = \sum_j (\partial f/\partial x_j)\,(dx_j/dx_i)$ measures the system's ability to adjust to changes in x_i, some of which might be correlated. A (linear) combination of all derivatives df/dx_i might suggest a measure of resilience of the given system. In this paper we study the resilience of variance reduction analysis applied to a sampling scheme.

VARIANCE REDUCTION ANALYSIS

The proposed sampling algorithm is based on kriging, a linear interpolation method for variable random fields. Given the values $Z(X_i)$, $i = 1, \ldots, N$ of a field $Z(X)$ at the data points X_i, $i = 1, \ldots, N$, kriging provides a technique for estimating the value of linear functionals of Z at additional points.

In point kriging one estimates:

$$\hat{Z}(X_0) = \sum_i \lambda_{i0} Z(X_i) \tag{1}$$

where $\hat{Z}(X_0)$ is the kriging estimate at an arbitrary point X_0, and λ_{i0} is the kriging weight for $Z(X_i)$ to estimate $Z(X_0)$. The λ_{i0} are defined by two criteria:

$$\begin{aligned} &E[\hat{Z}(X_0) - Z(X_0)] = 0 \\ &E[\hat{Z}(X_0) - Z(X_0)]^2 = \text{kriging variance, to be minimum.} \end{aligned} \tag{2}$$

Without any loss of generality, it may be assumed that the expected value of $Z(X_0)$ is a polynomial of kth order:

$$E[Z(X_0)] = \sum_{p=1}^{l(k)} b_p f_p(X_0), \tag{3}$$

where b_p are fixed unknown coefficients, $f_p(X)$ is the pth monomial, and $l(k)$ is the number of these monomials in the above kth order polynomial. In a 2-dimensional space with Cartesian coordinates (x_0,y_0), a 2nd order polynomial ($k = 2$ and $l = 6$) has the following form:

$$E[Z(X_0)] = b_1 + b_2 x_0 + b_3 y_0 + b_4 x_0^2 + b_5 y_0^2 + b_6 x_0 y_0 \tag{4}$$

In our study we measure the accuracy of an estimated value in terms of its kriging variance, so it can be a guideline for optimal sampling of the field at new data points. For example, the area around which the kriging variance is largest can be selected for further data collection. However, such an approach ignores the effect of a new measurement on the level of accuracy of the estimated field as a whole. Rouhani (1985) proposes a new method to establish a measure for such an influence; this resembles a common response function by calculating the level of improvement in the accuracy of $\hat{Z}(X_0)$ due to a new measurement at $X*$ (the arbitrary location of a new sampling site). The level of improvement is measured in terms of reductions in the kriging variances. Furthermore, this variance reduction can be expanded to cover the entire field. Such an expansion enables the user to rank the prospective locations for further data collection and, from this ordered list, along with other criteria that are not expressed, to select the sites.

Rouhani (1983) shows that this response function, which represents the amount of information gain, can be written as:

$$VR_0* = \frac{1}{V*(N)} [K*_0 - \sum_{i=1}^{N} \lambda_i* K_{i0} - \sum_{p=1}^{l(k)} \mu_p* f_p(X_0)]^2 \tag{5}$$

where VR_0* is the variance reduction at X_0 due to a measurement at $X*$; $V*(N)$ is the estimation variance at $X*$ prior to the new measurement; $K*_0$ is the covariance function between $X*$ and X_0; λ_i* is the optimal weight of $Z(X_i)$ in estimation of $Z(X*)$ prior to the new measurement; μ_p* is the Lagrange multiplier for the pth monomial constraint in the kriging system for the estimation of $Z(X*)$ prior to the new measurement; and N is the number of existing data points prior to the new measurement.

In this work, estimation of the covariance function is accomplished using the structural analysis proposed by Delfiner (1975). The suggested polynomial covariance function has the following form:

$$K(h) = C\,\delta(h) + \sum_{p=0}^{k} a_{2p+1} h^{2p+1} \tag{6}$$

where h is the length of vector distance between any two points; C is the point variance; δ is the Dirac delta function; and k is the order of polynomial expected value.

The following definitions are useful:

$$\begin{aligned} \text{TOTV} &= \sum_j \text{V}_j \\ \text{TVR}* &= \sum_j \text{VR}_j* \end{aligned} \tag{7}$$

where TOTV is the total variances of estimation; TVR$*$ is the total variance reduction due to a measurement at $X*$; and j is the set of estimated points. In variance reduction analysis, at each round of sampling the site among potential sampling locations with maximal TVR$*$ is selected as the next measurement location. This yields a sequence of n points among m sites for further sampling. Equations (5) and (7) show that the above sampling scheme depends on the location of points and the assumed covariance function.

Proposed methodology

For planning a sampling activity, one may assume that the estimated covariance function remains unchanged as new data are collected. This yields an off-line or non-sequential ranking of n points for further sampling, or a ranking which is invariant with respect to the acquisition of new data. The rank list can be used as a shopping list; we utilize it, from the top down, until the budget is exhausted or some information criterion is met.

To study the resilience of variance reduction analysis, we compare these non-sequentially selected sites to points selected by a sequential procedure, in which the point $X*$ with maximal TVR$*$, is selected as the next new added site at each round of sampling; however, the new measured value $Z(X*)$ is assumed to be over- or underestimated by a random additive term or white noise perturbation:

$$Z^{i-1}(X*) = \hat{Z}^i(X*) \pm t_z[V^i(X*)]^{1/2} \tag{8}$$

where $Z^{i+1}(X*)$ is the $(i + 1)$th added measured value located at $X*$; $\hat{Z}^i(X*)$ is the estimated value at $X*$ based on K^i; $V^i(X*)$ is the kriging variance at $X*$ based on K^i; K^i is the estimated covariance function at the ith round of sampling; t_z is a standardized normally distributed random variable with $Pr(t \leqslant t_z) = \alpha$; and α is the level of deviation.

After adding the perturbed value to the data set a new structural analysis is conducted to estimate K^{i+1}. This procedure is sequential in the sense that at each round of sampling the statistical structure of the field is re-evaluated to accommodate the perturbation term.

Three main schemes are defined to generate perturbed inputs. In the first, it is assumed that all the new measured values are smaller than their estimated levels: the overestimated scheme (O). In the second, all new measured values

are larger than their corresponding predicted levels: the underestimated scheme (U). Finally, in the third scheme the added measured values alternatively vary around their estimated values: the sinusoidal scheme (S). These schemes can be shown to be equivalent to:

Type of scheme	Simulated measured values
O	$Z^{i+1} = \hat{Z}^i - t_\alpha (V^i)^{1/2}$
U	$Z^{i+1} = \hat{Z}^i + t_\alpha (V^i)^{1/2}$
S	$Z^{i+1} = \hat{Z}^i + (-1)^{i+1} t_\alpha (V^i)^{1/2}$

Each of these is simulated under three levels of deviation as follows:

Level of deviation (%):	t_α:
70	0.525
90	1.280
99	2.327

Consequently, nine cases of noisy inputs are generated, each of which can be identified by its type of scheme and its level of deviation (e.g., O–70).

The data set used in this study is described in Rouhani (1983). The existing data points are 84 spatially distributed values of piezometric heads measured in wells in northwest Kansas during January, 1979. For further information about the geohydrology of this region, readers are referred to Pearl et al. (1972). Their values and locations are given in Table 1. This area is divided into a 5×5 grid with $\Delta x = 8$ miles and $\Delta y = 16$ miles. The nodes are defined as potential sampling sites as shown in Fig. 1.

ANALYSIS OF RESULTS

Nine studies of sequential data collection planning are conducted. Each is characterized by a scheme type and a level of deviation. Some of the generated data with large perturbations might be unrealistic. For example, in scheme U large additions to Z might yield a water table significantly higher than the ground level. These values are included in this study to test the reliability of the proposed algorithm under some extreme, unexpected or counter-expected events (Fiering and Kindler, 1981).

In general, the effects of noisy data are accommodated by adjustments to the parameters of the covariance function (i.e., the parameter space). This is akin to an absorptive process whose mechanism can be described as follows: when the level of deviation in the added data is small, the structural analysis considers it simply as noise (e.g., measurement error). Consequently, the chaotic component of the covariance function [C in eqn. (6)] is selected for absorption of the simulated noise. As the flow of low level noisy data continues the chaotic component becomes stronger.

On the other hand, when the level of deviation is large, the structural

TABLE 1

Existing data locations and values (1 mile = 1.609 km; 1 ft. = 0.3048 m)

Point	y (mile) down	x across	$Z(x, y)$ (ft.)	Point	y (mile) down	x across	$Z(x, y)$ (ft.)
(1)	1.18	4.16	3239.00	(43)	32.02	7.35	3356.00
(2)	1.93	7.08	3196.00	(44)	32.27	7.84	3349.00
(3)	3.31	9.95	3175.00	(45)	33.16	6.81	3372.00
(4)	6.06	8.43	3205.00	(46)	31.04	0.70	3445.00
(5)	4.18	0.33	3295.00	(47)	31.16	5.08	3391.00
(6)	5.94	2.11	3292.00	(48)	32.54	1.84	3443.00
(7)	7.44	0.81	3308.00	(49)	36.55	4.33	3433.00
(8)	12.58	7.62	3244.00	(50)	43.17	9.84	3228.00
(9)	12.20	3.46	3312.00	(51)	17.10	29.52	3039.00
(10)	13.33	1.57	3339.00	(52)	21.61	29.79	3050.00
(11)	14.83	2.16	3341.00	(53)	17.85	27.41	3060.00
(12)	48.94	26.87	3099.00	(54)	17.98	27.90	3045.00
(13)	46.55	13.52	3219.00	(55)	18.74	24.54	3101.00
(14)	46.67	13.14	3216.00	(56)	16.97	21.90	3112.00
(15)	61.59	22.49	2912.00	(57)	20.36	22.00	3146.00
(16)	2.18	26.65	2886.00	(58)	16.47	11.52	3219.00
(17)	3.43	24.38	2926.00	(59)	17.46	15.95	3189.00
(18)	1.55	21.35	2973.00	(60)	20.86	11.89	3263.00
(19)	0.53	14.27	3128.00	(61)	21.85	15.19	3229.00
(20)	0.92	11.25	3161.00	(62)	23.37	29.90	3041.00
(21)	6.82	29.30	2981.00	(63)	23.13	24.49	3124.00
(22)	8.82	31.03	2972.00	(64)	26.12	29.41	3055.00
(23)	9.70	24.33	3039.00	(65)	26.75	28.11	3065.00
(24)	4.92	16.92	3092.00	(66)	24.00	20.49	3169.00
(25)	6.69	22.22	2986.00	(67)	24.99	12.38	3273.00
(26)	8.94	18.16	3068.00	(68)	27.63	15.41	3252.00
(27)	4.06	10.22	3176.00	(69)	29.13	30.28	3041.00
(28)	5.93	12.60	3172.00	(70)	28.26	28.76	3056.00
(29)	14.34	29.73	3023.00	(71)	28.77	23.52	3124.00
(30)	10.08	27.36	3021.00	(72)	31.77	27.63	3062.00
(31)	14.47	25.09	3072.00	(73)	30.13	12.00	3302.00
(32)	11.95	19.68	3099.00	(74)	38.65	29.84	3017.00
(33)	13.46	12.33	3200.00	(75)	35.39	29.36	3043.00
(34)	16.47	8.38	3261.00	(76)	43.79	29.57	3057.00
(35)	17.09	5.51	3294.00	(77)	40.15	29.09	3015.00
(36)	18.97	7.79	3288.00	(78)	41.54	28.60	3038.00
(37)	17.22	2.00	3354.00	(79)	42.55	27.46	3076.00
(38)	19.48	0.87	3391.00	(80)	45.59	16.87	3189.00
(39)	22.24	7.41	3318.00	(81)	52.10	4.60	3229.00
(40)	22.86	6.38	3336.00	(82)	61.94	13.46	3011.00
(41)	23.36	9.89	3290.00	(83)	39.69	19.30	3173.00
(42)	24.37	6.76	3337.00	(84)	25.75	32.06	3026.00

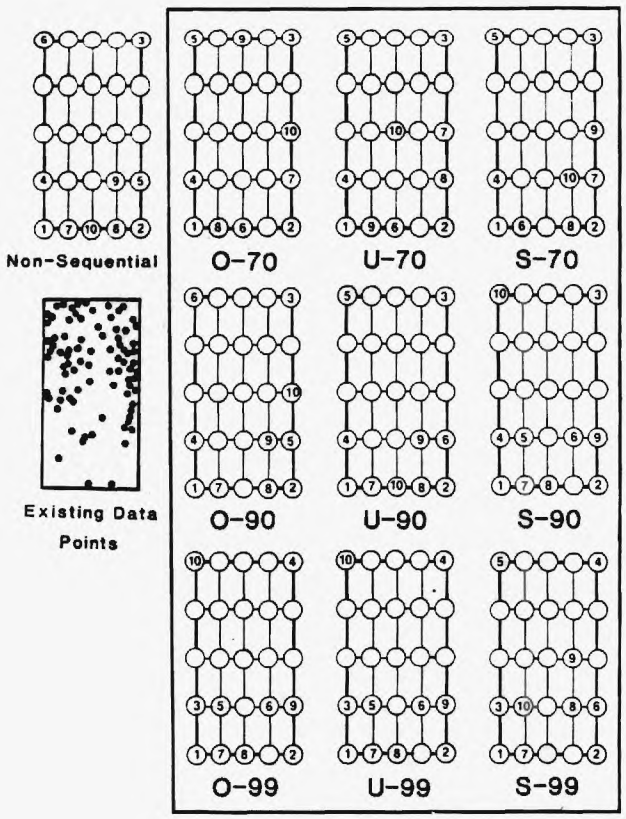


Fig. 1. Selected sets by sequential and non-sequential sampling schemes along with the set of existing data points.

analysis does not consider it only as a measurement error. Instead, it responds by trying to re-evaluate the over-all correlation structure of the field. As a result, the correlated part of eqn. (6) takes the burden of reflecting these added perturbations. However, as the flow of added data continues the chaotic components also increase. An interpretation of this is that there might be measurement errors superimposed on the potentially stronger correlation structure.

The amount of added noise might in fact cause the parameters of the covariance function to go up so much that the actual TVR (i.e., information gain) of the added data becomes negative. In such instances, the addition of noisy data deteriorates the predicted reliability of the estimated field. This phenomenon, dilution of good information with weakly correlated surrogates, was first studied systematically by Fiering (1962).

Effects of unstable parameters on the action space

The action space consists of selected sets of the new added sampling sites. As discussed earlier, these sets are directly related to parameters of the covariance

TABLE 2

Comparative rankings of the selected points of the non-sequential sampling

Rank				No. of sets containing the point
Non-sequential case	Sequential cases			
	High	Median	Low	
1	1	1	1	9
2	2	2	2	9
3	3	3	4	9
4	3	4	4	9
5	5	7	4	9
6	5	5	9	9
7	6	7	10	9
8	8	8	9	3
9	6	8	8	7
10	6	8	10	6

function. To study the effects of the unstable parameters on the system's action space, the sequential sets and the original non-sequential set are contrasted. This provides an estimate of the resilience of the variance reduction analysis.

Figure 1 illustrates the patterns of the selected sets for all nine cases plus the original (non-sequential) set. Table 2 compares the ranking of the first ten points of the original set to these same points in other sets. All sets share the first seven points of the original set, with the ranking patterns of the first four points being similar in all sets. The last three points are not included in all sets. For instance, point 8 is eliminated in six of nine cases. These patterns reveal an exchange mechanism within the action space, initiated as a function of variations in parameter space. The operational guideline is as follows: When the chaotic component of the covariance function overshadows the correlated part, there is a tendency to select more boundary nodes, and conversely.

All of the selected sequences reveal strongly similar central tendencies, with the points generally selected on either side of the field. In contrast to the existing data set, all the selected sets favor the sparsely sampled part of the region.

Regret analysis

In this section the sequential sets and the original set are compared using regret analysis. Suppose that at each round of sequential ranking, the estimated K^i is the true representative of the state of nature θ, as defined by Marin (1983). The decision set D^* (i.e., the sequentially selected sequence) is therefore a function of θ. The regret $\mathrm{IR}(D,D^*)$ is the incremental loss (in information gain) incurred by taking a non-optimal action D (i.e., the original set), instead of D^*:

TABLE 3

Average information regrets for non-sequential sampling in ft.^{2} (values in parentheses correspond to average percentage regrets; 1 ft.^{2} = 0.0929 m^2)

Level of deviation	Scheme Type			Avg.
	O	U	S	
70	45.1	90.2	12.7	49.3
	(0.8)	(1.3)	(0.2)	(0.7)
90	3.5	5.0	431.8	146.8
	(0.1)	(~0)	(4.2)	(1.4)
99	1663.7	3175.3	68.3	1635.8
	(4.2)	(3.7)	(0.1)	(2.7)
Avg.	570.8	1090.2	170.9	610.6
	(1.7)	(1.7)	(1.5)	(1.6)

$$\mathrm{IR}(D, D^*) = \mathrm{TOTV}(D, \theta) - \mathrm{TOTV}(D^*, \theta) \quad (9)$$

The average values of IR are shown in Table 3. As expected, the average regret increases with the level of deviation. The striking fact is that the values of regret as percentages of the sequential TOTV(D^*, θ) are negligible; their average is only 1.6%. It seems that use of the original set instead of the sequential set causes an insignificant loss in information.

The above result is in fact a case of near-optimality. Harrington (cf. Matalas and Fiering, 1977) gives an example of near-optimality which has many characteristics similar to those of our problem. In his study four treatment plants were to be built over a number of years to meet growing municipal demands. The least-cost solution is identified, but eleven other solutions generated at random fall within 3.3% of that minimal cost, well within the anticipated noise.

Further examples of near-optimality in water resources can be found in Harrington and So (1978), Gidley (1981) and Rogers and Fiering (1983). Near-optimality implies that, in spite of drastic differences between the estimated covariance functions of the sequential cases and the K^0 of the non-sequential case, the resulting actions D^* and D are so closely similar that their differences are insignificant.

UNSTABLE PARAMETER SPACE VERSUS RESILIENT ACTION SPACE

Comparison between the sequential sets and the original set shows that the parameter space is very sensitive to perturbations in the data set. Even slight levels of simulated noise in the input data cause significant changes in the general pattern of the estimated covariance function. On the contrary, the instability of the parameter space has a negligible effect on the action space. Actual results reveal the following patterns in the behavior of the parameter space and the action space.

(1) When the level of noise is low, kriging treats it primarily as measurement

error. Consequently, the structural analysis produces covariance functions with larger chaotic components. In such instances the priorities are further shifted towards border nodes.

(2) When the level of noise is high, kriging considers it an indication of error caused by an underestimated correlation function. As a result, the correlated part of the covariance function gets stronger to reflect the more unreliable results. This in turn causes an increase in the influence of the internal nodes on their neighboring points. Consequently, internal points become more advantageous as sampling sites.

(3) Despite the large amount of simulated noise, all selected sequences show a great degree of similarity. Furthermore, a regret analysis shows a case of near-optimality among all selected sets.

It can be concluded that the variance reduction analysis is a reliable method with an unstable parameter space but a resilient action space.

ACKNOWLEDGEMENTS

This research was supported in part by grant ECE-8503897 from the National Science Foundation.

REFERENCES

Bastin, G., Lorent, B., Duque, C. and Gevers, M., 1984. Optimal estimation of the average rainfall and optimal selection of rain gauge locations. Water Resour. Res., 20(4): 463–470.

Brady, P.J., 1978. Optimal sampling and analysis using two variables and modeled cross-covariance functions. J. Appl. Meteorol., 17: 12–21.

Chou, D. and Scheck, D.E., 1984. Selecting optimum drilling locations by groups. Paper presented at the SME-AIME Annual Meeting, Los Angeles, Calif.

Crawford, K.C., 1979. Considerations for the design of a hydrologic data network using multivariate sensors. Water Resour. Res., 15(6): 1752–1762.

Delfiner, P., 1975. Linear estimation of non-stationary spatial phenomena. In: Guaraseio et al. (Editors), Advanced Geostatistics in Mining Industry. Nato Advanced Study Inst., Reidel, Boston, Mass.

Fiering, M.B., 1962. On the use of correlation to augment data. J. Am. Stat. Assoc., 57: 20–32.

Fiering, M.B., 1965. An optimization scheme for gaging. Water Resour. Res., 1(4): 463–470.

Fiering, M.B., 1982. A screening model to quantify resilience. Water Resources Res., 18(1): 27–32.

Fiering, M.B. and Kindler, J., 1981. Surprise! in Water Resources Design. IIASA, Laxenburg, Austria.

Gidley, J.S., 1981. Nearly optimal decisions in water resources planning. Ph.D. thesis, Division of Applied Sciences, Harvard Univ., Cambridge, Mass.

Harrington, J.J. and So, S.K., 1978. Near-optimality in Capital-investment planning for water resources. In: Fiering, M.B., et al. (Editors), Standards, Optimality, and Resilience in Water-Resource Management. Division of Applied Sciences, Harvard Univ., Cambridge, Mass.

Hughes, J.P. and Lettenmaier, D.L., 1981. Data requirement for kriging: Estimation and network design. Water Resources Res., 17(6): 1641–1650.

Kemeny, J.G. and Snell, J.L., 1962. Mathematical Models in the Social Sciences. Blaisdell, New York, N.Y.

Kitanidis, P.K. and Vomvoris, E.G., 1982. A geostatistical approach to the inverse problem in groundwater modeling (steady state), 1. Institute of Hydraulic Research, Univ. Iowa, Iowa City, Iowa.

Marin, C., 1983. Parameter uncertainty in water resource planning. Ph.D. thesis, Division of Applied Sciences, Harvard Univ., Cambridge, Mass.

Matalas, N.C., 1968. Optimum gaging station location. Paper presented at Proc. IBM Sci. Comput. Symp., Water and Air Resources Management, IBM, White Plains, N.Y.

Matalas, N.C. and Fiering, M.B., 1977. Water-resources systems planning. Climate, Climatic Changes, and Water Supply, N.A.S., Washington, D.C.

Pearl, R.H., Roberts, R.S., Keene, K.M. and McClain, T.J., 1972. Water resources of northwest Kansas. Hydrologic Investigations, Atlas HA-429, Washington, D.C.

Rogers, P.P. and Fiering, M.B., 1983. Use of System Analysis in Water Management in Developing Countries. Harvard Univ., Cambridge, Mass.

