DISSERTATION

STATISTICAL ERROR ANALYSIS OF GROUND WATER SYSTEMS

Submitted by

Robert Bibby

In partial fulfillment of the requirements

for the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

March, 1971

CED70-71RB24

COLORADO STATE UNIVERSITY

MARCH 19 71

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER

OUR SUPERVISION BY ROBERT BIBBY

ENTITLED STATISTICAL ERROR ANALYSIS OF GROUND WATER SYSTEMS

BE ACCEPTED AS FULFILLING THIS PART OF THE REQUIREMENTS FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

Committee on Graduate Work

James P. Waltz

Daniel K. Sunada
Adviser

Head of Department

ABSTRACT OF DISSERTATION

STATISTICAL ERROR ANALYSIS OF GROUND WATER SYSTEMS

A method is developed, which, by considering the input variables to a numerical model of flow in porous media as random variables, enables the accuracy of these input variables to be related to the accuracy of the output. The input variables considered are initial head, permeability, discharge, storage coefficient and saturated thickness and the output variable is head after a period of time. The method involves the use of the Monte Carlo technique to generate a random sample of the final head, the computation of a tolerance limit width and a coefficient of variation on the final head which are used as measures of its accuracy, and a regression analysis to determine a predictive relation between the accuracy of the input variables and the accuracy of the final head. The results indicate that if only one of the input variables contains error then this error is linearly related to the error in final head. If all input variables contain error, then only the error on initial head is significant in predicting the error in final head.

In addition, a method of estimating the parameters of the probability density functions of the input variables from available field data is described and the relation is determined between the accuracy of these estimates and the number of data points used to make the estimate. The significance and application of the results in ground water system management is discussed.

Robert Bibby
Civil Engineering Department
Colorado State University
Fort Collins, Colorado 80521
March, 1971

ACKNOWLEDGMENTS

The author is deeply indebted to his adviser, Dr. D. K. Sunada, Associate Professor, Civil Engineering Department, Colorado State University, for the inspiration and guidance he has always provided. The author wishes to express his sincere appreciation to Dr. D. C. Boes, Associate Professor, Department of Mathematics and Statistics, Colorado State University, for the advice he has generously given throughout this study. The help and encouragement of Dr. J. P. Waltz, Associate Professor, Geology Department, Colorado State University and Dr. E. A. Breitenbach, Faculty Affiliate, Colorado State University and President, Scientific Software Corporation, Denver, Colorado, is gratefully acknowledged.

Thanks are also extended to the United States Department of the Interior as authorized under the Water Resources Act of 1964, Public Law 88-379, Project B-022-Colo, which provided financial support for this study, and to the Colorado State University Experiment Station, which provided funds for Project 110 from which a part of this study was sponsored. The funds provided by Colorado State University for use on the CDC 6400 computer are greatly appreciated.

The author wishes to thank Miss Lyn Koch who transformed his handwritten notes into typescript.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

A numerical model, obtained by approximating the partial differential equation of flow in porous media by finite differences, is commonly used to analyze groundwater systems. Its application requires that values be assigned to the input variables, permeability, storage coefficient, saturated thickness, initial head and discharge, and then the model is used to compute values of head at various times. It is an entirely deterministic model and is frequently used in situations in which nothing is known of the accuracy of the estimates of the input variables or how errors in these estimates are related to the accuracy of the results. Further, the relation between the amount of field data available and the accuracy of the estimates of the input variables is usually not known. This study is aimed at establishing these relations by a combined use of deterministic and stochastic methods.

RESEARCH OBJECTIVES AND THEIR SIGNIFICANCE

The principal objective of this study is to relate the accuracy of the estimates of the input variables of the numerical model to the accuracy of the estimate of the output. Thus, if $\tilde{X}_1, \tilde{X}_2, \ldots, \tilde{X}_p$ are the estimates of p input variables $X_1, X_2, \ldots, X_p$ and have errors $\epsilon_1, \epsilon_2, \ldots, \epsilon_p$, respectively, and $\tilde{Y}$ is the estimate of the output variable, Y, with error ϵ_Y, then a relation of the form,

$$\epsilon_Y = E_1(\epsilon_1, \epsilon_2, \ldots, \epsilon_p) \quad ,$$

was sought, where, E_1 denotes the function relating the errors.

This error relation was determined by considering the estimates of the input variables as random variables. Further research objectives were then to establish a technique for estimating the parameters of the probability density functions of the input variables from available field data and to find the relation between the amount of field data used to make the estimate and the accuracy of the estimate. Thus, if n field observations $X_{11}, X_{12}, \ldots, X_{1n}$ are available to estimate X_1, it was required to find,

$$\tilde{X}_1 = E_2(X_{11}, X_{12}, \ldots, X_{1n}),$$

and

$$\epsilon_1 = E_3(n) \quad ,$$

where, E_2, E_3 denote functions relating the variables.

With the results of this research it is hoped that future numerical model studies of groundwater systems will be made more effective, since, the relative importance of the input variables will be known from the point of view of the influence that they have on the accuracy of the results. Also, future data collection should be conducted more efficiently since the amount of data needed to obtain a certain accuracy in the estimate of an input variable will be known. In addition, data can be collected on only the more important variables influencing flow in ground water basins. It should also be possible to determine how much more data it would be necessary to collect on each input variable to obtain a specified increase in the accuracy of the estimate of the output variable.

LITERATURE REVIEW

The study of ground water systems has only recently been made tractable with the introduction of numerical models. In consequence, little work has been reported on the study of the sensitivity of the response of ground water systems to changes in the system parameters. However, the other tools used in this study, namely Regression Analysis, Non-parametric statistics and Monte Carlo techniques are widely used, although not simultaneously, and in conjunction with a deterministic model. Since these three methods are so well known, the literature on them will not be reviewed. Only that literature which pertains to the analysis of the sensitivity of the hydraulic response of a ground water system will be discussed. Literature on the numerical model will be cited in the section in which the model is described.

McMillan[13] investigated the relative importance of basin-wide heterogeneity of permeability in operational analysis of ground water basins. His analysis consisted of repeated numerical solutions of Laplace's Equation with variable coefficients. In his analysis he allowed the following factors to vary one at a time:

i) basin-wide mean value of permeability,

ii) basin-wide standard deviation of permeability