Rouhani, S., 1983. Optimal data collection in random fields. Ph.D. Thesis, Division of Applied Sciences, Harvard Univ., Cambridge, Mass.

Rouhani, S., 1985. Variance reduction analysis. Water Resour. Res., 21(6): 837–846.

Water Resources Monitoring:

A Combined Information-Economic Approach

Shahrokh Rouhani, Ph.D.
Assistant Professor
School of Civil Engineering
Georgia Institute of Technology
Atlanta, GA 30332 U.S.A.

Synopsis

Water resources management demands an efficient strategy for sampling activities. This policy involves two conflicting objectives, which are the information accuracy and the economic efficiency. Water experts have traditionally used approaches which emphasize one objective, while ignore the other. The author proposes a combined information-economic procedure on the basis of the above conflicting goals. Variance Reduction Analysis, a statistical algorithm, is utilized to quantify the information gain due to a new measurement. A loss function is then defined to convert the above gain function into a monetary value. This method is applied to a ground water monitoring problem, and its efficiency is illustrated by comparing it to a simple plan based on the criterion of maximum distance.

ERC 03-87 March 1987

OPTIMAL SCHEMES FOR GROUND WATER QUALITY MONITORING IN THE SHALLOW AQUIFER, DOUGHERTY PLAIN, SOUTHWESTERN GEORGIA

by

SHAHROKH ROUHANI, Ph.D.
Assistant Professor
of Hydrology and Water Resources

and

TIMOTHY J. HALL
Graduate Research Assistant

Technical Completion Report
USDI/USGS Project G-1219 (05)

The research on which this report is based was financed in part by the U.S. Department of the Interior, Geological Survey as authorized by the Water Resources Research Act of 1984 (P.L. 98-242)

Contents of this publication do not necessarily reflect the views and policies of the U.S. Department of the Interior, nor does mention of trade names or commercial products constitute their endorsement or recommendation for use by the U.S. Government.

School of Civil Engineering
in cooperation with
Environmental Resources Center
Georgia Institute of Technology
Atlanta, Georgia 30332

ABSTRACT

Geostatistical schemes for ground water quality monitoring in the shallow aquifer of Dougherty Plain, Georgia are presented. This aquifer is not generally used for water supply purposes. However, it is the main recharge route to the principal artesian aquifer which is the primary source of water supply in this rapidly growing agricultural region. The desired monitoring network acts as an early warning system for ground water pollution in deeper layers. We have utilized the available data on hydraulic properties of the shallow aquifer to identify the zones which should be the primary locations for our sampling activities. The one variable which appears to be most suitable for our study is leakance. Statistical analyses indicate that leakance has a log-normal distribution with a constant trend and a linear covariance function. Ranking criteria for the selection of the best sampling points are: the variance reductions, the medians, and the risk values. Due to the nature of our monitoring network we suggest to use mainly risk ranking as the basis of our sampling activities. The results of our risk rankings demonstrate that the southern tip of the Dougherty Plain and its upper central zone should be the prime targets of our monitoring activities.

Keywords: Network Design, Statistical Methods, Regional Analysis, Water Quality, Water Management (Applied), Georgia.

GEOSTATISTICAL SCHEMES FOR GROUND WATER QUALITY MONITORING IN SOUTHWEST GEORGIA

by

Shahrokh Rouhani and Timothy J.Hall

School of Civil Engineering

Georgia Institute of Technology

Atlanta, Georgia 30332

ABSTRACT

Regional schemes for shallow ground water quality monitoring in southwest Georgia are presented. The aquifer of concern is not generally used for water supply purposes. However, it is the main recharge route to the lower principal artesian aquifer which is the primary source of water supply in this rapidly growing agricultural region. The desired monitoring network acts as an early warning system for ground water pollution in deeper layers. We have utilized the available data on hydraulic properties of the shallow aquifer to identify the zones which should be the primary locations for our sampling activities. The one variable which appears to be most suitable for our study is leakance. Statistical analyses indicate that leakance has a log-normal distribution with a constant trend and a linear covariance function. Ranking criteria for the selection of the best sampling points are: the variance reductions, the medians, and the risk values. Due to the nature of our monitoring network we suggest to use the risk ranking as the basis of our sampling activities. The results of our risk rankings demonstrate that the southern tip of the study area and its upper central zone should be the prime targets of our monitoring activities.

GEOSTATISTICAL SCHEMES FOR GROUNDWATER SAMPLING

by

SHAHROKH ROUHANI, Ph.D.

Assistant Professor

of Hydrology and Water Resources

and

TIMOTHY J. HALL

Graduate Research Assistant

School of Civil Engineering

Georgia Institute of Technology

Atlanta, Georgia 30332

ABSTRACT

Geostatistical techniques offer efficient tools for design of ground water sampling networks. They include procedures for the selection of the best sequences of sampling points, such as: variance reduction analysis, median ranking, and risk ranking. Variance reduction analysis considers primarily the accuracy of the estimated field, while median ranking is based only on the magnitude of the estimated values. Risk ranking is a compromise between these procedures that appears to yield a more balanced guideline for cases when planners desire to acquire maximum information, while monitoring areas where the variable of interest exhibits critical values. These procedures are used for the design of a regional shallow groundwater quality monitoring network in the Dougherty Plain, located in southwest Georgia. The shallow aquifer of concern is the main recharge route to a semi-confined aquifer which is the primary source of water in this region. The desired monitoring network acts as an early warning system for groundwater pollution in deeper layers. Leakance data is utilized to identify the primary sampling locations. Statistical analyses indicate that leakance has a log-normal distribution with a constant drift and a linear spatial covariance. The results of our risk rankings demonstrate that the southern tip of the Dougherty Plain and its upper central zone should be the prime targets of our monitoring activities.

Space-Time Kriging Analysis of Groundwater Data

by

Shahrokh Rouhani and Timothy J. Hall

School of Civil Engineering
Georgia Institute of Technology
Atlanta, Georgia 30332
USA

Abstract

A significant number of naturally occurring processes and parameters can be described as stochastic processes. These processes can be mapped by using Gauss-Markov estimators, such as kriging. Presently most kriging packages are designed for estimation of spatially random variables. It is shown that with certain modifications, kriging can be expanded to the space-time domain to be applicable to a more general class of stochastic processes. This is analogous to combining spatial kriging with time series analysis. In this study a series of groundwater elevation data from southern Georgia is simultaneously analyzed in time and space, using universal kriging, in the framework of intrinsic random functions with polynomial generalized covariances. The results are presented in a series of spatial maps for different time periods. In this way valuable new information has been gained by utilizing both the spatial and the temporal data. This new procedure yields more precise estimates of covariance functions, as well as, more accurate spatial maps. It also allows hindcasting and forecasting for periods when no sampling is conducted.

APPENDIX 5. COMPUTER PROGRAMS

The following sections include the listing of a number of selected programs, developed by the PI in the course of this project. These programs are written in Fortran.

1. VARED: Spatial Universal Kriging with variacne reduction analysis option;
2. TKRIG: Temporal Universal Kriging; and
3. STVARED: Universal Space-Time Kriging with varaince reduction analysis.

VARED

```
C*****************************************************************
C***** UNIT 99 =INPUT DATA,UNIT 98=ZKRIGE OUTPUT
C***** UNIT 97=VARSET OUTPUT
C***** UNIT 96 & 95 =I/O TERMINAL
      PROGRAM VARED(INPUT,OUTPUT,VAR,ZKR,ZRISK,DAT,GRID,TAPE4=ZRISK
     &,TAPE5=INPUT,TAPE6=OUTPUT,TAPE7=VAR,TAPE8=ZKR
     &,TAPE9=DAT,TAPE10=GRID)
C.......NEW MODIFICATIONS.....................................
C....UNIT 95 ==> UNIT 5 TERMINAL INPUT (OR DATA INPUT)
C....UNIT 99 ==> UNIT 5 TERMINAL INPUT
C....UNIT 96 ==> UNIT 6   TERMINAL OUTPUT
C....UNIT 97 ==> UNIT 7   OUTPUT FILE = VAR
C....UNIT 98 ==> UNIT 8   OUTPUT FILE = ZKR
C*****************************************************************
      DIMENSION FORM(4),KM(15),CM(15),A1M(15),TVARED(100,100)
     &,A3M(15),A5M(15),ERROR(15),RANK(15),T(4),X(4),ZKRIGE(100,100)
     $,VARSET(100,100),G(4,5),UU(100),VV(100),ZRISK(100,100)
     &,ZMED(100,100),UUU(100),VVV(100)
      COMMON/C3/U(100),V(100)
      COMMON/C5/A(106,107)
      COMMON/C6/Z(100)
      COMMON/C7/ILIST(100)
      COMMON/C8/RLIST(100)
      COMMON/C9/P(100)
C*****
      DATA STQ,ST1/1HQ,1H1/
      DATA STA,STB,STC,STD,STE,STQ/1HA,1HB,1HC,1HD,1HE,1HQ/
C***** READ INFORMATION ABOUT DATA POINTS
      WRITE(6,9001)
      READ(5,*)NDATA
      ONDATA=NDATA
      WRITE(6,9008)
      READ(5,*)DP
      READ(9,*)(V(I),U(I),Z(I),I=1,NDATA)
400   WRITE(6,9600)
      READ(5,*)XLOG
      IF (XLOG.EQ.0) GO TO 999
      IF (XLOG.EQ.1) GO TO 499
      GO TO 400
499   DO 500 I=1,NDATA
      Z(I)=LOG(Z(I))
500   CONTINUE
C***** CHECK FOR DOUBLE POINTS
999   CALL DOUBLE(NDATA,DP,NDOUBL)
C***** REARANGE DATA
      CALL SORT(NDATA)
C***** UPDATE NUMBER OF DATA POINTS
      NDATA=NDATA-NDOUBL
      CONTINUE
C***** READ THE NEW DATA POINTS
      ISEQ=1
      WRITE(6,9610)
      READ(5,*)NEW
      IF(NEW.EQ.0) GO TO 10
      DO 2 INEW=1,NEW
      WRITE(6,9611)
      NDATA=NDATA+1
      READ(5,*)U(NDATA),V(NDATA),Z(NDATA)
      IF(XLOG.EQ.1)Z(NDATA)=LOG(Z(NDATA))
      IF(NONREG.EQ.1.AND.NDATA.GE.(NNN+ONDATA))WRITE(6,9613)
      IF(NONREG.NE.1.AND.NDATA.GE.(ONDATA+(IUMAX+IVMAX)))WRITE(6,9613)
9613  FORMAT(" ALL ESTIMATION GRID POINTS ARE NOW DEFINED AS",/,
     &" MEASUREMENT POINTS.  FURTHER SAMPLING IS NOT POSSIBLE.",/,
     &" YOU MAY SELECT THE SEQUENTIAL SAMPLING AND THEN EXIT.")
2     CONTINUE
      CALL SORT(NDATA)
      WRITE(6,9612)
9612  FORMAT("*** IF SEQUENTIAL (ON-LINE) SAMPLING IS DESIRED",/,
     &       "     TYPE A 1, OTHERWISE TYPE A 0")
      READ(5,*)ISEQ
      IF(ISEQ.EQ.1) GO TO 10
      TOTV=0.
      TOTSD=0.
      GO TO 5001
9610  FORMAT("?  HOW MANY NEW DATA POINTS?")
9611  FORMAT("?  NEW U, V, AND Z")
10    CONTINUE
      TOTV=0.
      TOTSD=0.
C****** OPTION SELECTION
5001  WRITE(6,9004)
      READ(5,9005)ST
      IF(ST.EQ.STA)GO TO 1000
      IF(ST.EQ.STB)GO TO 2000
      IF(ST.EQ.STC)GO TO 3000
      IF(ST.EQ.STD)GO TO 4000
      IF(ST.EQ.STE)GO TO 5000
      IF(ST.EQ.STQ)STOP
      GO TO 1
1000  CONTINUE
C**** OPTION A ***************************************************
      WRITE(6,9006)STA
      WRITE(6,9107)
      READ(5,*)NG,NGC,ISTEP
      IF(ISTEP.EQ.0)ISTEP=1
C**** READ K,C,A1,A3,A5 FOR EACH GENERALIZED COVARIANCE
      DO 1050 IGC=1,NGC
      WRITE(6,9109)IGC,IGC,IGC,IGC,IGC
      READ(5,*)KM(IGC),CM(IGC),A1M(IGC),A3M(IGC),A5M(IGC)
1050  CONTINUE
      DO 1060 IGC=1,NGC
      RANK(IGC)=0.
1060  CONTINUE
      NGI=0
C**** START ITERATION ON DATA POINTS  IO IS
C**** THE TYPICAL DATA POINT
      DO 1200 IO=1,NDATA,ISTEP
      NGI=NGI+1
C**** START ITERATION ON GENERALIZED COVARIENCES - USE EACH
C**** ONE IN TURN TO KRIGE POINT IO
      DO 1100 IGC=1,NGC
      CALL FIND1(NDATA,IO,U(IO),V(IO),NO)
C**** KRIGE POINT IO
      CALL  KRIGPO(KM(IGC),CM(IGC),A1M(IGC),A3M(IGC),A5M(IGC),
     &U(IO),V(IO),NO)
      NEQN=NO+(KM(IGC)+1)*(KM(IGC)+2)/2
      CALL ELIMIN(A,P,NEQN,106,107)
      SUM=-Z(IO)
      DO 1080 L=1,NO
      I=ILIST(L)
      SUM=SUM+P(L)*Z(I)
1080  CONTINUE
C**** FIND KRIGING ERROR OF CURRENT GENERALIZED COVARIENCE
      ERROR(IGC)=ABS(SUM)
1100  CONTINUE
C**** END OF ITERATION ON GENERALIZED COVARIANCES
```

```
1190  CONTINUE
1200  CONTINUE
C***** END OF ITERATION ON DATA POINTS
      DO 1210 IGC=1,NGC
      RANK(IGC)=RANK(IGC)/FLOAT(NGI)
1210  CONTINUE
C***** WRITE RESULTS
      DO 1220 IGC=1,NGC
      WRITE(6,9111)IGC,RANK(IGC)
1220  CONTINUE
      GO TO 1
2000  CONTINUE
C
C***** OPTION B ************************************************************
C
      WRITE(6,9006)STB
      WRITE(6,9301)
      READ(5,*)NC,ISTEP
      IF(ISTEP.EQ.0)ISTEP=1
      WRITE(6,9303)
      READ(5,*)K,C,A1,A3,A5
      WRITE(6,9205)
      READ(5,9206)(FORM(I),I=1,4)
2010  CONTINUE
C***** SET MATRIX G EQUAL TO ZERO - MATRIX G IS THE AUGMENTED
C***** MATRIX OF THE SET OF EQUATIONS APPEARING IN TABLE 3.1 OF
C***** CHAPTER 3
      KP2=K+2
      KP3=K+3
      DO 2020 IROW=1,KP2
      DO 2020 ICOL=1,KP3
      G(IROW,ICOL)=0.
2020  CONTINUE
C**** START ITERATION ON DATA POINTS - IO IS THE TYPICAL DATA POINT
C***** FIND THE NO NEARSET DATA POINTS TO DATA POINT IO
      DO 2200 IO=1,NDATA,ISTEP
      CALL FINDI(NDATA,IO,U(IO),V(IO),NO)
C***** KRIGE POINT IO USING INITIAL  K,C,A1,A3,A5
      CALL KRIGP0(K,C,A1,A3,A5,U(IO),V(IO),NO)
      NEQN=NO+(K+1)*(K+2)/2
      CALL ELIMIN(A,P,NEQN,106,107)
C***** FILL IN THE VECTOR T-T(1),T(2),T(3),T(4) ARE GIVEN BY THE
C***** LAST FOUR EQUATIONS OF TABLE 3.2 OF CHAPTER 3- NOTICE THE
C***** CHAGE OF NOTATION
      T(1)=1.
      DO 2050 L=1,NO
      T(1)=T(1)+P(L)**2
2050  CONTINUE
      KP2=K+2
      DO 2060 N=2,KP2
      NEXP=2*(N-2)+1
      T(N)=0.
      DO 2060 L1=1,NO
      I1=ILIST(L1)
      H=(U(IO)-U(I1))**2+(V(IO)-V(I1))**2
      H=SQRT(H)
      T(N)=T(N)-2.*P(L1)*H**NEXP
      DO 2060 L2=1,NO
      I2=ILIST(L2)
      H=(U(I1)-U(I2))**2+(V(I1)-V(I2))**2
      H=SQRT(H)
      T(N)=T(N)+P(L1)*P(L2)*H**NEXP
2060  CONTINUE
C***** UPDATE MATRIX G
      KP2=K+2
      DO 2090 IROW=1,KP2
      DO 2090 ICOL=1,KP2
      G(IROW,ICOL)=G(IROW,ICOL)+T(IROW)*T(ICOL)
2090  CONTINUE
C***** CONSTRUCT GENERALIZED INCREMENT
      GINCR=-Z(IO)
      DO 2100 L=1,NO
      I=ILIST(L)
      GINCR=GINCR+P(L)*Z(I)
2100  CONTINUE
      GINCR2=GINCR**2
C**** UPDATE LAST COLOUMN OF MATRIX G,I.E. THE RIGHT HAND SIDE
C***** VECTOR OF THE SET OF EQUATIONS OF TABLE 3.2 OF CHAPTER 3
      KP2=K+2
      DO 2110 IROW=1,KP2
      G(IROW,K+3)=G(IROW,K+3)+T(IROW)*GINCR2
2110  CONTINUE
2200  CONTINUE
C***** END OF ITERATION ON DATA POINTS -AT THIS STAGE
C***** MATRIX G HAS BEEN FILLED IN -PROCEED TO TAKE INTO ACCOUNT THAT
C***** SOME OF THE COEFFICIENTS C,A1,A3,A5 MAY HAVE BEEN SET EQUAL
C***** TO ZERO A PRIORI
      KP2=K+2
      DO 2240 IROW=1,KP2
      IF(FORM(IROW).EQ.ST1)GO TO 2240
C     KP2=K+2
      DO 2220 ICOL=1,KP2
      G(IROW,ICOL)=0.
      G(ICOL,IROW)=0.
2220  CONTINUE
      G(IROW,K+3)=0.
      G(IROW,IROW)=1.
2240  CONTINUE
      NEQN=K+2
C***** SOLVE FOR C, A1,A3,A5
      CALL ELIMIN(G,X,NEQN,4,5)
      C=0.
      A1=0.
      A3=0.
      A5=0.
      IF(FORM(1).EQ.ST1) C=X(1)
      IF(FORM(2).EQ.ST1)A1=X(2)
      IF(FORM(3).EQ.ST1)A3=X(3)
      IF(FORM(4).EQ.ST1)A5=X(4)
      WRITE(6,9207)K,C,A1,A3,A5
C***** CHECK IF THIS IS A PROPER GENERALIZED COVARIANCE -
C***** WRITE  COEFFICIENTS - ASK IF USER WANTS TO START AGAIN USING
C***** THE NEWLY FOUND GEN. COVARIANCE AS INITIAL GEN. COVARIANCE
      CALL CHECK(C,A1,A3,A5,INDEX)
      IF(INDEX.EQ.1)WRITE(6,9208)
2280  CONTINUE
      WRITE(6,9209)
      READ(5,9005)ST
      IF(ST.EQ.ST1)GO TO 2010
      IF(ST.EQ.ST2)GO TO 1
      GO TO 2280
3000  CONTINUE
C
C***** OPTION C ************************************************************
C
      WRITE(6,9006)STC
      WRITE(6,9301)
      READ(5,*)NC,IST P
      IF(IST P.EQ.0)ISTEP=1
      WRITE(6,9303)
      READ(5,*)K,C,A1,A3,A5
```

```
      IMIN=2+(NA-1)*ISTEP
      IMAX=NDATA
      CALL FIT(NDATA,IMIN,IMAX,ISTEP,N0,K,C,A1,A3,A5,
     &SUM1B,SUM2B,NB)
      R=(SUM1A+SUM1B)/(SUM2A+SUM2B)
      RA=SUM1A/SUM2A
      RB=SUM1B/SUM2B
      R=2*R-(NA*RA+NB*RB)/(NA+NB)
      WRITE(6,9305)NA,RA,NB,RB,R
      GO TO 1
000   CONTINUE

***** OPTION D ************************************************

      WRITE(6,9006)STD
      WRITE(6,9401)
      READ(5,*)THETA
      THETA=THETA*1.7453292E-2
      CALL ROTATE(THETA,NDATA)
      CALL SORT(NDATA)
      GO TO 1
000   CONTINUE

***** OPTION E ************************************************

      WRITE(6,9006)STE
      WRITE(6,9303)
      READ(5,*)K,C,A1,A3,A5
      WRITE(6,9802)
802   FORMAT("*** DO YOU WANT TO USE YOUR OWN GRID POINTS?",/,
     &       "     IF YES TYPE A 1",/,
     &       "     IF NO (REGULAR RECTANGULAR GRIDS) TYPE A 0")
      READ(5,*)NONREG
      IF(NONREG.NE.1) GO TO 5100
      READ(10,*)NNN,(UUU(I),VVV(I),I=1,NNN)
      GO TO 5101
100   WRITE(6,9501)
      READ(5,*)U0,V0,DU,DV,IUMAX,IVMAX,UINCR,VINCR
101   WRITE(6,9503)
      READ(5,*)ND,R0

***** THE OPTION FOR THE VARIANCE REDUCTION

      WRITE(6,9700)
700   FORMAT(" *****  IF VARIANCE REDUCTION ANALYSIS",/,
     &       "*****  IS DESIRED TYPE A 1",/,
     &       "*****  IS NOT DESIRED TYPE A 0")
      READ(5,*)IVRD
      WRITE(6,9801)
801   FORMAT(" *****  DO YOU ALSO DESIRE RISK VALUES?",/,
     &       "          IF YES TYPE IN THE RISK FOR",/,
     &       "          ?PISK=E(Z) + RISK * SQRT(V)",/,
     &       "          IF NO TYPE A 0")
      READ(5,*)RISK
      IF(NONREG.EQ.1) GO TO 5102
      UUU0=U0
      VVV0=V0
      DO 5700 IU=1,IUMAX
      DO 5700 IV=1,IVMAX
      UU(IU)=U0+FLOAT(IU-1)*UINCR
      VV(IV)=V0+FLOAT(IV-1)*VINCR
700   CONTINUE
001   IF(ISEQ.EQ.1)GO TO 5002
      IF(NONREG.NE.0)GO TO 5102
      U0=UUU0
      V0=VVV0
002   U0=U0-UINCR
      V0=V0-VINCR
***** KRIGE POINT OR BLOCK IU,IV -SEE FIGURE 5.2 OF CHAPTER 5
      DO 5600 IUU=1,IUMAX
      U0=U0+UINCR
      U1=U0-DU/2.
      U2=U0+DU/2.
      DO 5500 IVV=1,IVMAX
      V0=V0+VINCR
      V1=V0-DV/2.
      V2=V0+DV/2.
      IU=IUU
      IV=IVV
      IF(NONREG.NE.1) GO TO 5103
102   DO 5505 INR=1,NNN
      U0=UUU(INR)
      V0=VVV(INR)
      IU=INR
      IV=INR
103   DO 5501 III=1,NDATA
      IF(U0.NE.U(III).OR.V0.NE.V(III)) GO TO 5501
      ZKRIGE(IU,IV)=Z(III)
      VARSET(IU,IV)=C
      IVARED(IU,IV)=0.
      IF(RISK.NE.0)ZRISK(IU,IV)=Z(III)+RISK*C
      IF(NEW.NE.0.AND.RISK.NE.0)ZRISK(IU,IV)=-100.
      IF(NONREG.NE.0) GO TO 5505
      GO TO 5500
501   CONTINUE
      IF(R0.EQ.0.)GO TO 5200
      CALL FIND2(NDATA,0,U0,V0,R0,M0)
      IF(M0.LT.ND)GO TO 5200
      GO TO 5300
200   CONTINUE
      M0=N0
      CALL FIND1(NDATA,0,U0,V0,M0)
300   CONTINUE
      IF(DU.NE.0..AND.DV.NE.0.)CALL KRIGBL(K,C,A1,A3,A5,U1,U2,V1,V2,M0)
      IF(DU.EQ.0..OR.DV.EQ.0.)CALL KRIGP0(K,C,A1,A3,A5,U0,V0,M0)
      NEQN=M0+(K+1)*(K+2)/2
      CALL ELIMIN(A,P,NEQN,106,107)
      ZKRIGE(IU,IV)=0.
      VARSET(IU,IV)=0.
      DO 5350 L=1,M0
      I=ILIST(L)
      ZKRIGE(IU,IV)=ZKRIGE(IU,IV)+P(L)*Z(I)
350   CONTINUE
      IF(DU.NE.0..AND.DV.NE.0.)GO TO 5450
      DO 5400 L=1,M0
      I=ILIST(L)
      H=(U0-U(I))**2+(V0-V(I))**2
      H=SQRT(H)
      VARSET(IU,IV)=VARSET(IU,IV)-P(L)*GENCOV(K,C,A1,A3,A5,H)
400   CONTINUE
      VARSET(IU,IV)=VARSET(IU,IV)-P(M0+1)*C
      IF(K.EQ.0)GO TO 5450
      VARSET(IU,IV)=VARSET(IU,IV)-P(M0+2)*U0-P(M0+3)*V0
      IF(K.EQ.1)GO TO 5450
      VARSET(IU,IV)=VARSET(IU,IV)-P(M0+4)*U0*V0-P(M0+5)*U0**2-P(M0+6)
     &*V0**2
      IF(VARSET(IU,IV).GE.0.) GO TO 5450
      WRITE(6,9620)U0,V0,VARSET(IU,IV)
620   FORMAT(" NEGATIVE VAR AT U, V, & VAR",2F5.0,F10.1)
      VARSET(IU,IV)=-VARSET(IU,IV)
450   CONTINUE
      TOTV=TOTV+VARSET(IU,IV)
```

```
      TVARED(IU,IV)=0
      IF(NONREG.EQ.1) GO TO 5108
      DO 5451 IUUU = 1,IUMAX
      DO 5451 IVVV = 1,IVMAX
      VARED=0
      UU0=UU(IUUU)
      VV0=VV(IVVV)
      IF(NONREG.NE.1) GO TO 5107
5108  DO 5111 INRG=1,NNN
      VARED=0
      UU0=UUU(INRG)
      VV0=VVV(INRG)
5107  DO 5460 IIII=1,NDATA
      IF(UU0.NE.U(IIII).OR.VV0.NE.V(IIII)) GO TO 5460
      GO TO 5440
5460  CONTINUE
      CALL FIND1(NDATA,0,UU0,VV0,M0)
      IM0=ILIST(M0)
      RMAX=(UU0-U(IM0))**2+(VV0-V(IM0))**2
      RMAX=SQRT(RMAX)
      HR=(U0-UU0)**2+(V0-VV0)**2
      HR=SQRT(HR)
      IF(HR.GT.RMAX) GO TO 5440
      CALL FIND1(NDATA,0,U0,V0,M0)
      DO 5452 L=1,M0
      I=ILIST(L)
      DEP=(UU0-U(I))**2+(VV0-V(I))**2
      DEP=SQRT(DEP)
      VARED=VARED+P(L)*GENCOV(K,C,A1,A3,A5,DEP)
5452  CONTINUE
      VARED=VARED+P(M0+1)-GENCOV(K,C,A1,A3,A5,HR)
      IF(K.EQ.0) GO TO 5453
      VARED=VARED+P(M0+2)*UU0+P(M0+3)*VV0
      IF(K.EQ.1) GO TO 5453
      VARED=VARED+P(M0+4)*UU0*VV0+P(M0+5)*UU0**2
     &+P(M0+6)*VV0**2
5453  VARED=VARED**2/VARSET(IU,IV)
      TVARED(IU,IV)=TVARED(IU,IV)+VARED
5440  IF(NONREG.NE.1) GO TO 5451
5111  CONTINUE
      IF(NONREG.EQ.1) GO TO 5109
5451  CONTINUE
5109  IF(RISK.EQ.0) GO TO 5110
      ZRISK(IU,IV)=ZKRIGE(IU,IV)+RISK*SQRT(VARSET(IU,IV))
5110  IF(NONREG.NE.1) GO TO 5500
5505  CONTINUE
      IF(NONREG.EQ.1) GO TO 5510
5500  CONTINUE
      V0=V0-IVMAX*VINCR
5600  CONTINUE
C *****  THE MATRICES ZKRIGE, VARSET, AND ZRISK IN THE
C        CASE OF POINT KRIGING AND THE MATRIX ZKRIGE IN THE
C        CASE OF BLOCK KRIGING HAVE BEEN FILLED AND CAN BE
C        RETRIEVED FROM THIS LOCATION.
C
      IF(XLOG.EQ.0) GO TO 5698
      DO 5655 I=1,IUMAX
      DO 5655 J=1,IVMAX
      ZMED(I,J)=2.718282**(ZKRIGE(I,J))
      ZRISK(I,J)=2.718282**(ZRISK(I,J))
      ZKRIGE(I,J)=2.718282**(ZKRIGE(I,J)+(VARSET(I,J)/2))
5655  VARSET(I,J)=(ZKRIGE(I,J)**2)*((2.718282**VARSET(I,J))-1)
C
C****  IDENTIFYING THE BEST SAMPLING SITES AMONG
C      REGULAR GRID POINTS
C
5698  TVMAX=0
      VARMAX=0.
      RISKMAX=0.
      DO 5804 IU=1,IUMAX
      DO 5804 IV=1,IVMAX
      IF(IVRD.EQ.0) GO TO 5805
      TVMAX=MAX(TVMAX,TVARED(IU,IV))
5805  IF(RISK.EQ.0) GO TO 5806
      RISKMAX=MAX(RISKMAX,ZRISK(IU,IV))
5806  VARMAX=MAX(VARMAX,VARSET(IU,IV))
5804  CONTINUE
      DO 5807 IU=1,IUMAX
      DO 5807 IV=1,IVMAX
      IF(IVRD.EQ.0) GO TO 5808
      IF(TVMAX.NE.TVARED(IU,IV)) GO TO 5808
      WRITE(6,9706)UU(IU),VV(IV),ZKRIGE(IU,IV),VARSET(IU,IV)
     &,TVARED(IU,IV)
      IF(XLOG.EQ.1)WRITE(6,9709)ZMED(IU,IV)
5808  IF(RISK.EQ.0) GO TO 5809
      IF(RISKMAX.NE.ZRISK(IU,IV)) GO TO 5809
      WRITE(6,9707)UU(IU),VV(IV),ZKRIGE(IU,IV),VARSET(IU,IV)
     &,ZRISK(IU,IV)
      IF(XLOG.EQ.1)WRITE(6,9709)ZMED(IU,IV)
5809  IF(VARMAX.NE.VARSET(IU,IV)) GO TO 5807
      WRITE(6,9708)UU(IU),VV(IV),ZKRIGE(IU,IV),VARSET(IU,IV)
     &,TOTV,TOTSD
      IF(XLOG.EQ.1)WRITE(6,9709)ZMED(IU,IV)
5807  CONTINUE
9706  FORMAT("POINT WITH MAX TVR:  U,V,Z,VARSET,TVR:",/,
     &2F10.1,/,3(F20.5,1X),/)
9707  FORMAT("POINT WITH MAX RISK:  U,V,Z,VARSET,ZRISK:",/,
     &2F10.1,/,3(F20.5,1X),/)
9708  FORMAT("POINT WITH MAX VARIENCE:  U,V,Z,VARSET,TOTV,TOTSD:",/,
     &2F10.1,/,2(F20.5,1X),2(F15.5,1X),/)
9709  FORMAT("THE MEDIAN OF THIS POINT IS:",F20.5,/)
      GO TO 5699
5510  IF(XLOG.EQ.0) GO TO 5520
      DO 5930 I=1,NNN
      ZMED(I,I)=2.718282**(ZKRIGE(I,I))
      ZRISK(I,I)=2.718282**(ZRISK(I,I))
      ZKRIGE(I,I)=2.718282**(ZKRIGE(I,I)+(VARSET(I,I)/2))
5930  VARSET(I,I)=(ZKRIGE(I,I)**2)*((2.718282**VARSET(I,I))-1)
C
C****  IDENTIFYING THE BEST SAMPLING SITES AMONG
C      NON-REGULAR GRID POINTS
C
5520  TVMAX=0
      VARMAX=0
      RISKMAX=0
      DO 5900 I=1,NNN
      IF(IVRD.EQ.0) GO TO 5901
      TVMAX=MAX(TVMAX,TVARED(I,I))
5901  IF(RISK.EQ.0) GO TO 5900
      RISKMAX=MAX(RISKMAX,ZRISK(I,I))
      VARMAX=MAX(VARMAX,VARSET(I,I))
5900  CONTINUE
      DO 5910 I=1,NNN
      IF(IVRD.EQ.0) GO TO 5911
      IF(TVMAX.NE.TVARED(I,I)) GO TO 5911
      WRITE(6,9706)UUU(I),VVV(I),ZKRIGE(I,I),VARSET(I,I)
     &,TVARED(I,I)
      IF(XLOG.EQ.1)WRITE(6,9709)ZMED(I,I)
5911  IF(RISK.EQ.0) GO TO 5912
      IF(RISKMAX.NE.ZRISK(I,I)) GO TO 5912
      WRITE(6,9707)UUU(I),VVV(I),ZKRIGE(I,I),VARSET(I,I)
```

```
     &,TOTV,TOTSD
      IF(XLOG.EQ.1)WRITE(6,9709)ZMED(I,I)
910   CONTINUE
      IF(IVRD.NE.0) GO TO 1

**** PREPARING ZKR,VAR, AND ZRISK FILES FOR DISSPLA GRAPHICS

699   WRITE(6,9805)
805   FORMAT("DO YOU DESIRE TO PRODUCE MAPS?",/,
     &       "IF YES: TYPE 1, OTHERWISE 0",/)
      READ(5,*)IMAP
      IF(IMAP.EQ.0) GO TO 1
*****  NUMBER OF DATA POINTS IS WRITTEN INTO
*****  EACH OUTPUT FILE FOR MAPPING PURPOSES
      IF(NONREG.EQ.0)  SIZE=IUMAX*IVMAX
      IF(NONREG.EQ.1)  SIZE=NNN
      WRITE(4,*)SIZE
      WRITE(7,*)SIZE
      WRITE(8,*)SIZE
      IF(NONREG.NE.1) GO TO 5697
      DO 5696 I=1,NNN
      WRITE(4,*)UUU(I),VVV(I),ZRISK(I,I)
      WRITE(8,*)UUU(I),VVV(I),ZKRIGE(I,I)
      WRITE(7,*)UUU(I),VVV(I),VARSET(I,I)
696   CONTINUE
697   DO 5701 I=1,IUMAX
      DO 5701 J=1,IVMAX
      WRITE(4,*)UU(I),VV(J),ZRISK(I,J)
      WRITE(8,*)UU(I),VV(J),ZKRIGE(I,J)
      WRITE(7,*)UU(I),VV(J),VARSET(I,J)
701   CONTINUE
      GO TO 1

***** FORMAT STATEMENTS**************************************************

001   FORMAT("?     NDATA")
004   FORMAT("| SELECT AN OPTION :A,B,C,D,E OR E-TYPE Q TO STOP")
005   FORMAT(A1)
008   FORMAT("?       DR")
011   FORMAT(3F10.5)
006   FORMAT("*********************************",/,
     &"* YOU ARE NOW USING OPTION ",A1," *",/,
     &"*********************************")
107   FORMAT("?       NO        NGC       ISTEP")
109   FORMAT("?   K(",I2,")       C(",I2,")       A1(",
     &I2,")       A3(",I2,")       A5(",I2,")")
111   FORMAT("| GEN. COV. FUNCT. NO",I2," AVERAGE RANK= ",G15.5)
205   FORMAT("| GIVE FORM OF GEN. COV. FUNCT.")
206   FORMAT(4A1)
207   FORMAT("| K=",I1," C=" ,G13.5," A1=",G13.5," A3=",G13.5,
     &" A5=",G13.5)
208   FORMAT("| THIS IS NOT A PROPER GEN. COV. FUNCT.")
209   FORMAT("| TYPE 1 TO ITERATE ON THIS GEN. COV. FUNCT.",/,
     &"| TYPE 0 TO MOVE TO ANOTHER OPTION")
301   FORMAT("?      NO      ISTEP")
303   FORMAT("?       K                C                A1",
     &"       A3               A5")
305   FORMAT("| DOMAIN A-",I5," POINTS-R=",G15.5,/,
     &       "| DOMAIN B-",I5," POINTS-R=",G15.5,
     &      /,"| JACKKNIFEV=",G15.5)
401   FORMAT("?     THETA(IN DEGREES)")
501   FORMAT("?       U0        V0        DU        DV     IUMAX",
     &"     IVMAX      UINCR     VINCR")
503   FORMAT("?      NO        RO")
600   FORMAT("**** DO YOU DESIRE CONVERSION TO LOG VALUES?",/,
     2       "      **** IF NO CONVERSION:          TYPE 0",/,
     3       "      **** IF CONVERSION IS DESIRED:  TYPE 1",/)
      END
**********************************************************************
**** SUBROUTINE CHECK
**********************************************************************
      SUBROUTINE CHECK(C,A0,A1,A2,INDEX)
      INDEX=0
      IF(C.LT.0.)INDEX=1
      IF(A0.GT.0.)INDEX=1
      IF(A2.GT.0.)INDEX=1
      IF(INDEX.EQ.1)RETURN
      A=-10./3.*SQRT(A0*A2)
      IF(A1.LT.A)INDEX=1
      RETURN
      END
**********************************************************************
**** SUBROUTINE DOUBLE
**********************************************************************
      SUBROUTINE DOUBLE(NDATA,DR,NDOUBL)
      COMMON/C3/U(100),V(100)
      COMMON/C6/Z(100)
      NDOUBL=0
      NDAT=NDATA-1
      DO 20 I0=1,NDAT
      JDOUBL=0
      IF(U(I0).EQ.1.E+10.AND.V(I0).EQ.1.E+10)GO TO 20
      I1=I0+1
      DO 10 I=I1,NDATA
      R=(U(I)-U(I0))**2+(V(I)-V(I0))**2
      R=SQRT(R)
      IF(R.GT.DR)GO TO 10
      Z(I0)=Z(I0)+Z(I)
      U(I)=1.E+10
      V(I)=1.E+10
      NDOUBL=NDOUBL+1
      JDOUBL=JDOUBL+1
C     CONTINUE
      Z(I0)=Z(I0)/FLOAT(JDOUBL+1)
D     CONTINUE
      RETURN
      END
**********************************************************************
**** SUBROUTINE ELIMIN
**********************************************************************
      SUBROUTINE ELIMIN(A,X,NEQN,NROW,NCOL)
      DIMENSION A(NROW,NCOL),X(NROW)
      NEQN=NEQN
      NCOLM1=NCOL-1
      IF(NEQN.LE.NROW.AND.NEQN.LE.NCOLM1)GO TO 1
      WRITE(6,61)
1     FORMAT("STOP-DIMENSION ERROR IN ELIMIN")
      STOP
      CONTINUE
      JMAX=NEQN+1
      NEQNM1=NEQN-1
      DO 6 IEQN=1,NEQNM1
      IMIN=IEQN+1
      IMAX=IEQN
      DO 3 I=IMIN,NEQN
      IF(ABS(A(I,IEQN)).GT.ABS(A(IMAX,IEQN))) IMAX=I
      IF(IMAX.EQ.IEQN)GO TO 5
      DO 4 J=IEQN,JMAX
      AA=A(IEQN,J)
      A(IEQN,J)=A(IMAX,J)
      A(IMAX,J)=AA
      CONTINUE
```

```
      X(NEQN)=A(NEQN,JMAX)/A(NEQN,NEQN)
      DO 8 L=2,NEQN
      I=NEQN+1-L
      SUM=A(I,JMAX)
      IP1=I+1
      DO 7 J=IP1,NEQN
    7 SUM=SUM-A(I,J)*X(J)
    8 X(I)=SUM/A(I,I)
      RETURN
      END
C*******************************************************************
C**** SUBROUTINE FIND1
C*******************************************************************
      SUBROUTINE FIND1(NDATA,I0,U0,V0,N0)
      COMMON/C7/ILIST(100)
      COMMON/C8/RLIST(100)
      COMMON/C3/U(100),V(100)
      IF(I0.EQ.0)GO TO 5
      U0=U(I0)
      V0=V(I0)
      IP=I0
      IM=I0
      GO TO 6
    5 CONTINUE
      CALL POS(NDATA,IN,U0)
      IP=IN-1
      IM=IN
    6 CONTINUE
      I=0
   10 CONTINUE
      IP=IP+1
      IF(IP.GT.NDATA)GO TO 15
      I=I+1
      ILIST(I)=IP
      RLIST(I)=(U0-U(IP))**2+(V0-V(IP))**2
      RLIST(I)=SQRT(RLIST(I))
      IF(I.EQ.N0)GO TO 20
   15 CONTINUE
      IM=IM-1
      IF(IM.LE.0)GO TO 10
      I=I+1
      ILIST(I)=IM
      RLIST(I)=(U0-U(IM))**2+(V0-V(IM))**2
      RLIST(I)=SQRT(RLIST(I))
      IF(I.EQ.N0)GO TO 20
      GO TO 10
   20 CONTINUE
      CALL HLPFND(N0)
      IFP=0
      IFM=0
   30 CONTINUE
      IP=IP+1
      IF(IP.GT.NDATA)GO TO 40
      IF(ABS(U(IP)-U0).GE.RLIST(N0))GO TO 40
      R=(U0-U(IP))**2+(V0-V(IP))**2
      R=SQRT(R)
      IF(R.LT.RLIST(N0))GO TO 35
      GO TO 50
   35 CONTINUE
      ILIST(N0)=IP
      RLIST(N0)=R
      CALL HLPFND(N0)
      GO TO 50
   40 CONTINUE
      IFP=1
      IF(IFP.EQ.1.AND.IFM.EQ.1)RETURN
   50 CONTINUE
      IM=IM-1
      IF(IM.LE.0)GO TO 140
      IF(ABS(U(IM)-U0).GE.RLIST(N0))GO TO 140
      R=(U0-U(IM))**2+(V0-V(IM))**2
      R=SQRT(R)
      IF(R.LT.RLIST(N0))GO TO 135
      GO TO 30
  135 CONTINUE
      ILIST(N0)=IM
      RLIST(N0)=R
      CALL HLPFND(N0)
      GO TO 30
  140 CONTINUE
      IFM=1
      IF(IFP.EQ.1.AND.IFM.EQ.1)RETURN
      GO TO 30
      END
C*******************************************************************
C**** SUBROUTINE FIND2
C*******************************************************************
      SUBROUTINE FIND2(NDATA,I0,U0,V0,R0,N0)
      COMMON/C7/ILIST(100)
      COMMON/C8/RLIST(100)
      COMMON/C3/U(100),V(100)
      IF(I0.EQ.0)GO TO 5
      U0=U(I0)
      V0=V(I0)
      IP=I0
      IM=I0
      GO TO 6
    5 CONTINUE
      CALL POS(NDATA,IN,U0)
      IP=IN-1
      IM=IN
    6 CONTINUE
      IFP=0
      IFM=0
      I=0
   30 CONTINUE
      IP=IP+1
      IF(IP.GT.NDATA)GO TO 40
      IF(ABS(U0-U(IP)).GE.R0)GO TO 140
      R=(U0-U(IP))**2+(V0-V(IP))**2
      R=SQRT(R)
      IF(R.LT.R0)GO TO 35
      GO TO 50
   35 CONTINUE
      I=I+1
      ILIST(I)=IP
      RLIST(I)=R
      GO TO 50
   40 CONTINUE
      IFP=1
      IF(IFP.EQ.1.AND.IFM.EQ.1)GO TO 200
   50 CONTINUE
      IM=IM-1
      IF(IM.LE.0)GO TO 140
      IF (ABS(U0-U(IM)).GE.R0)GO TO 140
      R=(U0-U(IM))**2+(V0-V(IM))**2
      R=SQRT(R)
      IF(R.LE.R0)GO TO 135
      GO TO 30
  135 CONTINUE
      I=I+1
      ILIST(I)=IM
```

```
      IF(IFP.EQ.1.AND.IFM.EQ.1)GO TO 200
      GO TO 30
 20   CONTINUE
      NO=I
      CALL HLPFND(NO)
      RETURN
      END
C*********************************************************************
C**** SUBROUTINE FIT
C*********************************************************************
      SUBROUTINE FIT(NDATA,IMIN,IMAX,ISTEP,NO,K,C,A1,A3,A5,
     &SUM1,SUM2,N)
      COMMON/C7/ILIST(100)
      COMMON/C8/RLIST(100)
      COMMON/C3/U(100),V(100)
      COMMON/C6/Z(100)
      COMMON/C5/A(106,107)
      COMMON/C9/P(100)
      SUM1=0.
      SUM2=0.
      N=0
      DO 100 I0=IMIN,IMAX,ISTEP
      N=N+1
      CALL FIND1(NDATA,I0,U(I0),V(I0),NO)
      NEQN=NO+(K+1)*(K+2)/2
      CALL KRIGPO(K,C,A1,A3,A5,U(I0),V(I0),NO)
      CALL ELIMIN(A,P,NEQN,106,107)
      SUM=-Z(I0)
      DO 50 L=1,NO
      I=ILIST(L)
      SUM=SUM+P(L)*Z(I)
 50   CONTINUE
      SUM1=SUM1+SUM**2
      SUM2=SUM2+C
      DO 60 L=1,NO
      I=ILIST(L)
      H=(U(I0)-U(I))**2+(V(I0)-V(I))**2
      H=SQRT(H)
      SUM2=SUM2-P(L)*GENCOV(K,C,A1,A3,A5,H)
 60   CONTINUE
      SUM2=SUM2-P(NO+1)
      IF(K.EQ.0)GO TO 100
      SUM2=SUM2-P(NO+2)*U(I0)-P(NO+3)*V(I0)
      IF(K.EQ.1)GO TO 100
      SUM2=SUM2-P(NO+4)*U(I0)*V(I0)-P(NO+5)*U(I0)**2-P(NO+6)*V(I0)**2
100   CONTINUE
      RETURN
      END
C*********************************************************************
C**** FUNCTION GENCOV
C*********************************************************************
      FUNCTION GENCOV(K,C,A1,A3,A5,H)
      IF(H.NE.0.)GO TO 10
      GENCOV=C
      RETURN
 10   CONTINUE
      GENCOV=A1*H
      IF(K.EQ.0)RETURN
      GENCOV=GENCOV+A3*H**3
      IF(K.EQ.1)RETURN
      GENCOV=GENCOV+A5*H**5
      RETURN
      END
C*********************************************************************
C**** SUBROUTINE HLPFND
C*********************************************************************
      SUBROUTINE HLPFND(NO)
      COMMON/C8/RLIST(100)
      COMMON/C7/ILIST(100)
      IMAX=NO-1
 10   CONTINUE
      IFLAG=0
      DO 40 I=1,IMAX
      IF(RLIST(I).GT.RLIST(I+1))GO TO 20
      GO TO 40
 20   CONTINUE
      ITEMP=ILIST(I)
      RTEMP=RLIST(I)
      ILIST(I)=ILIST(I+1)
      RLIST(I)=RLIST(I+1)
      ILIST(I+1)=ITEMP
      RLIST(I+1)=RTEMP
      IFLAG=1
 40   CONTINUE
      IF(IFLAG.EQ.0)RETURN
      IMAX=IMAX-1
      GO TO 10
      END
C*********************************************************************
C**** SUBROUTINE KRIGBL
C*********************************************************************
      SUBROUTINE KRIGBL(K,C,A1,A3,A5,U1,U2,V1,V2,NO)
      COMMON/C5/A(106,107)
      COMMON/C7/ILIST(100)
      COMMON/C3/U(100),V(100)
      NROW=NO+(K+1)*(K+2)/2
      NCOL=NROW+1
      DO 5 IROW=1,NROW
      DO 5 ICOL=1,NCOL
      A(IROW,ICOL)=0.
 5    CONTINUE
      DO 40 J1=1,NO
      DO 40 J2=1,NO
      I1=ILIST(J1)
      I2=ILIST(J2)
      H=(U(I1)-U(I2))**2+(V(I1)-V(I2))**2
      H=SQRT(H)
      A(J1,J2)=GENCOV(K,C,A1,A3,A5,H)
 40   CONTINUE
      DO 50 ICOL=1,NO
      A(NO+1,ICOL)=1.
      A(ICOL,NO+1)=A(NO+1,ICOL)
 50   CONTINUE
      IF(K.EQ.0)GO TO 80
      DO 60 ICOL=1,NO
      I=ILIST(ICOL)
      A(NO+2,ICOL)=U(I)
      A(ICOL,NO+2)=A(NO+2,ICOL)
      A(NO+3,ICOL)=V(I)
      A(ICOL,NO+3)=A(NO+3,ICOL)
 60   CONTINUE
      IF(K.EQ.1)GO TO 80
      DO 70 ICOL=1,NO
      I=ILIST(ICOL)
      A(NO+4,ICOL)=U(I)*V(I)
      A(ICOL,NO+4)=A(NO+4,ICOL)
      A(NO+5,ICOL)=U(I)**2
      A(ICOL,NO+5)=A(NO+5,ICOL)
      A(NO+6,ICOL)=V(I)**2
      A(ICOL,NO+6)=A(NO+6,ICOL)
 70   CONTINUE
      CONTINUE
```

```
      V20=V2-V(I)
      CALL QDR(K,C,A1,A3,A5,U10,U20,V10,V20,F)
      A(IROW,NCOL)=F
90    CONTINUE
      A(N0+1,NCOL)=(U2-U1)*(V2-V1)
      IF (K.EQ.0)RETURN
      A(N0+2,NCOL)=(U2**2-U1**2)*(V2-V1)/2.
      A(N0+3,NCOL)=(V2**2-V1**2)*(U2-U1)/2.
      IF(K.EQ.1)RETURN
      A(N0+4,NCOL)=(U2**2-U1**2)*(V2**2-V1**2)/4.
      A(N0+5,NCOL)=(U2**3-U1**3)*(V2-V1)/3.
      A(N0+6,NCOL)=(V2**3-V1**3)*(U2-U1)/3.
      RETURN
      END
C*********************************************************************
C***** SUBROUTINE KRIGPO
C*********************************************************************
      SUBROUTINE KRIGPO(K,C,A1,A3,A5,U0,V0,N0)
      COMMON/C7/ILIST(100)
      COMMON/C5/A(106,107)
      COMMON/C3/U(100),V(100)
      NROW=N0+(K+1)*(K+2)/2
      NCOL=NROW+1
      DO 5 IROW=1,NROW
      DO 5 ICOL=1,NCOL
      A(IROW,ICOL)=0.
5     CONTINUE
      DO 40 J1=1,N0
      DO 40 J2=1,N0
      I1=ILIST(J1)
      I2=ILIST(J2)
      H=(U(I1)-U(I2))**2+(V(I1)-V(I2))**2
      H=SQRT(H)
      A(J1,J2)=GENCOV(K,C,A1,A3,A5,H)
40    CONTINUE
      DO 50 ICOL=1,N0
      A(N0+1,ICOL)=1.
      A(ICOL,N0+1)=A(N0+1,ICOL)
50    CONTINUE
      IF(K.EQ.0)GO TO 80
      DO 60 ICOL=1,N0
      I=ILIST(ICOL)
      A(N0+2,ICOL)=U(I)
      A(ICOL,N0+2)=A(N0+2,ICOL)
      A(N0+3,ICOL)=V(I)
      A(ICOL,N0+3)=A(N0+3,ICOL)
60    CONTINUE
      IF(K.EQ.1)GO TO 80
      DO 70 ICOL=1,N0
      I=ILIST(ICOL)
      A(N0+4,ICOL)=U(I)*V(I)
      A(ICOL,N0+4)=A(N0+4,ICOL)
      A(N0+5,ICOL)=U(I)**2
      A(ICOL,N0+5)=A(N0+5,ICOL)
      A(N0+6,ICOL)=V(I)**2
      A(ICOL,N0+6)=A(N0+6,ICOL)
70    CONTINUE
80    CONTINUE
      DO 90 IROW=1,N0
      I=ILIST(IROW)
      H=(U0-U(I))**2+(V0-V(I))**2
      H=SQRT(H)
      A(IROW,NCOL)=GENCOV(K,C,A1,A3,A5,H)
90    CONTINUE
      A(N0+1,NCOL)=1.
      IF(K.EQ.0)RETURN
      A(N0+2,NCOL)=U0
      A(N0+3,NCOL)=V0
      IF(K.EQ.1)RETURN
      A(N0+4,NCOL)=U0*V0
      A(N0+5,NCOL)=U0**2
      A(N0+6,NCOL)=V0**2
      RETURN
      END
C*********************************************************************
C**** SUBROUTINE POS
C*********************************************************************
      SUBROUTINE POS(NDATA,IN,U0)
      COMMON/C3/U(100),V(100)
      DO 4 I=1,NDATA
      IF(U0.LE.U(I))GO TO 5
4     CONTINUE
      IN=NDATA
      RETURN
5     CONTINUE
      IN=I
      RETURN
      END
C*********************************************************************
C**** SUBROUTINE QDR
C*********************************************************************
      SUBROUTINE QDR(K,C,A1,A3,A5,U1,U2,V1,V2,F)
      F=0.
      IF(C.EQ.0.)GO TO 100
      F=F+C*(U2-U1)*(V2-V1)
100   CONTINUE
      IF(A1.EQ.0.)GO TO 200
      CALL QDR1(U1,V1,F1)
      CALL QDR1(U2,V2,F2)
      CALL QDR1(U1,V2,F3)
      CALL QDR1(U2,V1,F4)
      F=F+A1*(F1+F2-F3-F4)
200   CONTINUE
      IF(K.EQ.0)RETURN
      IF(A3.EQ.0.)GO TO 300
      CALL QDR3(U1,V1,F1)
      CALL QDR3(U2,V2,F2)
      CALL QDR3(U1,V2,F3)
      CALL QDR3(U2,V1,F4)
      F=F+A3*(F1+F2-F3-F4)
300   CONTINUE
      IF(K.EQ.1)RETURN
      IF(A5.EQ.0.)RETURN
      CALL QDR5(U1,V1,F1)
      CALL QDR5(U2,V2,F2)
      CALL QDR5(U1,V2,F3)
      CALL QDR5(U2,V1,F4)
      F=F+A5*(F1+F2-F3-F4)
      RETURN
      END
C*********************************************************************
C**** SUBROUTINE QDR1
C*********************************************************************
      SUBROUTINE QDR1(X,Y,F)
      T=X**2+Y**2
      T=SQRT(T)
      IF(X.EQ.0.)A=0.
      IF(X.NE.0.)A=X**3*ALOG(Y+T)/6.
      IF(Y.EQ.0.)B=0.
      IF(Y.NE.0.)B=Y**3*ALOG(X+T)/6.
      F=Y*T/3.+A+B
```

```
**********
      SUBROUTINE QDR3(X,Y,F)
      T=X**2+Y**2
      T=SQRT(T)
      IF(Y.EQ.0.)A=0.
      IF(X.NE.0.)A=.075*X**5*ALOG(Y+T)
      IF(Y.EQ.0.)B=0.
      IF(Y.NE.0.)B=.075*Y**5*ALOG(X+T)
      F=.175*X*Y*T**3+A+B
      RETURN
      END
**********
**** SUBROUTINE QDR5
**********
      SUBROUTINE QDR5(X,Y,F)
      T=X**2+Y**2
      T=SQRT(T)
      IF(X.EQ.0.)A=0.
      IF(X.NE.0.)A=.44642857E-1*X**7*ALOG(Y+T)
      IF(Y.EQ.0.)B=0.
      IF(Y.NE.0.)B=.44642857E-1*Y**7*ALOG(X+T)
      F=.1220238*X*Y*T**5-.89285714E-1*X**3*Y**3*T+A+B
      RETURN
      END
**********
**** SUBROUTINE ROTATE
**********
      SUBROUTINE ROTATE(THETA,NDATA)
      COMMON/C3/U(100),V(100)
      CS=COS(THETA)
      SN=SIN(THETA)
      DO 10 I=1,NDATA
      UU=U(I)
      U(I)=U(I)*CS+V(I)*SN
      V(I)=-UU*SN+V(I)*CS
0     CONTINUE
      RETURN
      END
**********
**** SUBROUTINE SORT
**********
      SUBROUTINE SORT(NDATA)
      COMMON/C6/Z(100)
      COMMON/C3/U(100),V(100)
      IMAX=NDATA-1
0     CONTINUE
      IFLAG=0
      DO 40 I=1,IMAX
      IF(U(I).GT.U(I+1))GO TO 20
      GO TO 40
0     CONTINUE
      ZZ=Z(I)
      UU=U(I)
      VV=V(I)
      Z(I)=Z(I+1)
      U(I)=U(I+1)
      V(I)=V(I+1)
      Z(I+1)=ZZ
      U(I+1)=UU
      V(I+1)=VV
      IFLAG=1
0     CONTINUE
      IF(IFLAG.EQ.0)RETURN
      IMAX=IMAX-1
      GO TO 10
      END
4.40.23.UCLP, 64, ELP13  .       1.408KLNS.
```

```
C***** UNIT 96 & 95 =I/O TERMINAL
C
      PROGRAM TKRIG(INPUT,OUTPUT,VAR,ZKR,ZRISK,DAT,VARO,ZKRO,ZRISKO
     &,TAPE4=ZRISK,TAPE5=INPUT,TAPE6=OUTPUT,TAPE7=VAR,TAPE8=ZKR
     &,TAPE9=DAT,TAPE10=VARO,TAPE11=ZKRO,TAPE12=ZRISKO)
C*****NEW MODIFICATIONS.....                                          TKRIG
C*****UNIT 95 ==> UNIT 5 TERMINAL INPUT (OR DATA INPUT)
C*****UNIT 99 ==> UNIT 5 TERMINAL INPUT
C*****UNIT 96 ==> UNIT 6   TERMINAL OUTPUT
C*****UNIT 97 ==> UNIT 7   OUTPUT FILE = VAR
C*****UNIT 98 ==> UNIT 8   OUTPUT FILE = ZKR
C*************************************************************
C
      DIMENSION FORM(4),KM(15),CM(15),A1M(15)
     &,A3M(15),A5M(15),ERROR(15),RANK(15),T(4),X(4),ZKRIGE(100,100)
     &,VARSET(100,100),G(4,5),UU(100),VV(100),ZRISK(100,100)
     &,VARSETO(100,100),ZKRIGEO(100,100),ZRISKO(100,100)
      COMMON/C3/U(100),V(100)
      COMMON/C5/A(106,107)
      COMMON/C6/Z(100)
      COMMON/C7/ILIST(100)
      COMMON/C8/RLIST(100)
      COMMON/C9/P(100)
C*****
      DATA STO,ST1/1HO,1H1/
      DATA STA,STB,STC,STD,STE,STG/1HA,1HB,1HC,1HD,1HE,1HG/
C***** READ INFORMATION ABOUT DATA POINTS
  300 WRITE(6,9700)
      READ(5,*)TIME
      IF(TIME.EQ.0) GO TO 150
      IF(TIME.EQ.1) GO TO 150
      GO TO 300
  150 WRITE(6,9001)
      READ(5,*)NDATA
      WRITE(6,9008)
      READ(5,*)DR
      IF(TIME.EQ.1) THEN
      READ(9,*)(U(I),Z(I),I=1,NDATA)
      DO 175 I=1,NDATA
      V(I)=0.
  175 CONTINUE
      ELSE
      READ(9,*)(V(I),U(I),Z(I),I=1,NDATA)
      END IF
  400 WRITE(6,9600)
      READ(5,*)XLOG
      IF (XLOG.EQ.0) GO TO 999
      IF (XLOG.EQ.1) GO TO 499
      GO TO 400
  499 DO 500 I=1,NDATA
      Z(I)=LOG(Z(I))
  500 CONTINUE
C***** CHECK FOR DOUBLE POINTS
  999 CALL DOUBLE(NDATA,DR,NDOUBL)
C***** REARANGE DATA
      CALL SORT(NDATA)
C***** UPDATE NUMBER OF DATA POINTS
      NDATA=NDATA-NDOUBL
    1 CONTINUE
C
C****** OPTION SELECTION
C
      WRITE(6,9004)
      READ(5,9005)ST
      IF(ST.EQ.STA)GO TO 1000
      IF(ST.EQ.STB)GO TO 2000
      IF(ST.EQ.STC)GO TO 3000
      IF(ST.EQ.STD)GO TO 4000
      IF(ST.EQ.STE)GO TO 5000
      IF(ST.EQ.STG)STOP
      GO TO 1
 1000 CONTINUE
C
C***** OPTION A ***************************************************
C
      WRITE(6,9006)STA
      WRITE(6,9107)
      READ(5,*)NO,NGC,ISTEP
      IF(ISTEP.EQ.0)ISTEP=1
C***** READ K,C,A1,A3,AE FOR EACH GENERALIZED COVARIANCE
      DO 1050 IGC=1,NGC
      WRITE(6,9109)IGC,IGC,IGC,IGC,IGC
      READ(5,*)KM(IGC),CM(IGC),A1M(IGC),A3M(IGC),A5M(IGC)
 1050 CONTINUE
      DO 1060 IGC=1,NGC
      RANK(IGC)=0.
 1060 CONTINUE
      NGI=0
C***** START ITERATION ON DATA POINTS- IO IS
C***** THE TYPICAL DATA POINT
      DO 1200 IO=1,NDATA,ISTEP
      NGI=NGI+1
C***** START ITERATION ON GENERALIZED COVARIENCES - USE EACH
C***** ONE IN TURN TO KRIGE POINT IO
      DO 1100 IGC=1,NGC
      CALL FIND1(NDATA,IO,U(IO),V(IO),NO)
C***** KRIGE POINT IO
      CALL  KRIGPO(KM(IGC),CM(IGC),A1M(IGC),A3M(IGC),A5M(IGC),
     &L(IO),V(IO),NO)
      NEQN=NO+(KM(IGC)+1)*(KM(IGC)+2)/2
      CALL ELIMIN(TIME,KM(IGC),A,P,NEQN,106,107)
      SUM=-Z(IO)
      DO 1080 L=1,NO
      I=ILIST(L)
      SUM=SUM+P(L)*Z(I)
 1080 CONTINUE
C***** FIND KRIGING ERROR OF CURRENT GENERALIZED COVARIENCE
      ERROR(IGC)=ABS(SUM)
 1100 CONTINUE
C***** END OF ITERATION ON GENERALIZED COVARIENCES
C***** UPDATE RANKS OF GENERALIZED COVARIENCES
      DO 1190 IGC1=1,NGC
      RANK(IGC1)=RANK(IGC1)+1.
      DO 1190 IGC2=1,NGC
      IF(ERROR(IGC2).LT.ERROR(IGC1))RANK(IGC1)=RANK(IGC1)+1
 1190 CONTINUE
 1200 CONTINUE
C***** END OF ITERATION ON DATA POINTS
      DO 1210 IGC=1,NGC
      RANK(IGC)=RANK(IGC)/FLOAT(NGI)
 1210 CONTINUE
C***** WRITE RESULTS
      DO 1220 IGC=1,NGC
      WRITE(6,9111)IGC,RANK(IGC)
 1220 CONTINUE
      GO TO 1
 2000 CONTINUE
C
C***** OPTION B ***************************************************
C
      WRITE(6,9006)STB
```

```
      READ(5,2067)(FORM(I),I=1,4)
2010  CONTINUE
C***** SET MATRIX G EQUAL TO ZERO - MATRIX G IS THE AUGMENTED
C***** MATRIX OF THE SET OF EQUATIONS APPEARING IN TABLE 3.I OF
C***** CHAPTER 3
      KP2=K+2
      KP3=K+3
      DO 2020 IROW=1,KP2
      DO 2020 ICOL=1,KP3
      G(IROW,ICOL)=0.
2020  CONTINUE
C**** START ITERATION ON DATA POINTS - IO IS THE TYPICAL DATA POINT
C***** FIND THE NO NEAREST DATA POINTS TO DATA POINT IO
      DO 2200 IO=1,NDATA,ISTEP
      CALL FIND1(NDATA,IO,U(IO),V(IO),NO)
C***** KRIGE POINT IO USING INITIAL K,C,A1,A3,A5
      CALL KRIGPO(K,C,A1,A3,A5,U(IO),V(IO),NO)
      NEQN=NO+(K+1)*(K+2)/2
      CALL LIMIN(TIME,K,A,P,NEQN,106,107)
C***** FILL IN THE VECTOR T-T(1),T(2),T(3),T(4) ARE GIVEN BY THE
C***** LAST FOUR EQUATIONS OF TABLE 3.2 OF CHAPTER 3- NOTICE THE
C***** CHANGE OF NOTATION
      T(1)=1.
      DO 2050 L=1,NO
      T(1)=T(1)+P(L)**2
2050  CONTINUE
      KP2=K+2
      DO 2060 N=2,KP2
      NEXP=2*(N-2)+1
      T(N)=0.
      DO 2060 L1=1,NO
      I1=ILIST(L1)
      H=(U(IO)-U(I1))**2+(V(IO)-V(I1))**2
      H=SQRT(H)
      T(N)=T(N)-2.*P(L1)*H**NEXP
      DO 2060 L2=1,NO
      I2=ILIST(L2)
      H=(U(I1)-U(I2))**2+(V(I1)-V(I2))**2
      H=SQRT(H)
      T(N)=T(N)+P(L1)*P(L2)*H**NEXP
2060  CONTINUE
C***** UPDATE MATRIX G
      KP2=K+2
      DO 2090 IROW=1,KP2
      DO 2090 ICOL=1,KP2
      G(IROW,ICOL)=G(IROW,ICOL)+T(IROW)*T(ICOL)
2090  CONTINUE
C*****-CONSTRUCT GENERALIZED INCREMENT
      GINCR=-Z(IO)
      DO 2100 L=1,NO
      I=ILIST(L)
      GINCR=GINCR+P(L)*Z(I)
2100  CONTINUE
      GINCR2=GINCR**2
C**** UPDATE LAST COLOUMN OF MATRIX G.I.E. THE RIGHT HAND SIDE
C***** VECTOR OF THE SET OF EQUATIONS OF TABLE 3.2 OF CHAPTER 3
      KP2=K+2
      DO 2110 IROW=1,KP2
      G(IROW,K+3)=G(IROW,K+3)+T(IROW)*GINCR2
2110  CONTINUE
2200  CONTINUE
C***** END OF ITERATION ON DATA POINTS -AT THIS STAGE
C***** MATRIX G HAS BEEN FILLED IN -PROCEED TO TAKE INTO ACCOUNT THAT
C***** SOME OF THE COEFFICIENTS C,A1,A3,A5 MAY HAVE BEEN SET EQUAL
C***** TO ZERO A PRIORI
      KP2=K+2
      DO 2240 IROW=1,KP2
      IF(FORM(IROW).EQ.ST1)GO TO 2240
C     KP2=K+2
      DO 2220 ICOL=1,KP2
      G(IROW,ICOL)=0.
      G(ICOL,IROW)=0.
2220  CONTINUE
      G(IROW,K+3)=0.
      G(IROW,IROW)=1.
2240  CONTINUE
      NEQN=K+2
C***** SOLVE FOR C, A1,A3,A5
      CALL ELIMINS(TIME,G,X,NEQN,4,5)
      C=0.
      A1=0.
      A3=0.
      A5=0.
      IF(FORM(1).EQ.ST1) C=X(1)
      IF(FORM(2).EQ.ST1)A1=X(2)
      IF(FORM(3).EQ.ST1)A3=X(3)
      IF(FORM(4).EQ.ST1)A5=X(4)
      WRITE(6,9207)K,C,A1,A3,A5
C***** CHECK IF THIS IS A PROPER GENERALIZED COVARIANCE -
C***** WRITE COEFFICIENTS - ASK IF USER WANTS TO START AGAIN USING
C***** THE NEWLY FOUND GEN. COVARIANCE AS INITIAL GEN. COVARIANCE
      CALL CHECK(C,A1,A3,A5,INDEX)
      IF(INDEX.EQ.1)WRITE(6,9208)
2280  CONTINUE
      WRITE(6,9209)
      READ(5,9005)ST
      IF(ST.EQ.ST1)GO TO 2010
      IF(ST.EQ.ST0)GO TO 1
      GO TO 2280
3000  CONTINUE
C
C***** OPTION C ****************************************************
C
      WRITE(6,9006)STC
      WRITE(6,9301)
      READ(5,*)NO,ISTEP
      IF(ISTEP.EQ.0)ISTEP=1
      WRITE(6,9303)
      READ(5,*)K,C,A1,A3,A5
      IMIN=1
      IMAX=NDATA/2
      CALL FIT(NDATA,IMIN,IMAX,ISTEP,NO,K,C,A1,A3,A5,
     *SUM1A,SUM2A,NA)
      IMIN=2+(NA-1)*ISTEP
      IMAX=NDATA
      CALL FIT(NDATA,IMIN,IMAX,ISTEP,NO,K,C,A1,A3,A5,
     *SUM1B,SUM2B,NB)
      R=(SUM1A+SUM1B)/(SUM2A+SUM2B)
      RA=SUM1A/SUM2A
      RB=SUM1B/SUM2B
      R=2*R-(NA*RA+NB*RB)/(NA+NB)
      WRITE(6,9305)NA,RA,NB,RB,R
      GO TO 1
4000  CONTINUE
C
C***** OPTION D ****************************************************
C
      WRITE(6,9006)STD
      WRITE(6,9401)
      READ(5,*)TH,IA
      [illegible]TA,1.7453[illegible]
      [illegible]
```

```
** OPTION 2 ********************************************************
   WRITE(6,9002)ST
   WRITE(6,9303)
   READ(5,*)K,C,A1,A3,A5
   WRITE(6,9501)
   READ(5,*)U0,V0,DU,DV,IUMAX,IVMAX,UINCR,VINCR
   WRITE(6,9503)
   READ(5,*)NO,RO
   UU0=U0
   VV0=V0
   U0=U0-UINCR
   V0=V0-VINCR
** KRIGE POINT OR BLOCK IU,IV -SEE FIGURE 5.2 OF CHAPTER 5
   DO 5600 IU=1,IUMAX
   U0=U0+UINCR
   U1=U0-DU/2.
   U2=U0+DU/2.
   DO 5500 IV=1,IVMAX
   V0=V0+VINCR
   V1=V0-DV/2.
   V2=V0+DV/2.
   IF(RO.EQ.0.)GO TO 5200
   CALL FIND2(NDATA,0,U0,V0,RO,MO)
   IF(MO.LT.NO)GO TO 5200
   GO TO 5300
   CONTINUE
   MO=NO
   CALL FIND1(NDATA,C,U0,V0,MO)
   CONTINUE
   IF(DU.NE.0..AND.DV.NE.0.)CALL KRIGEL(K,C,A1,A3,A5,U1,U2,V1,V2,MO)
   IF(DU.EQ.0..OR.DV.EQ.0.)CALL KRIGPO(K,C,A1,A3,A5,U0,V0,MO)
   NEQN=MO+(K+1)*(K+2)/2
   CALL ELIMIN(TIME,K,A,P,NEQN,106,107)
   ZKRIGE(IU,IV)=0.
   VARSET(IU,IV)=0.
   DO 5350 L=1,MO
   I=ILIST(L)
   ZKRIGE(IU,IV)=ZKRIGE(IU,IV)+P(L)*Z(I)
   CONTINUE
   IF(DU.NE.0..AND.DV.NE.0.)GO TO 5450
   DO 5400 L=1,MO
   I=ILIST(L)
   H=(U0-U(I))**2+(V0-V(I))**2
   H=SQRT(H)
   VARSET(IU,IV)=VARSET(IU,IV)-P(L)*GENCOV(K,C,A1,A3,A5,H)
   CONTINUE
   VARSET(IU,IV)=VARSET(IU,IV)-P(MO+1)+C
   IF(K.EQ.0)GO TO 5450
   VARSET(IU,IV)=VARSET(IU,IV)-P(MO+2)*U0-P(MO+3)*V0
   IF(K.EQ.1)GO TO 5450
   VARSET(IU,IV)=VARSET(IU,IV)-P(MO+4)*U0*V0-P(MO+5)*U0**2-P(MO+6)
  &*V0**2
   CONTINUE
   CONTINUE
   V0=V0-IVMAX*VINCR
   CONTINUE
** THE MATRICES ZKRIGE AND VARSET IN THE CASE OF POINT KRIGING
** AND THE MATRIX ZKRIGE IN THE CASE OF BLOCK KRIGING
** HAVE BEEN FILLED IN AND CAN BE RETRIEVED FROM THIS LOCATION
   DO 5650 IU=1,IUMAX
   DO 5650 IV=1,IVMAX
   UU(IU)=UU0+FLOAT(IU-1)*UINCR
   VV(IV)=VV0+FLOAT(IV-1)*VINCR
   WRITE(6,9801)
   FORMAT("*** DO YOU DESIRE RISK VALUES?",/,
  &      "      IF YES: TYPE IN THE RISK. FOR",/,
  &                ZRISK=Z(X)+RISK*SQRT(V)",/,
  &      "      IF NO:  TYPE A 0.",/)
   READ(5,*)RISK
   IF (XLOG.EQ.0) GO TO 5698
   DO 5655 I=1,IUMAX
   DO 5655 J=1,IVMAX
   ZRISK0(I,J)=ZKRIGE(I,J)+RISK*SQRT(VARSET(I,J))
   ZKRIGE0(I,J)=ZKRIGE(I,J)
   VARSET0(I,J)=VARSET(I,J)
   ZRISK(I,J)=2.718282**(ZKRIGE(I,J)+RISK*SQRT(VARSET(I,J)))
   ZKRIGE(I,J)=2.718282**(ZKRIGE(I,J)+(VARSET(I,J)/2))
   VARSET(I,J)=ZKRIGE(I,J)**2*((2.718282**VARSET(I,J))-1)
*************************************************************
** PREPARING A FILE FOR DISSPLA GRAPHICS
*************************************************************
   GO TO 5699
   DO 5702 I=1,IUMAX
   DO 5702 J=I,IVMAX
   ZRISK(I,J)=ZKRIGE(I,J)+RISK*SQRT(VARSET(I,J))
   CONTINUE
** NUMBER OF DATA POINTS IS WRITTEN TO OUTPUT FILES
** FOR MAPPING PURPOSES
   SIZE=IUMAX*IVMAX
   WRITE(4,*)SIZE
   WRITE(8,*)SIZE
   WRITE(7,*)SIZE
   WRITE(10,*)SIZE
   WRITE(11,*)SIZE
   WRITE(12,*)SIZE
   DO 5701 I=1,IUMAX
   DO 5701 J=1,IVMAX
   WRITE(4,*)UU(I),VV(J),ZRISK(I,J)
   WRITE(8,*)UU(I),VV(J),ZKRIGE(I,J)
   WRITE(7,*)UU(I),VV(J),VARSET(I,J)
   WRITE(10,*)UU(I),VV(J),VARSET0(I,J)
   WRITE(11,*)UU(I),VV(J),ZKRIGE0(I,J)
   WRITE(12,*)UU(I),VV(J),ZRISK0(I,J)
   CONTINUE
   GO TO 1

** FORMAT STATEMENTS************************************************

   FORMAT("?      KDATA")
   FORMAT("| SELECT AN OPTION (A,B,C,D,E...) - TYPE 0 TO STOP")
   FORMAT(A1)
   FORMAT("?          DF")
   FORMAT(3F10.5)
   FORMAT("**********************************",/,
  &"* YOU ARE NOW USING OPTION ",A1," *",/,
  &"**********************************")
   FORMAT("?        NO          NGC          ISTEP")
   FORMAT("?   K(",I2,")              C(",I2,")              A1(",
  &I2,")          A3(",I2,")            A5(",I2,")")
   FORMAT("| GEN. COV. FUNCT. NO",I2," AVERAGE RANK= ",G15.5)
   FORMAT("| GIVE FORM OF GEN. COV. FUNCT.")
   FORMAT(4A1)
   FORMAT("| K=",I1," C=",G13.5," A1=",G13.5," A3=",G13.5,
  &" A5=",G13.5)
   FORMAT("| THIS IS NOT A PROPER GEN. COV. FUNCT.")
   FORMAT("| TYPE 1 TO ITERATE ON THIS GEN. COV. FUNCT.",/,
  &"| TYPE 0 TO MOVE TO ANOTHER OPTION")
   FORMAT("?        C         ISTEP")
   FORMAT("?                                        A1",
  &"                  ")
```

```
     5"        IVMAY        CIMCR        VIMCM")
9503  FORMAT("?             NO           RO")
9600  FORMAT("**** DO YOU DESIRE CONVERSION TO LOG VALUES?",/,
     2        "     **** IF NO CONVERSION:          TYPE 0",/,
     3        "     **** IF CONVERSION IS DESIRED:  TYPE 1",/)
9700  FORMAT("**** ENTER A 0 FOR SPACE KRIGING",/,
     *        "     ENTER A 1 FOR TIME KRIGING",/)
      END
C*****************************************************************
C**** SUBROUTINE CHECK
C*****************************************************************
      SUBROUTINE CHECK(C,A0,A1,A2,INDEX)
      INDEX=0
      IF(C.LT.0.)INDEX=1
      IF(A0.GT.0.)INDEX=1
      IF(A2.GT.0.)INDEX=1
      IF(INDEX.EQ.1)RETURN
      A=-10./3.*SQRT(A0*A2)
      IF(TIME.EQ.1) A=-2./3.*SQRT(30.*A0*A2)
      IF(A1.LT.A)INDEX=1
      RETURN
      END
C*****************************************************************
C**** SUBROUTINE DOUBLE
C*****************************************************************
      SUBROUTINE DOUBLE(NDATA,DR,NDOUBL)
      COMMON/C3/U(100),V(100)
      COMMON/C6/Z(100)
      NDOUBL=0
      NDAT=NDATA-1
      DO 20 I0=1,NDAT
      JDOUBL=0
      IF(U(I0).EQ.1.E+10.AND.V(I0).EQ.1.E+10)GO TO 20
      I1=I0+1
      DO 10 I=I1,NDATA
      R=(U(I)-U(I0))**2+(V(I)-V(I0))**2
      R=SQRT(R)
      IF(R.GT.DR)GO TO 10
      Z(I0)=Z(I0)+Z(I)
      U(I)=1.E+10
      V(I)=1.E+10
      NDOUBL=NDOUBL+1
      JDOUBL=JDOUBL+1
10    CONTINUE
      Z(I0)=Z(I0)/FLOAT(JDOUBL+1)
20    CONTINUE
      RETURN
      END
C*****************************************************************
C**** SUBROUTINE ELIMIN
C*****************************************************************
      SUBROUTINE ELIMIN(TIME,K,A,X,NEQN,NROW,NCOL)
      DIMENSION A(NROW,NCOL),X(NROW)
      MROW=NEQN
      MCOL=NEQN+1
      IF(TIME.NE.0.)NEQN=NEQN-K*(K+1.)/2
      NEGN=NEQN
      NCOLM1=NCOL-1
      IF(NEQN.LE.NROW.AND.NEQN.LE.NCOLM1)GO TO 1
      WRITE(6,61)
61    FORMAT("STOP-DIMENSION ERROR IN ELIMIN")
      STOP
1     CONTINUE
      IF (TIME.NE.1) GO TO 99
      CHEKK=0.
      DO 200 III=1,MROW
      I=III-CHEKK
      CHEKK=0.
      DO 210 J=1,MCOL
      IF (A(I,J).NE.0) GO TO 200
210   CONTINUE
      MROW=MROW-1
      DO 220 KK=I,MROW
      DO 230 L=1,MCOL
      A(KK,L)=A(KK+1,L)
230   CONTINUE
220   CONTINUE
      CHEKK=1.
200   CONTINUE
      CHEKK=0.
      DO 300 III=1,MCOL
      I=III-CHEKK
      CHEKK=0.
      DO 310 J=1,MROW
      IF(A(J,I).NE.0) GO TO 300
310   CONTINUE
      MCOL=MCOL-1
      DO 320 KK=I,MCOL
      DO 330 L=1,MROW
      A(L,KK)=A(L,KK+1)
330   CONTINUE
320   CONTINUE
      CHEKK=1.
300   CONTINUE
99    CONTINUE
      JMAX=NEQN+1
      NEQNM1=NEQN-1
      DO 6 IEQN=1,NEQNM1
      IMIN=IEQN+1
      IMAX=IEQN
      DO 3 I=IMIN,NEQN
3     IF(ABS(A(I,IEQN)).GT.ABS(A(IMAX,IEQN))) IMAX=I
      IF(IMAX.EQ.IEQN)GO TO 5
      DO 4 J=IEQN,JMAX
      AA=A(IEQN,J)
      A(IEQN,J)=A(IMAX,J)
4     A(IMAX,J)=AA
5     CONTINUE
      DO 6 I=IMIN,NEQN
      FACT=A(I,IEQN)/A(IEQN,IEQN)
      DO 6 J=IMIN,JMAX
6     A(I,J)=A(I,J)-FACT*A(IEQN,J)
      X(NEQN)=A(NEQN,JMAX)/A(NEQN,NEQN)
      DO 8 L=2,NEQN
      I=NEQN+1-L
      SUM=A(I,JMAX)
      IP1=I+1
      DO 7 J=IP1,NEQN
7     SUM=SUM-A(I,J)*X(J)
8     X(I)=SUM/A(I,I)
      IF(TIME.EQ.0) GO TO 15
      IF(K.EQ.0) GO TO 15
      IF(K.EQ.1) GO TO 11
      NEQN=NEQN+3
      X(NEQN)=0.
      X(NEQN-1)=X(NEQN-3)
      X(NEQN-2)=0.
      X(NEQN-3)=0.
      GO TO 15
11    [illegible]
      [illegible]
```

```
      SUBROUTINE ELIMING(IM,A,X,NDIM,NEQN,NCOL)
      DIMENSION A(NDIM,NCOL),X(NEQN)
      NEQN=NEQN
      NCOLM1=NCOL-1
      IF(NEQN.LE.NCOLM1)GO TO 1
      WRITE(6,61)
   61 FORMAT("STOP-DIMENSION ERROR IN ELIMING")
      STOP
    1 CONTINUE
      JMAX=NEQN+1
      NEQNM1=NEQN-1
      DO 6 I=1,NEQNM1
      IMIN=I+1
      IMAX=I
      DO 3 II=IMIN,NEQN
      IF(ABS(A(II,I)).GT.ABS(A(IMAX,I))) IMAX=II
    3 CONTINUE
      IF(IMAX.EQ.I)GO TO 5
      DO 4 J=I,JMAX
      AA=A(I,J)
      A(I,J)=A(IMAX,J)
      A(IMAX,J)=AA
    4 CONTINUE
    5 DO 6 II=IMIN,NEQN
      FACT=A(II,I)/A(I,I)
      DO 6 J=IMIN,JMAX
      A(II,J)=A(II,J)-FACT*A(I,J)
    6 X(NEQN)=A(NEQN,JMAX)/A(NEQN,NCOL)
      DO 9 L=1,NEQNM1
      I=NEQN+1-L
      SUM=A(I,JMAX)
      IP1=I+1
      DO 7 J=IP1,NEQN
      SUM=SUM-A(I,J)*X(J)
    7 X(I)=SUM/A(I,I)
      RETURN
      END
*****************************************************************
* SUBROUTINE FIND1
*****************************************************************
      SUBROUTINE FIND1(NDATA,IO,U0,V0,N0)
      COMMON/C7/ILIST(100)
      COMMON/C8/RLIST(100)
      COMMON/C3/U(100),V(100)
      IF(IO.NE.0)GO TO 5
      U0=U(IO)
      V0=V(IO)
      IP=IO
      IM=IO
      GO TO 6
    5 CONTINUE
      CALL POS(NDATA,IO,UO)
      IP=IO-1
      IM=IO
    6 CONTINUE
      I=0
      CONTINUE
      IP=IP+1
      IF(IP.GT.NDATA)GO TO 15
      I=I+1
      ILIST(I)=IP
      RLIST(I)=(U0-U(IP))**2+(V0-V(IP))**2
      RLIST(I)=SQRT(RLIST(I))
      IF(I.EQ.N0)GO TO 20
      CONTINUE
      IM=IM-1
      IF(IM.LE.0)GO TO 10
      I=I+1
      ILIST(I)=IM
      RLIST(I)=(U0-U(IM))**2+(V0-V(IM))**2
      RLIST(I)=SQRT(RLIST(I))
      IF(I.EQ.N0)GO TO 20
      GO TO 10
      CONTINUE
      CALL HLPEND(N0)
      IFP=0
      IFM=0
      CONTINUE
      IP=IP+1
      IF(IP.GT.NDATA)GO TO 40
      IF(ABS(U(IP)-U0).GE.RLIST(N0))GO TO 40
      R=(U0-U(IP))**2+(V0-V(IP))**2
      R=SQRT(R)
      IF(R.LT.RLIST(N0))GO TO 35
      GO TO 50
      CONTINUE
      ILIST(N0)=IP
      RLIST(N0)=R
      CALL HLPFND(N0)
      GO TO 50
      CONTINUE
      IFP=1
      IF(IFP.EQ.1.AND.IFM.EQ.1)RETURN
      CONTINUE
      IM=IM-1
      IF(IM.LE.0)GO TO 140
      IF(ABS(U(IM)-U0).GE.RLIST(N0))GO TO 140
      R=(U0-U(IM))**2+(V0-V(IM))**2
      R=SQRT(R)
      IF(R.LT.RLIST(N0))GO TO 135
      GO TO 30
      CONTINUE
      ILIST(N0)=IM
      RLIST(N0)=R
      CALL HLPEND(N0)
      GO TO 30
      CONTINUE
      IFM=1
      IF(IFP.EQ.1.AND.IFM.EQ.1)RETURN
      GO TO 30
      END
*****************************************************************
* SUBROUTINE FIND2
*****************************************************************
      SUBROUTINE FIND2(NDATA,TU,U0,V0,KE,*U)
      COMMON/C7/ILIST(100)
      COMMON/C8/RLIST(100)
      COMMON/C3/U(100),V(100)
      IF(IO.EQ.0)GO TO 5
      U0=U(IO)
      V0=V(IO)
      IP=IO
      IM=IO
      GO TO 6
      CONTINUE
      CALL POS(NDATA,IL,UO)
      IP=IL-1
      IM=IL
      CONTINUE
      IFP=0
      IFM=0
      IF
```

```
      R=SQRT(R)
      IF(R.LT.R0)GO TO 35
      GO TO 50
35    CONTINUE
      I=I+1
      ILIST(I)=IP
      RLIST(I)=R
      GO TO 50
40    CONTINUE
      IFP=1
      IF(IFP.EQ.1.AND.IFM.EQ.1)GO TO 200
50    CONTINUE
      IM=IM-1
      IF(IM.LE.0)GO TO 140
      IF (ABS(U0-U(IM)).GE.R0)GO TO 140
      R=(U0-U(IM))**2+(V0-V(IM))**2
      R=SQRT(R)
      IF(R.LE.R0)GO TO 135
      GO TO 30
135   CONTINUE
      I=I+1
      ILIST(I)=IM
      RLIST(I)=R
      GO TO 30
140   CONTINUE
      IFM=1
      IF(IFP.EQ.1.AND.IFM.EQ.1)GO TO 200
      GO TO 30
200   CONTINUE
      NO=I
      CALL HLPFND(NO)
      RETURN
      END
C*****************************************************************
C**** SUBROUTINE FIT
C*****************************************************************
      SUBROUTINE FIT(NDATA,IMIN,IMAX,ISTEP,NO,K,C,A1,A3,A5,
     &SUM1,SUM2,N)
      COMMON/C7/ILIST(100)
      COMMON/C8/RLIST(100)
      COMMON/C3/U(100),V(100)
      COMMON/C4/Z(100)
      COMMON/C5/A(106,107)
      COMMON/C9/P(100)
      SUM1=0.
      SUM2=0.
      N=0
      DO 100 IO=IMIN,IMAX,ISTEP
      N=N+1
      CALL FIND1(NDATA,IO,U(IO),V(IO),NO)
      NEQN=NO+(K+1)*(K+2)/2
      CALL KRIGEQ(K,C,A1,A3,A5,U(IO),V(IO),NO)
      CALL [illegible](TIML,K,A,P,NEQN,106,107)
      SUM=-Z(IO)
      DO 50 L=1,NO
      I=ILIST(L)
      SUM=SUM+P(L)*Z(I)
50    CONTINUE
      SUM1=SUM1+SUM**2
      SUM2=SUM2+C
      DO 60 L=1,NO
      I=ILIST(L)
      H=(U(IO)-U(I))**2+(V(IO)-V(I))**2
      H=SQRT(H)
      SUM2=SUM2-P(L)*GENCOV(K,C,A1,A3,A5,H)
60    CONTINUE
      SUM2=SUM2-P(NO+1)
      IF(K.EQ.0)GO TO 100
      SUM2=SUM2-P(NO+2)*U(IO)-P(NO+3)*V(IO)
      IF(K.EQ.1)GO TO 100
      SUM2=SUM2-P(NO+4)*U(IO)*V(IO)-P(NO+5)*U(IO)**2-P(NO+6)*V(IO)**2
100   CONTINUE
      RETURN
      END
C*****************************************************************
C**** FUNCTION GENCOV
C*****************************************************************
      FUNCTION GENCOV(K,C,A1,A3,A5,H)
      IF(H.NE.0.)GO TO 10
      GENCOV=C
      RETURN
10    CONTINUE
      GENCOV=A1*H
      IF(K.EQ.0)RETURN
      GENCOV=GENCOV+A3*H**3
      IF(K.EQ.1)RETURN
      GENCOV=GENCOV+A5*H**5
      RETURN
      END
C*****************************************************************
C**** SUBROUTINE HLPFND
C*****************************************************************
      SUBROUTINE HLPFND(NO)
      COMMON/C8/RLIST(100)
      COMMON/C7/ILIST(100)
      IMAX=NO-1
10    CONTINUE
      IFLAG=0
      DO 40 I=1,IMAX
      IF(RLIST(I).GT.RLIST(I+1))GO TO 20
      GO TO 40
20    CONTINUE
      ITEMP=ILIST(I)
      RTEMP=RLIST(I)
      ILIST(I)=ILIST(I+1)
      RLIST(I)=RLIST(I+1)
      ILIST(I+1)=ITEMP
      RLIST(I+1)=RTEMP
      IFLAG=1
40    CONTINUE
      IF(IFLAG.EQ.0)RETURN
      IMAX=IMAX-1
      GO TO 10
      END
C*****************************************************************
C**** SUBROUTINE KRIGEL
C*****************************************************************
      SUBROUTINE KRIGEL(K,C,A1,A3,A5,U1,U2,V1,V2,NO)
      COMMON/C5/A(106,107)
      COMMON/C7/ILIST(100)
      COMMON/C3/U(100),V(100)
      NROW=NO+(K+1)*(K+2)/2
      NCOL=NROW+1
      DO 5 IROW=1,NROW
      DO 5 ICOL=1,NCOL
      A(IROW,ICOL)=0.
5     CONTINUE
      DO 40 J1=1,NO
      DO 40 J2=1,NO
      I1=ILIST(J1)
      [illegible]
      [illegible]
```

```
      CALL GDR3(U2,V1,F4)
      F=F+A3*(F1+F2-F3-F4)
      CONTINUE
      IF(M.EQ.1)RETURN
      IF(A5.EQ.0.)RETURN
      CALL GDR3(U1,V1,F1)
      CALL GDR3(U2,V2,F2)
      CALL GDR3(U1,V2,F3)
      CALL GDR3(U2,V1,F4)
      F=F+A5*(F1+F2-F3-F4)
      RETURN
      END
*************************************************************
* SUBROUTINE GDR1
*************************************************************
      SUBROUTINE GDR1(X,Y,F)
      T=X**2+Y**2
      T=SQRT(T)
      IF(X.EQ.0.)A=0.
      IF(X.NE.0.)A=X**3*ALOG(Y+T)/6.
      IF(Y.EQ.0.)B=0.
      IF(Y.NE.0.)B=Y**3*ALOG(X+T)/6.
      F=A-Y*T/3.+B
      RETURN
      END
*************************************************************
* SUBROUTINE GDR3
*************************************************************
      SUBROUTINE GDR3(X,Y,F)
      T=Y**2+Y**2
      T=SQRT(T)
      IF(X.EQ.0.)A=0.
      IF(X.NE.0.)A=.075*X**5*ALOG(Y+T)
      IF(Y.EQ.0.)B=0.
      IF(Y.NE.0.)B=.075*Y**5*ALOG(X+T)
      F=.175*)*Y*T**3+A+B
      RETURN
      END
*************************************************************
* SUBROUTINE GDR5
*************************************************************
      SUBROUTINE GDR5(X,Y,F)
      T=X**2+Y**2
      T=SQRT(T)
      IF(X.EQ.0.)A=0.
      IF(Y.NE.0.)A=.44642857E-1*X**7*ALOG(Y+T)
      IF(Y.EQ.0.)B=0.
      IF(Y.NE.0.)B=.44642857E-1*Y**7*ALOG(X+T)
      F=-.1220238*Y*Y*T**5-.69285714E-1*X**3*Y**3*T+A+B
      RETURN
      END
*************************************************************
* SUBROUTINE ROTATE
*************************************************************
      SUBROUTINE ROTATE(THETA,NDATA)
      COMMON/C3/U(100),V(100)
      CS=COS(THETA)
      SN=SIN(THETA)
      DO 10 I=1,NDATA
      UU=U(I)
      U(I)=U(I)*CS+V(I)*SN
      V(I)=-UU*SN+V(I)*CS
   10 CONTINUE
      RETURN
      END
*************************************************************
* SUBROUTINE SORT
*************************************************************
      SUBROUTINE SORT(NDATA)
      COMMON/C6/Z(100)
      COMMON/C3/U(100),V(100)
      IMAX=NDATA-1
   20 CONTINUE
      IFLAG=0
      DO 40 I=1,IMAX
      IF(U(I).GT.U(I+1))GO TO 20
      GO TO 40
   20 CONTINUE
      ZZ=Z(I)
      UU=U(I)
      VV=V(I)
      Z(I)=Z(I+1)
      U(I)=U(I+1)
      V(I)=V(I+1)
      Z(I+1)=ZZ
      U(I+1)=UU
      V(I+1)=VV
      IFLAG=1
   40 CONTINUE
      IF(IFLAG.EQ.0)RETURN
      IMAX=IMAX-1
      GO TO 10
      END
```

```
C*****
C*****   DEVELOPED BY:  DR. S. ROUHANI
C*****                  SCHOOL OF CIVIL ENGINEERING
C*****                  GEORGIA INSTITUTE OF TECHNOLOGY
C*****                  ATLANTA, GEORGIA  30332
C*****                  AUGUST, 1987
C*****
C*****
C
C
      PROGRAM STVARED(INPUT,OUTPUT,VAR,ZKR,DAT,VARO,MED
     &,TAPE5=INPUT,TAPE6=OUTPUT,TAPE7=VAR,TAPE8=ZKR
     &,TAPE9=DAT,TAPE10=VARO,TAPE11=MED)
C
      DIMENSION FORM(4),KM(15),CM(15),T(4),X(4),A1M(15)
     &,A3M(15),A5M(15),ERROR(15),RANK(15),ZKRIGE(20,20,12)
     &,VARSET(20,20,12),G(4,5),UU(20),VV(20),TT(12)
     &,TVARED(20,20,12),TLIST(200),TVMAX(12)
     &,VARSETO(20,20,12),MKRIGE(20,20,12)
      COMMON/C1/KS,KT,CS,CT,A1S,A1T,A3S,A3T,A5S,A5T
      COMMON/C3/UDAT(200),VDAT(200),TDAT(200),IOUT(200),CHOICE
      COMMON/C5/A(106,107)
      COMMON/C6/Z(200)
      COMMON/C7/ILIST(200)
      COMMON/C8/RLIST(200),U(200),V(200)
      COMMON/C9/P(200)
      COMMON/C10/UU0,VV0,TT0
C*****
      DATA ST0,ST1/1H0,1H1/
      DATA STA,STB,STC,STD,STE,STS,STQ/1HA,1HB,1HC,1HD,1HE,1HS,1HQ/
C***** READ INFORMATION ABOUT DATA POINTS
      WRITE(6,9001)
      READ(5,*)NDATA
      ONDATA=NDATA
      READ(9,*)(UDAT(I),VDAT(I),TDAT(I),Z(I),I=1,NDATA)
      WRITE(6,9701)
      READ(5,*)UFAC,VFAC,TFAC
      DO 300 I=1,NDATA
      UDAT(I)=UDAT(I)/UFAC
      VDAT(I)=VDAT(I)/VFAC
  300 TDAT(I)=TDAT(I)/TFAC
      WRITE(6,9008)
      READ(5,*)DRS,DRT
      DRS=DRS/SQRT((UFAC**2+VFAC**2)/2)
      DRT=DRT/TFAC
C****  CHECK FOR DUPLICATE POINTS
      CALL DOUBLE(NDATA,DRS,DRT,NDOUBL)
      NDATA=NDATA-NDOUBL
      WRITE(6,9009)
      READ(5,*)DX,DT
      DX=DX/SQRT((UFAC**2+VFAC**2)/2)
      DT=DT/TFAC
  400 WRITE(6,9600)
      READ(5,*)XLOG
      IF(XLOG.EQ.0) GO TO 999
      IF(XLOG.EQ.1) GO TO 499
      GO TO 400
  499 DO 500 I=1,NDATA
      Z(I)=LOG(Z(I))
  500 CONTINUE
   11 CONTINUE
C***** READ THE NEW DATA POINTS
      ISEQ=1
      WRITE(6,9610)
      READ(5,*)NEW
      IF(NEW.EQ.0)GO TO 10
      DO 2 INEW=1,NEW
      WRITE(6,9611)
      READ(5,*)UNEW,VNEW,ZNEW
      IF(XLOG.EQ.1)ZNEW=LOG(ZNEW)
      WRITE(6,9612)
      READ(5,*)TSTART,TEND
      ISTART=TSATRT/TINCR
      IEND=TEND/TINCR
      DO 15 ITT=ISTART,IEND
      NDATA=NDATA+1
      UDAT(NDATA)=UNEW
      VDAT(NDATA)=VNEW
      TDAT(NDATA)=ITT*TINCR
      Z(NDATA)=ZNEW
   15 CONTINUE
      IF(NDATA.GE.(ONDATA+IUMAX*IVMAX*ITMAX))WRITE(6,9613)
    2 CONTINUE
      WRITE(6,9614)
      READ(5,*)ISEQ
      IF(ISEQ.EQ.1)GO TO 10
      TOTV=0.
      TOTSD=0.
      GO TO 5001
   10 CONTINUE
      TOTV=0.
      TOTSD=0.
  999 WRITE(6,9700)
      READ(5,*)CHOICE
      IF(CHOICE.EQ.2)GO TO 5000
      DO 150 I=1,NDATA
      IF(CHOICE.EQ.0) THEN
      U(I)=TDAT(I)
      V(I)=0.
      ELSE
      U(I)=UDAT(I)
      V(I)=VDAT(I)
      END IF
  150 CONTINUE
C***** REARANGE DATA
      CALL SORT(NDATA)
    1 CONTINUE
C
C****** OPTION SELECTION
C
      WRITE(6,9004)
      READ(5,9005)ST
      IF(ST.EQ.STA)GO TO 1000
      IF(ST.EQ.STB)GO TO 2000
      IF(ST.EQ.STC)GO TO 3000
      IF(ST.EQ.STD)GO TO 4000
      IF(ST.EQ.STE)GO TO 5000
      IF(ST.EQ.STS)GO TO 999
      IF(ST.EQ.STQ)STOP
      GO TO 1
 1000 CONTINUE
C
C***** OPTION A ***************************************************
C
      WRITE(6,9006)STA
 1040 WRITE(6,9107)
      READ(5,*)NO,NGC,ISTEP
      IF(NO.GE.NDATA) THEN
      WRITE(6,9108)NDATA
      GO TO 1040
      ELSE
      END IF
```

```
      WRITE(6,9207)IGC,IGC,IGC,IGC
      READ(5,*)KM(IGC),CM(IGC),A1M(IGC),A3M(IGC),A5M(IGC)
 050  CONTINUE
      DO 1060 IGC=1,NGC
      RANK(IGC)=0.
 060  CONTINUE
      NGI=0
***** START ITERATION ON DATA POINTS- I0 IS
***** THE TYPICAL DATA POINT
      DO 1200 I0=1,NDATA,ISTEP
      NGI=NGI+1
***** START ITERATION ON GENERALIZED COVARIENCES - USE EACH
***** ONE IN TURN TO KRIGE POINT I0
      DO 1100 IGC=1,NGC
      CALL FIND1(NDATA,I0,U(I0),V(I0),N0,DX,DT)
***** KRIGE POINT I0
      CALL  KRIGP0(KM(IGC),CM(IGC),A1M(IGC),A3M(IGC),A5M(IGC),
     &U(I0),V(I0),N0)
      NEQN=N0+(KM(IGC)+1)*(KM(IGC)+2)/2
      CALL ELIMIN(CHOICE,KM(IGC),A,P,NEQN,106,107)
      SUM=-Z(I0)
      DO 1080 L=1,N0
      I=ILIST(L)
      SUM=SUM+P(L)*Z(I)
 080  CONTINUE
***** FIND KRIGING ERROR OF CURRENT GENERALIZED COVARIENCE
      ERROR(IGC)=ABS(SUM)
 100  CONTINUE
***** END OF ITERATION ON GENERALIZED COVARIENCES
***** UPDATE RANKS OF GENERALIZED COVARIENCES
      DO 1190 IGC1=1,NGC
      RANK(IGC1)=RANK(IGC1)+1.
      DO 1190 IGC2=1,NGC
      IF(ERROR(IGC2).LT.ERROR(IGC1))RANK(IGC1)=RANK(IGC1)+1
 190  CONTINUE
 200  CONTINUE
***** END OF ITERATION ON DATA POINTS
      DO 1210 IGC=1,NGC
      RANK(IGC)=RANK(IGC)/FLOAT(NGI)
 210  CONTINUE
***** WRITE RESULTS
      DO 1220 IGC=1,NGC
      WRITE(6,9111)IGC,RANK(IGC)
 220  CONTINUE
      GO TO 1
 000  CONTINUE
***** OPTION B **********************************************************

      WRITE(6,9006)STB
      WRITE(6,9301)
 005  READ(5,*)N0,ISTEP
      IF(N0.GE.NDATA) THEN
      WRITE(6,9108)NDATA
      GO TO 2005
      ELSE
      END IF
      IF(ISTEP.EQ.0)ISTEP=1
      WRITE(6,9303)
      READ(5,*)K,C,A1,A3,A5
      WRITE(6,9205)
      READ(5,9206)(FORM(I),I=1,4)
 010  CONTINUE
***** SET MATRIX G EQUAL TO ZERO - MATRIX G IS THE AUGMENTED
***** MATRIX OF THE SET OF EQUATIONS APPEARING IN TABLE 3.1 OF
***** CHAPTER 3
      KP2=K+2
      KP3=K+3
      DO 2020 IROW=1,KP2
      DO 2020 ICOL=1,KP3
      G(IROW,ICOL)=0.
 020  CONTINUE
**** START ITERATION ON DATA POINTS - I0 IS THE TYPICAL DATA POINT
***** FIND THE N0 NEARSET DATA POINTS TO DATA POINT I0
      DO 2200 I0=1,NDATA,ISTEP
      CALL FIND1(NDATA,I0,U(I0),V(I0),N0,DX,DT)
***** KRIGE POINT I0 USING INITIAL  K,C,A1,A3,A5
      CALL KRIGP0(K,C,A1,A3,A5,U(I0),V(I0),N0)
      NEQN=N0+(K+1)*(K+2)/2
      CALL ELIMIN(CHOICE,K,A,P,NEQN,106,107)
***** FILL IN THE VECTOR T=T(1),T(2),T(3),T(4) ARE GIVEN BY THE
***** LAST FOUR EQUATIONS OF TABLE 3.2 OF CHAPTER 3- NOTICE THE
***** CHAGE OF NOTATION
      T(1)=1.
      DO 2050 L=1,N0
      T(1)=T(1)+P(L)**2
 050  CONTINUE
      KP2=K+2
      DO 2060 N=2,KP2
      NEXP=2*(N-2)+1
      T(N)=0.
      DO 2060 L1=1,N0
      I1=ILIST(L1)
      H=(U(I0)-U(I1))**2+(V(I0)-V(I1))**2
      H=SQRT(H)
      T(N)=T(N)-2.*P(L1)*H**NEXP
      DO 2060 L2=1,N0
      I2=ILIST(L2)
      H=(U(I1)-U(I2))**2+(V(I1)-V(I2))**2
      H=SQRT(H)
      T(N)=T(N)+P(L1)*P(L2)*H**NEXP
 060  CONTINUE
***** UPDATE MATRIX G
      KP2=K+2
      DO 2090 IROW=1,KP2
      DO 2090 ICOL=1,KP2
      G(IROW,ICOL)=G(IROW,ICOL)+T(IROW)*T(ICOL)
 090  CONTINUE
***** CONSTRUCT GENERALIZED INCREMENT
      GINCR=-Z(I0)
      DO 2100 L=1,N0
      I=ILIST(L)
      GINCR=GINCR+P(L)*Z(I)
 00   CONTINUE
      GINCR2=GINCR**2
*** UPDATE LAST COLOUMN OF MATRIX G,I.E. THE RIGHT HAND SIDE
**** VECTOR OF THE SET OF EQUATIONS OF TABLE 3.2 OF CHAPTER 3
      KP2=K+2
      DO 2110 IROW=1,KP2
      G(IROW,K+3)=G(IROW,K+3)+T(IROW)*GINCR2
 10   CONTINUE
 00   CONTINUE
**** END OF ITERATION ON DATA POINTS -AT THIS STAGE
**** MATRIX G HAS BEEN FILLED IN -PROCEED TO TAKE INTO ACCOUNT THAT
**** SOME OF THE COEFFICIENTS C,A1,A3,A5 MAY HAVE BEEN SET EQUAL
**** TO ZERO A PRIORI
      KP2=K+2
      DO 2240 IROW=1,KP2
      IF(FORM(IROW).EQ.ST1)GO TO 2240
      KP2=K+2
      DO 2220 ICOL=1,KP2
      G(IROW,ICOL)=0.
      G(ICOL,IROW)=0.
```

```
      NEQN=K+2
C***** SOLVE FOR C, A1,A3,A5
      CALL ELIMING(G,X,NEQN,4,5,EFLAG)
      C=0.
      A1=0.
      A3=0.
      A5=0.
      IF(FORM(1).EQ.ST1) C=X(1)
      IF(FORM(2).EQ.ST1)A1=X(2)
      IF(FORM(3).EQ.ST1)A3=X(3)
      IF(FORM(4).EQ.ST1)A5=X(4)
      WRITE(6,9207)K,C,A1,A3,A5
C***** CHECK IF THIS IS A PROPER GENERALIZED COVARIANCE -
C***** WRITE COEFFICIENTS - ASK IF USER WANTS TO START AGAIN USING
C***** THE NEWLY FOUND GEN. COVARIANCE AS INITIAL GEN. COVARIANCE
      CALL CHECK(CHOICE,C,A1,A3,A5,INDEX)
      IF(INDEX.EQ.1)WRITE(6,9208)
2280  CONTINUE
      WRITE(6,9209)
      READ(5,9005)ST
      IF(ST.EQ.ST1)GO TO 2010
      IF(ST.EQ.STO)GO TO 1
      GO TO 2280
3000  CONTINUE
C
C***** OPTION C ***************************************************
C
      WRITE(6,9006)STC
3100  WRITE(6,9301)
      READ(5,*)NO,ISTEP
      IF(ISTEP.EQ.0)ISTEP=1
      IF(NO.GE.NDATA)THEN
      WRITE(6,9108)NDATA
      GO TO 3100
      ELSE
      END IF
      WRITE(6,9303)
      READ(5,*)K,C,A1,A3,A5
      IMIN=1
      IMAX=NDATA/2
      IF(NO.GE.NDATA)NO=NDATA-1
      CALL FIT(NDATA,IMIN,IMAX,ISTEP,NO,K,C,A1,A3,A5,
     &SUM1A,SUM2A,NA,DX,DT)
      IMIN=2+(NA-1)*ISTEP
      IMAX=NDATA
      CALL FIT(NDATA,IMIN,IMAX,ISTEP,NO,K,C,A1,A3,A5,
     &SUM1B,SUM2B,NB,DX,DT)
      R=(SUM1A+SUM1B)/(SUM2A+SUM2B)
      RA=SUM1A/SUM2A
      RB=SUM1B/SUM2B
      R=2*R-(NA*RA+NB*RB)/(NA+NB)
      WRITE(6,9305)NA,RA,NB,RB,R
      GO TO 1
4000  CONTINUE
C
C***** OPTION D ***************************************************
C
      WRITE(6,9006)STD
      WRITE(6,9401)
      READ(5,*)THETA
      THETA=THETA*1.74532925E-2
      CALL ROTATE(THETA,NDATA)
      CALL SORT(NDATA)
      GO TO 1
5000  CONTINUE
C
C***** OPTION E ***************************************************
      WRITE(6,9006)STE
      WRITE(6,9350)
      WRITE(6,9351)
      READ(5,*)KS,CS,A1S,A3S,A5S
      KSS=KS
      WRITE(6,9352)
      READ(5,*)U0,V0,IUMAX,IVMAX,UINCR,VINCR
      U0=U0/UFAC
      V0=V0/VFAC
      UINCR=UINCR/UFAC
      VINCR=VINCR/VFAC
      WRITE(6,9355)
      WRITE(6,9351)
      READ(5,*)KT,CT,A1T,A3T,A5T
      KTT=KT
      WRITE(6,9357)
      READ(5,*)T0,ITMAX,TINCR
      T0=T0/TFAC
      TINCR=TINCR/TFAC
5196  WRITE(6,9353)
      READ(5,*)MO
      IF(MO.GT.NDATA)THEN
      WRITE(6,9354)NDATA
      GO TO 5196
      ELSE
      END IF
C
C***** THE OPTION FOR THE VARIANCE REDUCTION ANALYSIS
C
      WRITE(6,9650)
      READ(5,*)IVRD
      NFLAG=0.
      UU0=U0
      VV0=V0
      TT0=T0
5001  U0=UU0
      V0=VV0
      T0=TT0
      U0=U0-UINCR
      V0=V0-VINCR
      T0=T0-TINCR
C*********************************************
C** BEGIN ITERATION FOR ALL DESIRED POINTS
C*********************************************
      DO 5650 IT=1,ITMAX
      T0=T0+TINCR
      DO 5600 IU=1,IUMAX
      U0=U0+UINCR
      DO 5500 IV=1,IVMAX
      V0=V0+VINCR
C****  LOCATE NEAREST MO POINTS FOR KRIGING
      CALL FIND3(NDATA,U0,V0,T0,MO)
      DO 5200 III=1,MO-1
      LLL=ILIST(III)
      LLLP1=ILIST(III+1)
      IF(UDAT(LLL).EQ.UDAT(LLLP1).AND.VDAT(LLL).EQ.VDAT(LLLP1))THEN
      GO TO 5200
      ELSE
      END IF
      KS=KSS
      GO TO 5215
5200  CONTINUE
      KS=1
5215  CONTINUE
      DO 5250 III=1,MO-1
```

```
      GO TO 5265
250   CONTINUE
      KT=0
265   CONTINUE
      CALL KRGPST(U0,V0,T0,M0)
      NEQN=M0+(KS+1)*(KS+2)/2+KT
      CALL ELIMING(A,P,NEQN,106,107,EFLAG)
      IF(EFLAG.EQ.1)THEN
      ZKRIGE(IU,IV,IT)=-1E+15
      VARSET(IU,IV,IT)=-1E+15
      NFLAG=NFLAG+1
      GO TO 5500
      ELSE
      END IF
      ZKRIGE(IU,IV,IT)=0
      VARSET(IU,IV,IT)=0
      DO 5350 L=1,M0
      I=ILIST(L)
      TLIST(L)=I
      ZKRIGE(IU,IV,IT)=ZKRIGE(IU,IV,IT)+P(L)*Z(I)
350   CONTINUE
      DO 5400 L=1,M0
      I=ILIST(L)
      HS=(U0-UDAT(I))**2+(V0-VDAT(I))**2
      HS=SQRT(HS)
      HT=ABS(T0-TDAT(I))
      VARSET(IU,IV,IT)=VARSET(IU,IV,IT)-P(L)
     &*STGC(KS,KT,CS,CT,A1S,A1T,A3S,A3T,A5S,A5T,HS,HT)
400   CONTINUE
      VARSET(IU,IV,IT)=VARSET(IU,IV,IT)-P(M0+1)*CS*CT
      IF(KS.EQ.0)THEN
      MM0=M0+1
      GO TO 5450
      ELSE
      END IF
      VARSET(IU,IV,IT)=VARSET(IU,IV,IT)-P(M0+2)*U0-P(M0+3)*V0
      IF(KS.EQ.1)THEN
      MM0=M0+3
      GO TO 5450
      ELSE
      END IF
      VARSET(IU,IV,IT)=VARSET(IU,IV,IT)-P(M0+4)*U0*V0-P(M0+5)*U0**2
     &-P(M0+6)*V0**2
      MM0=M0+6
450   CONTINUE
      IF(KT.EQ.0)GO TO 5451
      VARSET(IU,IV,IT)=VARSET(IU,IV,IT)-P(MM0+1)*T0
      IF(KT.EQ.1)GO TO 5451
      VARSET(IU,IV,IT)=VARSET(IU,IV,IT)-P(MM0+2)*T0**2
451   CONTINUE
      TOTV=TOTV+VARSET(IU,IV,IT)
      TOTSD=TOTSD+SQRT(VARSET(IU,IV,IT))

***** VARINANCE REDUCTION ANALYSIS

      IF(IVRD.EQ.0)GO TO 5500
      TVARED(IU,IV,IT)=0.
      DO 6000 IUUU=1,IUMAX
      DO 6000 IVVV=1,IVMAX
      DO 6000 ITTT=1,ITMAX
      VARED=0.
      U00=UU0+FLOAT(IUUU-1)*UINCR
      V00=VV0+FLOAT(IVVV-1)*VINCR
      T00=TT0+FLOAT(ITTT-1)*TINCR
      CALL FIND3(NDATA,U00,V00,T00,M0)
      IMAX=ILIST(M0)
      KMAX=-RLIST(IMAX)
      HRS=SQRT((U0-U00)**2+(V0-V00)**2)
      HRT=ABS(T0-T00)
      K00=STGC(KS,KT,CS,CT,A1S,A1T,A3S,A3T,A5S,A5T,HRS,HRT)
      IF(K00.LT.KMAX)GO TO 6000
      DO 6020 L=1,M0
      I=TLIST(L)
      DEP=SQRT((U00-UDAT(I))**2+(V00-VDAT(I))**2)
      TEP=ABS(T00-TDAT(I))
      VARED=VARED+P(L)*STGC(KS,KT,CS,CT,A1S,A1T,A3S,A3T,A5S,A5T,DEP,TEP)
020   CONTINUE
      VARED=VARED+P(M0+1)-K00
      IF(KS.EQ.0)GO TO 6030
      VARED=VARED+P(M0+2)*U00+P(M0+3)*V00
      IF(KS.EQ.1)GO TO 6030
      VARED=VARED+P(M0+4)*U00*V00+P(M0+5)*UU0**2
     &+P(M0+6)*V00**2
030   CONTINUE
      IF(KT.EQ.0)GO TO 6040
      VARED=VARED+P(MM0+1)*T00
      IF(KT.EQ.1)GO TO 6040
      VARED=VARED+P(MM0+2)*T00**2
040   VARED=VARED**2/VARSET(IU,IV,IT)
      TVARED(IU,IV,IT)=TVARED(IU,IV,IT)+VARED
000   CONTINUE
500   CONTINUE
      V0=V0-IVMAX*VINCR
400   CONTINUE
      U0=U0-IUMAX*UINCR
150   CONTINUE
      DO 5660 IU=1,IUMAX
      DO 5660 IV=1,IVMAX
      DO 5660 IT=1,ITMAX
      UU(IU)=(UU0+FLOAT(IU-1)*UINCR)*UFAC
      VV(IV)=(VV0+FLOAT(IV-1)*VINCR)*VFAC
      TT(IT)=(TT0+FLOAT(IT-1)*TINCR)*TFAC
660   CONTINUE
**************************************************************
***   PREPARE A FILE FOR USE WITH DISSPLA GRAPHICS
**************************************************************
      SIZE=IUMAX*IVMAX*ITMAX-NFLAG
      IF(XLOG.EQ.0)THEN
      WRITE(7,*)SIZE
      WRITE(8,*)SIZE
      DO 5700 I=1,IUMAX
      DO 5700 J=1,IVMAX
      DO 5700 K=1,ITMAX
      IF(ZKRIGE(I,J,K).EQ.-1E+15.AND.VARSET(I,J,K).EQ.
     &-1E+15)GO TO 5700
      WRITE(7,9702)UU(I),VV(J),TT(K),VARSET(I,J,K)
      WRITE(8,9702)UU(I),VV(J),TT(K),ZKRIGE(I,J,K)
600   CONTINUE
      ELSE
      WRITE(7,*)SIZE
      WRITE(9,*)SIZE
      WRITE(10,*)SIZE
      DO 5701 I=1,IUMAX
      DO 5701 J=1,IVMAX
      DO 5701 K=1,ITMAX
      IF(ZKRIGE(I,J,K).EQ.-1E+15.AND.VARSET(I,J,K).EQ.
     &-1E+15)GO TO 5701
      ZKRIGE(I,J,K)=2.718282**ZKRIGE(I,J,K)
      VARSET(I,J,K)=VARSET(I,J,K)
      ZKRIGE(I,J,K)=2.718282**(ZKRIGE(I,J,K)+(VARSET(I,J,K)/2))
      VARSET(I,J,K)=ZKRIGE(I,J,K)**2*((2.718282**VARSET(I,J,K))-1)
```

```
      IF(IVRD.EQ.0)GO TO 1
C
C***** IDENTIFYING BEST SAMPLING POINTS
C
      DO 6100 ITT=1,ITMAX
      TVMAX(ITT)=0.
      DO 6150 IUU=1,IUMAX
      DO 6150 IVV=1,IVMAX
      TVMAX(ITT)=MAX(TVMAX(ITT),TVARED(IUU,IVV,ITT))
 6150 CONTINUE
 6100 CONTINUE
      WRITE(6,9800)
      DO 6200 ITT=1,ITMAX
      DO 6250 IUU=1,IUMAX
      DO 6250 IVV=1,IVMAX
      IF(TVMAX(ITT).NE.TVARED(IUU,IVV,ITT))GO TO 6250
      IF(XLOG.EQ.1)THEN
      WRITE(6,9801)TT(ITT),UU(IUU),VV(IVV)
     &,MKRIGE(IUU,IVV,ITT),VARSET(IUU,IVV,ITT),TVARED(IUU,IVV,ITT)
      ELSE
      WRITE(6,9801)TT(ITT),UU(IUU),VV(IVV)
     &,ZKRIGE(IUU,IVV,ITT),VARSET(IUU,IVV,ITT),TVARED(IUU,IVV,ITT)
      END IF
 6250 CONTINUE
 6200 CONTINUE
      GO TO 11
C
C***** FORMAT STATEMENTS*********************************************
C
 9001 FORMAT("?    NDATA")
 9004 FORMAT("***  SELECT AN OPTION:  A, B, C, D, OR E",/,
     &       "      ***   TYPE S TO SWITCH TIME/SPACE",/,
     &       "            FOR OPTIONS A, B, C, OR D",/,
     &       "      ***   TYPE Q TO QUIT")
 9005 FORMAT(A1)
 9008 FORMAT("***  ENTER DR FOR SPACE AND DR FOR TIME (DRS,DRT)  ****")
 9009 FORMAT("***  INPUT DX AND DT VALUES  ***")
 9011 FORMAT(3F10.5)
 9006 FORMAT("*******************************",/,
     &"** YOU ARE NOW USING OPTION ",A1," **",/,
     &"*******************************")
 9107 FORMAT("?      NO       NGC       ISTEP")
 9108 FORMAT(//,"***  NO MUST BE LESS THAN ",I4,//)
 9109 FORMAT("?   K(",I2,")          C(",I2,")          A1(",
     &I2,")          A3(",I2,")          A5(",I2,")")
 9111  FORMAT("| GEN. COV. FUNCT. NO",I2," AVERAGE RANK= ",G15.5)
 9205 FORMAT("| GIVE FORM OF GEN. COV. FUNCT.")
 9206 FORMAT(4A1)
 9207 FORMAT("| K=",I1," C=" ,G13.5," A1=",G13.5," A3=",G13.5,
     &" A5=",G13.5)
 9208 FORMAT("| THIS IS NOT A PROPER GEN. COV. FUNCT.")
 9209 FORMAT("| TYPE 1 TO ITERATE ON THIS GEN. COV. FUNCT.",/,
     &"| TYPE 0 TO MOVE TO ANOTHER OPTION")
 9301 FORMAT("?      NO      ISTEP")
 9303 FORMAT("?       K                C                A1",
     &"               A3               A5")
 9305 FORMAT("| DOMAIN A-",I5," POINTS-R=",G15.5,/,
     &       "| DOMAIN B-",I5," POINTS-R=",G15.5,
     &       /,"| JACKKNIFEV=",G15.5)
 9350 FORMAT(//,"****  ENTER INFORMATION FOR SPACE DATA  ****")
 9351 FORMAT(/," ***  ENTER COVARIANCE FUNCTION",/,
     &       "       K   C    A1    A3   A5")
 9352 FORMAT(/," ***  ENTER UO,VO,UMAX,VMAX,UINCR,VINCR")
 9353 FORMAT(/,">>>>  ENTER NO  <<<<")
 9354 FORMAT(//,">>>>  NO MUST BE LESS THAN",I4,/,)
 9355 FORMAT(//,"****  ENTER INFORMATION FOR TIME DATA  ****")
 9358 FORMAT(/,"****  ENTER A WEIGHT FOR TIME COVARIANCE")
 9357 FORMAT(/," ***  ENTER TO,TMAX,TINCR")
 9401 FORMAT("?    THETA(IN DEGREES)")
 9501 FORMAT("?      UO        VO        DU        DV     IUMAX",
     &"     IVMAX      UINCR      VINCR")
 9503 FORMAT("?      NO        RO")
 9600 FORMAT("**** DO YOU DESIRE CONVERSION TO LOG VALUES?",/,
     2       "      **** IF NO CONVERSION:          TYPE 0",/,
     3       "      **** IF CONVERSION IS DESIRED:  TYPE 1",/)
 9610 FORMAT("? HOW MANY NEW DATA SITES?")
 9611 FORMAT("? NEW U, V, AND Z?")
 9612 FORMAT("? START AND FINISH TIME OF SAMPLING AT THIS SITE?")
 9613 FORMAT(" ALL ESTIMATION GRID POINTS ARE NOW DEFINED AS",/
     2," MEASUREMENT POINTS.  FURTHER SAMPLING IS NOT POSSIBLE.",/
     3," YOU MAY SELECT THE SEQUENTIAL SAMPLING AND THEN EXIT.")
 9614 FORMAT(" IF SEQUENTIAL SAMPLING IS DESIRED",/
     2," TYPE A 1, OTHERWISE A 0.")
 9650 FORMAT(" ***** IF VARIANCE REDUCTION ANALYSIS",/
     2,"     IS DESIRED TYPE A 1, OTHERWISE A 0.")
 9700 FORMAT("**** ENTER A 0 FOR TIME STRUCTURAL ANALYSIS",/,
     &       "      ENTER A 1 FOR SPACE STRUCTURAL ANALYSIS",/,
     &       "      ENTER A 2 FOR OPTION E",/)
 9701 FORMAT("****  ENTER SCALING FACTORS (UFAC, VFAC, TFAC)  ****")
 9702 FORMAT(3E7.1,1E9.2)
 9800 FORMAT("BEST SAMPLING POINTS",//,3X,"TIME",7X,"U",9X,"V",
     &8X,"Z OR MED",8X,"VARSET",9X,"TVARED",/)
 9801 FORMAT(3(F9.2,1X),3(F14.1,1X))
      END
C
C
C*************************************************
C*************************************************
C*****       LIST OF SUBROUTINES           *******
C*************************************************
C*************************************************
C
C
C*********************************************************************
C**** SUBROUTINE CHECK
C*********************************************************************
      SUBROUTINE CHECK(TIME,C,A0,A1,A2,INDEX)
      INDEX=0
      IF(C.LT.0.)INDEX=1
      IF(A0.GT.0.)INDEX=1
      IF(A2.GT.0.)INDEX=1
      IF(INDEX.EQ.1)RETURN
      A=-10./3.*SQRT(A0*A2)
      IF(TIME.EQ.0) A=-2./3.*SQRT(30.*A0*A2)
      IF(A1.LT.A)INDEX=1
      RETURN
      END
C*************************************************************
C***  SUBROUTINE DOUBLE
C*************************************************************
      SUBROUTINE DOUBLE(NDATA,DRS,DRT,NDOUBL)
      COMMON/C3/UDAT(200),VDAT(200),TDAT(200),IOUT(200),CHOICE
      COMMON/C6/Z(200)
      NDOUBL=0
      NDAT=NDATA-1
      DO 20 I0=1,NDAT
      JDOUBL=0
      IF(UDAT(I0).EQ.1.E+10.AND.VDAT(I0).EQ.1.E+10
     &.AND.TDAT(I0).EQ.1E+10)GO TO 20
      I1=I0+1
      DO 10 I=I1,NDATA
      RS=SQRT((UDAT(I)-UDAT(I0))**2+(VDAT(I)-VDAT(I0))**2)
```

```
      VDAT(I)=1.E+10
      TDAT(I)=1E+10
      NDOUBL=NDOUBL+1
      JDOUBL=JDOUBL+1
 10   CONTINUE
      Z(IO)=Z(IO)/FLOAT(JDOUBL+1)
 20   CONTINUE
      RETURN
      END
C*********************************************************************
C**** SUBROUTINE ELIMIN
C*********************************************************************
      SUBROUTINE ELIMIN(TIME,K,A,X,MEQN,NROW,NCOL)
      DIMENSION A(NROW,NCOL),X(NROW)
      MROW=MEQN
      MCOL=MEQN+1
      IF(TIME.NE.1.)MEQN=MEQN-K*(K+1.)/2
      NEQN=MEQN
      NCOLM1=NCOL-1
      IF(NEQN.LE.NROW.AND.NEQN.LE.NCOLM1)GO TO 1
      WRITE(6,61)
 1    FORMAT("STOP-DIMENSION ERROR IN ELIMIN")
      STOP
   1  CONTINUE
      IF (TIME.NE.0) GO TO 99
      CHEKK=0.
      DO 200 III=1,MROW
      I=III-CHEKK
      CHEKK=0.
      DO 210 J=1,MCOL
      IF (A(I,J).NE.0) GO TO 200
210   CONTINUE
      MROW=MROW-1
      DO 220 KK=I,MROW
      DO 230 L=1,MCOL
      A(KK,L)=A(KK+1,L)
230   CONTINUE
220   CONTINUE
      CHEKK=1.
200   CONTINUE
      CHEKK=0.
      DO 300 III=1,MCOL
      I=III-CHEKK
      CHEKK=0.
      DO 310 J=1,MROW
      IF(A(J,I).NE.0) GO TO 300
310   CONTINUE
      MCOL=MCOL-1
      DO 320 KK=I,MCOL
      DO 330 L=1,MROW
      A(L,KK)=A(L,KK+1)
330   CONTINUE
320   CONTINUE
      CHEKK=1.
300   CONTINUE
 99   CONTINUE
      JMAX=NEQN+1
      NEQNM1=NEQN-1
      DO 6 IEQN=1,NEQNM1
      IMIN=IEQN+1
      IMAX=IEQN
      DO 3 I=IMIN,NEQN
      IF(ABS(A(I,IEQN)).GT.ABS(A(IMAX,IEQN))) IMAX=I
      IF(IMAX.EQ.IEQN)GO TO 5
      DO 4 J=IEQN,JMAX
      AA=A(IEQN,J)
      A(IEQN,J)=A(IMAX,J)
      A(IMAX,J)=AA
      CONTINUE
      DO 6 I=IMIN,NEQN
      FACT=A(I,IEQN)/A(IEQN,IEQN)
      DO 6 J=IMIN,JMAX
      A(I,J)=A(I,J)-FACT*A(IEQN,J)
      X(NEQN)=A(NEQN,JMAX)/A(NEQN,NEQN)
      DO 8 L=2,NEQN
      I=NEQN+1-L
      SUM=A(I,JMAX)
      IP1=I+1
      DO 7 J=IP1,NEQN
      SUM=SUM-A(I,J)*X(J)
      X(I)=SUM/A(I,I)
      IF(TIME.EQ.1) GO TO 15
      IF(K.EQ.0) GO TO 15
      IF(K.EQ.1) GO TO 11
      NEQN=NEQN+3
      X(NEQN)=0.
      X(NEQN-1)=X(NEQN-3)
      X(NEQN-2)=0.
      X(NEQN-3)=0.
      GO TO 15
 11   NEQN=NEQN+1
      X(NEQN)=0.
 15   RETURN
      END
C*********************************************************************
C**** SUBROUTINE ELIMING
C*********************************************************************
      SUBROUTINE ELIMING(A,X,MEQN,NROW,NCOL,EFLAG)
      DIMENSION A(NROW,NCOL),X(NROW)
      NEQN=MEQN
      NCOLM1=NCOL-1
      IF(NEQN.LE.NROW.AND.NEQN.LE.NCOLM1)GO TO 1
      WRITE(6,61)
 61   FORMAT("STOP-DIMENSION ERROR IN ELIMING")
      STOP
   1  CONTINUE
      JMAX=NEQN+1
      NEQNM1=NEQN-1
      DO 6 IEQN=1,NEQNM1
      IMIN=IEQN+1
      IMAX=IEQN
      DO 3 I=IMIN,NEQN
   3  IF(ABS(A(I,IEQN)).GT.ABS(A(IMAX,IEQN))) IMAX=I
      IF(IMAX.EQ.IEQN) GO TO 5
      DO 4 J=IEQN,JMAX
      AA=A(IEQN,J)
      A(IEQN,J)=A(IMAX,J)
   4  A(IMAX,J)=AA
   5  CONTINUE
      DO 6 I=IMIN,NEQN
      IF(ABS(A(IEQN,IEQN)).LT.1E-5)GO TO 9
      FACT=A(I,IEQN)/A(IEQN,IEQN)
      DO 6 J=IMIN,JMAX
   6  A(I,J)=A(I,J)-FACT*A(IEQN,J)
      IF(ABS(A(NEQN,NEQN)).LT.1E-5)GO TO 9
      X(NEQN)=A(NEQN,JMAX)/A(NEQN,NEQN)
      DO 8 L=2,NEQN
      I=NEQN+1-L
      SUM=A(I,JMAX)
      IP1=I+1
      DO 7 J=IP1,NEQN
```

```
      IFLAG=1
   10 RETURN
      END
C*****************************************************************
C**** SUBROUTINE FIND1
C*****************************************************************
      SUBROUTINE FIND1(NDATA,IO,U0,V0,N0,DX,DT)
      COMMON/C7/ILIST(200)
      COMMON/C8/RLIST(200),U(200),V(200)
      COMMON/C3/UDAT(200),VDAT(200),TDAT(200),IOUT(200),CHOICE
      COMMON/C10/UUU,VVV,TTT
      IF(IO.EQ.0)THEN
      UUU=UUU
      VVV=VVV
      TTT=TTT
      ELSE
      UUU=UDAT(IO)
      VVV=VDAT(IO)
      TTT=TDAT(IO)
      END IF
      IF(CHOICE.EQ.0) CALL TDATA(NDATA,UUU,VVV,DX)
      IF(CHOICE.EQ.1) CALL SDATA(NDATA,TTT,DT)
      IF(IO.EQ.0)GO TO 5
      U0=U(IO)
      V0=V(IO)
      IP=IO
      IM=IO
      GO TO 6
    5 CONTINUE
      CALL POS(NDATA,IN,U0)
      IP=IN-1
      IM=IN
    6 CONTINUE
      I=0
   10 CONTINUE
      IP=IP+1
      IF(IP.GT.NDATA) GO TO 15
      IF(IOUT(IP).EQ.0) GO TO 19
      I=I+1
      ILIST(I)=IP
      RLIST(I)=(U0-U(IP))**2+(V0-V(IP))**2
      RLIST(I)=SQRT(RLIST(I))
      IF(I.EQ.N0)GO TO 20
   15 CONTINUE
      IM=IM-1
      IF(IM.LE.0) GO TO 10
      IF(IOUT(IM).EQ.0.) GO TO 19
      I=I+1
      ILIST(I)=IM
      RLIST(I)=(U0-U(IM))**2+(V0-V(IM))**2
      RLIST(I)=SQRT(RLIST(I))
      IF(I.EQ.N0)GO TO 20
   19 GO TO 10
   20 CONTINUE
      CALL HLPFND(N0)
      IFP=0
      IFM=0
   30 CONTINUE
      IP=IP+1
      IF(IP.GT.NDATA) GO TO 40
      IF(IOUT(IP).EQ.0.) GO TO 40
      IF(ABS(U(IP)-U0).GE.RLIST(N0))GO TO 40
      R=(U0-U(IP))**2+(V0-V(IP))**2
      R=SQRT(R)
      IF(R.LT.RLIST(N0))GO TO 35
      GO TO 50
   35 CONTINUE
      ILIST(N0)=IP
      RLIST(N0)=R
      CALL HLPFND(N0)
      GO TO 50
   40 CONTINUE
      IFP=1
      IF(IFP.EQ.1.AND.IFM.EQ.1)RETURN
   50 CONTINUE
      IM=IM-1
      IF(IM.LE.0) GO TO 140
      IF(IOUT(IM).EQ.0.) GO TO 140
      IF(ABS(U(IM)-U0).GE.RLIST(N0))GO TO 140
      R=(U0-U(IM))**2+(V0-V(IM))**2
      R=SQRT(R)
      IF(R.LT.RLIST(N0))GO TO 135
      GO TO 30
  135 CONTINUE
      ILIST(N0)=IM
      RLIST(N0)=R
      CALL HLPFND(N0)
      GO TO 30
  140 CONTINUE
      IFM=1
      IF(IFP.EQ.1.AND.IFM.EQ.1)RETURN
      GO TO 30
      END
C*****************************************************************
C**** SUBROUTINE FIND2
C*****************************************************************
      SUBROUTINE FIND2(NDATA,IO,U0,V0,R0,N0,DX,DT)
      COMMON/C7/ILIST(200)
      COMMON/C8/RLIST(200),U(200),V(200)
      COMMON/C3/UDAT(200),VDAT(200),TDAT(200),IOUT(200),CHOICE
      COMMON/C10/UUU,VVV,TTT
      IF(IO.EQ.0)THEN
      UUU=UUU
      VVV=VVV
      TTT=TTT
      ELSE
      UUU=UDAT(IO)
      VVV=VDAT(IO)
      TTT=TDAT(IO)
      END IF
      IF(CHOICE.EQ.0)CALL TDATA(NDATA,UUU,VVV,DX)
      IF(CHOICE.EQ.1)CALL SDATA(NDATA,TTT,DT)
      IF(IO.EQ.0)GO TO 5
      U0=U(IO)
      V0=V(IO)
      IP=IO
      IM=IO
      GO TO 6
    5 CONTINUE
      CALL POS(NDATA,IN,U0)
      IP=IN-1
      IM=IN
    6 CONTINUE
      IFP=0
      IFM=0
      I=0
   30 CONTINUE
      IP=IP+1
      IF(IP.GT.NDATA)GO TO 40
      IF(IOUT(IP).EQ.0.) GO TO 19
      IF(ABS(U0-U(IP)).GT.R0)GO TO 140
      R=(U0-U(IP))**2+(V0-V(IP))**2
```

```
      I=I+1
      ILIST(I)=IP
      RLIST(I)=R
      GO TO 50
0     CONTINUE
      IFP=1
      IF(IFP.EQ.1.AND.IFM.EQ.1)GO TO 200
0     CONTINUE
      IM=IM-1
      IF(IM.LE.0)GO TO 140
      IF(IOUT(IM).EQ.0) GO TO 19
      IF (ABS(U0-U(IM)).GT.R0)GO TO 140
      R=(U0-U(IM))**2+(V0-V(IM))**2
      R=SQRT(R)
      IF(R.LE.R0)GO TO 135
e     GO TO 30
35    CONTINUE
      I=I+1
      ILIST(I)=IM
      RLIST(I)=R
      GO TO 30
40    CONTINUE
      IFM=1
      IF(IFP.EQ.1.AND.IFM.EQ.1)GO TO 200
      GO TO 30
00    CONTINUE
      N0=I
      CALL HLPFND(N0)
      RETURN
      END
*****************************************************************
****  SUBROUTINE FIND3
*****************************************************************
      SUBROUTINE FIND3(NDATA,U0,V0,T0,M0)
      COMMON/C1/KS,KT,CS,CT,A1S,A1T,A3S,A3T,A5S,A5T
      COMMON/C8/RLIST(200),U(200),V(200)
      COMMON/C7/ILIST(200)
      COMMON/C3/UDAT(200),VDAT(200),TDAT(200),IOUT(200),CHOICE
      DO 1 I=1,NDATA
      HS=(U0-UDAT(I))**2+(V0-VDAT(I))**2
      HS=SQRT(HS)
      HT=ABS(T0-TDAT(I))
      RLIST(I)=-STGC(KS,KT,CS,CT,A1S,A1T,A3S,A3T,A5S,A5T,HS,HT)
      ILIST(I)=I
  1   CONTINUE
      CALL HLPFND(M0)
      DO 10 J=M0+1,NDATA
      IF(RLIST(J).LT.RLIST(M0))THEN
      RLIST(M0)=RLIST(J)
      ILIST(M0)=ILIST(J)
      CALL HLPFND(M0)
      ELSE
      END IF
 10   CONTINUE
      RETURN
      END
*****************************************************************************
**** SUBROUTINE FIT
*****************************************************************************
      SUBROUTINE FIT(NDATA,IMIN,IMAX,ISTEP,N0,K,C,A1,A3,A5,
     $SUM1,SUM2,N,DX,DT)
      COMMON/C7/ILIST(200)
      COMMON/C3/UDAT(200),VDAT(200),TDAT(200),IOUT(200),CHOICE
      COMMON/C8/RLIST(200),U(200),V(200)
      COMMON/C6/Z(200)
      COMMON/C5/A(106,107)
      COMMON/C9/P(200)
      SUM1=0.
      SUM2=0.
      N=0
      DO 100 I0=IMIN,IMAX,ISTEP
      N=N+1
      CALL FIND1(NDATA,I0,U(I0),V(I0),N0,DX,DT)
      NEQN=N0+(K+1)*(K+2)/2
      CALL KRIGP0(K,C,A1,A3,A5,U(I0),V(I0),N0)
      CALL ELIMIN(CHOICE,K,A,P,NEQN,106,107)
      SUM=-Z(I0)
      DO 50 L=1,N0
      I=ILIST(L)
      SUM=SUM+P(L)*Z(I)
 50   CONTINUE
      SUM1=SUM1+SUM**2
      SUM2=SUM2+C
      DO 60 L=1,N0
      I=ILIST(L)
      H=(U(I0)-U(I))**2+(V(I0)-V(I))**2
      H=SQRT(H)
      SUM2=SUM2-P(L)*GENCOV(K,C,A1,A3,A5,H)
 60   CONTINUE
      SUM2=SUM2-P(N0+1)
      IF(K.EQ.0)GO TO 100
      SUM2=SUM2-P(N0+2)*U(I0)-P(N0+3)*V(I0)
      IF(K.EQ.1)GO TO 100
      SUM2=SUM2-P(N0+4)*U(I0)*V(I0)-P(N0+5)*U(I0)**2-P(N0+6)*V(I0)**2
100   CONTINUE
      RETURN
      END
*****************************************************************************
*** FUNCTION GENCOV
*****************************************************************************
      FUNCTION GENCOV(K,C,A1,A3,A5,H)
      IF(H.NE.0.)GO TO 10
      GENCOV=C
      RETURN
      CONTINUE
      GENCOV=A1*H
      IF(K.EQ.0)RETURN
      GENCOV=GENCOV+A3*H**3
      IF(K.EQ.1)RETURN
      GENCOV=GENCOV+A5*H**5
      RETURN
      END
*************************************************************
***  FUNCTION STGC
*************************************************************
      FUNCTION STGC(KS,KT,CS,CT,A1S,A1T,A3S,A3T,A5S,A5T,HS,HT)
      IF(HS.NE.0.OR.HT.NE.0)GO TO 5
      STGC=CS+CT
      RETURN
  5   CONTINUE
      IF(HS.NE.0)GO TO 10
      STGC=A1T*HT+CS
      IF(KT.EQ.0)RETURN
      STGC=STGC+A3T*HT**3
      IF(KT.EQ.1)RETURN
      STGC=STGC+A5T*HT**5
      RETURN
10    CONTINUE
      IF(HT.NE.0)GO TO 15
      STGC=A1S*HS+CT
      IF(KS.EQ.0)RETURN
      STGC=STGC+A3S*HS**3
```

```
      S3=A3S*HS**3
      S5=A5S*HS**5
      T1=A1T*HT
      T3=A3T*HT**3
      T5=A5T*HT**5
      IF(KS.EQ.0.AND.KT.EQ.0)STGC=S1*T1
      IF(KS.EQ.0.AND.KT.EQ.1)STGC=S1*T1*T3
      IF(KS.EQ.1.AND.KT.EQ.0)STGC=S1*S3*T1
      IF(KS.EQ.1.AND.KT.EQ.1)STGC=S1*S3*T1*T3
      IF(KS.EQ.2.AND.KT.EQ.0)STGC=S1*S3*S5*T1
      IF(KS.EQ.0.AND.KT.EQ.2)STGC=S1*T1*T3*T5
      IF(KS.EQ.2.AND.KT.EQ.1)STGC=S1*S3*T1*T3*T5
      IF(KS.EQ.1.AND.KT.EQ.2)STGC=S1*S3*T1*T3*T5
      IF(KS.EQ.2.AND.KT.EQ.2)STGC=S1*S3*S5*T1*T3*T5
      RETURN
      END
C*********************************************************************
C**** SUBROUTINE HLPEND
C*********************************************************************
      SUBROUTINE HLPEND(N0)
      COMMON/C8/RLIST(200),U(200),V(200)
      COMMON/C7/ILIST(200)
      IMAX=N0-1
10    CONTINUE
      IFLAG=0
      DO 40 I=1,IMAX
      IF(RLIST(I).GT.RLIST(I+1))GO TO 20
      GO TO 40
20    CONTINUE
      ITEMP=ILIST(I)
      RTEMP=RLIST(I)
      ILIST(I)=ILIST(I+1)
      RLIST(I)=RLIST(I+1)
      ILIST(I+1)=ITEMP
      RLIST(I+1)=RTEMP
      IFLAG=1
40    CONTINUE
      IF(IFLAG.EQ.0)RETURN
      IMAX=IMAX-1
      GO TO 10
      END
C*********************************************************************
C***** SUBROUTINE KRIGPO
C*********************************************************************
      SUBROUTINE KRIGPO(K,C,A1,A3,A5,U0,V0,N0)
      COMMON/C7/ILIST(200)
      COMMON/C5/A(106,107)
      COMMON/C8/RLIST(200),U(200),V(200)
      NROW=N0+(K+1)*(K+2)/2
      NCOL=NROW+1
      DO 5 IROW=1,NROW
      DO 5 ICOL=1,NCOL
      A(IROW,ICOL)=0.
5     CONTINUE
      DO 40 J1=1,N0
      DO 40 J2=1,N0
      I1=ILIST(J1)
      I2=ILIST(J2)
      H=(U(I1)-U(I2))**2+(V(I1)-V(I2))**2
      H=SQRT(H)
      A(J1,J2)=GENCOV(K,C,A1,A3,A5,H)
40    CONTINUE
      DO 50 ICOL=1,N0
      A(N0+1,ICOL)=1.
      A(ICOL,N0+1)=A(N0+1,ICOL)
50    CONTINUE
      IF(K.EQ.0)GO TO 80
      DO 60 ICOL=1,N0
      I=ILIST(ICOL)
      A(N0+2,ICOL)=U(I)
      A(ICOL,N0+2)=A(N0+2,ICOL)
      A(N0+3,ICOL)=V(I)
      A(ICOL,N0+3)=A(N0+3,ICOL)
60    CONTINUE
      IF(K.EQ.1)GO TO 80
      DO 70 ICOL=1,N0
      I=ILIST(ICOL)
      A(N0+4,ICOL)=U(I)*V(I)
      A(ICOL,N0+4)=A(N0+4,ICOL)
      A(N0+5,ICOL)=U(I)**2
      A(ICOL,N0+5)=A(N0+5,ICOL)
      A(N0+6,ICOL)=V(I)**2
      A(ICOL,N0+6)=A(N0+6,ICOL)
70    CONTINUE
80    CONTINUE
      DO 90 IROW=1,N0
      I=ILIST(IROW)
      H=(U0-U(I))**2+(V0-V(I))**2
      H=SQRT(H)
      A(IROW,NCOL)=GENCOV(K,C,A1,A3,A5,H)
90    CONTINUE
      A(N0+1,NCOL)=1.
      IF(K.EQ.0)RETURN
      A(N0+2,NCOL)=U0
      A(N0+3,NCOL)=V0
      IF(K.EQ.1)RETURN
      A(N0+4,NCOL)=U0*V0
      A(N0+5,NCOL)=U0**2
      A(N0+6,NCOL)=V0**2
      RETURN
      END
C*********************************************************************
C****  SUBROUTINE KRGPST
C*********************************************************************
      SUBROUTINE KRGPST(U0,V0,T0,N0)
      COMMON/C1/KS,KT,CS,CT,A1S,A1T,A3S,A3T,A5S,A5T
      COMMON/C3/UDAT(200),VDAT(200),TDAT(200),IOUT(200),CHOICE
      COMMON/C7/ILIST(200)
      COMMON/C5/A(106,107)
      COMMON/C8/RLIST(200),U(200),V(200)
      NROW=N0+(KS+1)*(KS+2)/2+KT
      NCOL=NROW+1
      DO 5 IROW=1,NROW
      DO 5 ICOL=1,NCOL
      A(IROW,ICOL)=0.
5     CONTINUE
      DO 40 J1=1,N0
      DO 40 J2=1,N0
      I1=ILIST(J1)
      I2=ILIST(J2)
      HS=(UDAT(I1)-UDAT(I2))**2+(VDAT(I1)-VDAT(I2))**2
      HS=SQRT(HS)
      HT=ABS(TDAT(I1)-TDAT(I2))
      A(J1,J2)=STGC(KS,KT,CS,CT,A1S,A1T,A3S,A3T,A5S,A5T,HS,HT)
40    CONTINUE
      DO 50 ICOL=1,N0
      A(N0+1,ICOL)=1.
      A(ICOL,N0+1)=A(N0+1,ICOL)
50    CONTINUE
      IF(KS.EQ.0)GO TO 80
      DO 60 ICOL=1,N0
```

```
      A(ICOL,NO+3)=A(NO+3,ICOL)
 60   CONTINUE
      IF(KS.EQ.1)GO TO 80
      DO 70 ICOL=1,NO
      I=ILIST(ICOL)
      A(NO+4,ICOL)=UDAT(I)*VDAT(I)
      A(ICOL,NO+4)=A(NO+4,ICOL)
      A(NO+5,ICOL)=UDAT(I)**2
      A(ICOL,NO+5)=A(NO+5,ICOL)
      A(NO+6,ICOL)=VDAT(I)**2
      A(ICOL,NO+6)=A(NO+6,ICOL)
 70   CONTINUE
 80   CONTINUE
      DO 1000 I=1,NO
000   CONTINUE
      IF(KS.EQ.0)XXX=2
      IF(KS.EQ.1)XXX=4
      IF(KS.EQ.2)XXX=7
      IF(KT.EQ.0)GO TO 89
      DO 85 ICOL=1,NO
      I=ILIST(ICOL)
      A(NO+XXX,ICOL)=TDAT(I)
      A(ICOL,NO+XXX)=A(NO+XXX,ICOL)
 85   CONTINUE
      DO 1001 I=1,NO
001   CONTINUE
      IF(KT.EQ.1)GO TO 89
      DO 86 ICOL=1,NO
      I=ILIST(ICOL)
      A(NO+XXX+1,ICOL)=TDAT(I)**2
      A(ICOL,NO+XXX+1)=A(NO+XXX+1,ICOL)
 86   CONTINUE
 89   CONTINUE
      DO 90 IROW=1,NO
      I=ILIST(IROW)
      HS=(U0-UDAT(I))**2+(V0-VDAT(I))**2
      HS=SQRT(HS)
      HT=ABS(T0-TDAT(I))
      A(IROW,NCOL)=STGC(KS,KT,CS,CT,A1S,A1T,A3S,A3T,A5S,A5T,HS,HT)
 90   CONTINUE
      A(NO+1,NCOL)=1.
      IF(KS.EQ.0)GO TO 99
      A(NO+2,NCOL)=U0
      A(NO+3,NCOL)=V0
      IF(KS.EQ.1)GO TO 99
      A(NO+4,NCOL)=U0*V0
      A(NO+5,NCOL)=U0**2
      A(NO+6,NCOL)=V0**2
 99   CONTINUE
      IF(KS.EQ.0)XXX=2
      IF(KS.EQ.1)XXX=4
      IF(KS.EQ.2)XXX=7
      IF(KT.EQ.0)RETURN
      A(NO+XXX,NCOL)=T0
      IF(KT.EQ.1)RETURN
      A(NO+XXX+1,NCOL)=T0**2
      RETURN
      END
*****************************************************************************
**** SUBROUTINE POS
*****************************************************************************
      SUBROUTINE POS(NDATA,IN,U0)
      COMMON/C8/RLIST(200),U(200),V(200)
      DO 4 I=1,NDATA
      IF(U0.LE.U(I))GO TO 5
      CONTINUE
      IN=NDATA
      RETURN
      CONTINUE
      IN=I
      RETURN
      END
*****************************************************************************
**** SUBROUTINE QDR
*****************************************************************************
      SUBROUTINE QDR(K,C,A1,A3,A5,U1,U2,V1,V2,F)
      F=0.
      IF(C.EQ.0.)GO TO 100
      F=F+C*(U2-U1)*(V2-V1)
00    CONTINUE
      IF(A1.EQ.0.)GO TO 200
      CALL QDR1(U1,V1,F1)
      CALL QDR1(U2,V2,F2)
      CALL QDR1(U1,V2,F3)
      CALL QDR1(U2,V1,F4)
      F=F+A1*(F1+F2-F3-F4)
00    CONTINUE
      IF(K.EQ.0)RETURN
      IF(A3.EQ.0.)GO TO 300
      CALL QDR3(U1,V1,F1)
      CALL QDR3(U2,V2,F2)
      CALL QDR3(U1,V2,F3)
      CALL QDR3(U2,V1,F4)
      F=F+A3*(F1+F2-F3-F4)
00    CONTINUE
      IF(K.EQ.1)RETURN
      IF(A5.EQ.0.)RETURN
      CALL QDR5(U1,V1,F1)
      CALL QDR5(U2,V2,F2)
      CALL QDR5(U1,V2,F3)
      CALL QDR5(U2,V1,F4)
      F=F+A5*(F1+F2-F3-F4)
      RETURN
      END
*****************************************************************************
**** SUBROUTINE QDR1
*****************************************************************************
      SUBROUTINE QDR1(X,Y,F)
      T=X**2+Y**2
      T=SQRT(T)
      IF(X.EQ.0.)A=0.
      IF(X.NE.0.)A=X**3*ALOG(Y+T)/6.
      IF(Y.EQ.0.)B=0.
      IF(Y.NE.0.)B=Y**3*ALOG(X+T)/6.
      F=X*Y*T/3.+A+B
      RETURN
      END
*****************************************************************************
**** SUBROUTINE QDR3
*****************************************************************************
      SUBROUTINE QDR3(X,Y,F)
      T=X**2+Y**2
      T=SQRT(T)
      IF(X.EQ.0.)A=0.
      IF(X.NE.0.)A=.075*X**5*ALOG(Y+T)
      IF(Y.EQ.0.)B=0.
      IF(Y.NE.0.)B=.075*Y**5*ALOG(X+T)
      F=.175*X*Y*T**3+A+B
      RETURN
      END
*****************************************************************************
**** SUBROUTINE QDR5
```

```
      IF(X.EQ.0.)A=0.
      IF(X.NE.0.)A=.446428575-1*X**7*ALOG(Y+T)
      IF(Y.EQ.0.)B=0.
      IF(Y.NE.0.)B=.446428575-1*Y**7*ALOG(X+T)
      F=.1220238*Y*Y*T**5-.8928571E-1*X**3*Y**3*T+A+B
      RETURN
      END
C*****************************************************************************
C**** SUBROUTINE ROTATE
C*****************************************************************************
      SUBROUTINE ROTATE(THETA,NDATA)
      COMMON/C8/RLIST(200),U(200),V(200)
      CS=COS(THETA)
      SN=SIN(THETA)
      DO 10 I=1,NDATA
      UU=U(I)
      U(I)=U(I)*CS+V(I)*SN
      V(I)=-UU*SN+V(I)*CS
10    CONTINUE
      RETURN
      END
C*****************************************************************************
C****  SUBROUTINE SDATA
C*****************************************************************************
      SUBROUTINE SDATA(NDATA,T0,DT)
      COMMON/C3/UDAT(200),VDAT(200),TDAT(200),IOUT(200),CHOICE
      DO 10 I=1,NDATA
      IF(ABS(TDAT(I)-T0).GT.DT) THEN
      IOUT(I)=0
      ELSE
      IOUT(I)=1.
      END IF
10    CONTINUE
      RETURN
      END
C*****************************************************************************
C**** SUBROUTINE SORT
C*****************************************************************************
      SUBROUTINE SORT(NDATA)
      COMMON/C6/Z(200)
      COMMON/C8/RLIST(200),U(200),V(200)
      COMMON/C3/UDAT(200),VDAT(200),TDAT(200),IOUT(200),CHOICE
      IMAX=NDATA-1
10    CONTINUE
      IFLAG=0
      DO 40 I=1,IMAX
      IF(U(I).GT.U(I+1))GO TO 20
      GO TO 40
20    CONTINUE
      ZZ=Z(I)
      UU=U(I)
      VV=V(I)
      UUDAT=UDAT(I)
      VVDAT=VDAT(I)
      TTDAT=TDAT(I)
      Z(I)=Z(I+1)
      U(I)=U(I+1)
      V(I)=V(I+1)
      UDAT(I)=UDAT(I+1)
      VDAT(I)=VDAT(I+1)
      TDAT(I)=TDAT(I+1)
      Z(I+1)=ZZ
      U(I+1)=UU
      V(I+1)=VV
      UDAT(I+1)=UUDAT
      VDAT(I+1)=VVDAT
      TDAT(I+1)=TTDAT
      IFLAG=1
40    CONTINUE
      IF(IFLAG.EQ.0)RETURN
      IMAX=IMAX-1
      GO TO 10
      END
C*****************************************************************************
C****  SUBROUTINE TDATA
C*****************************************************************************
      SUBROUTINE TDATA(NDATA,U0,V0,DX)
      COMMON/C3/UDAT(200),VDAT(200),TDAT(200),IOUT(200),CHOICE
      DO 10 I=1,NDATA
      IF(SQRT((UDAT(I)-U0)**2+(VDAT(I)-V0)**2).GT.DX) THEN
      IOUT(I)=0.
      ELSE
      IOUT(I)=1.
      END IF
10    CONTINUE
      RETURN
      END
```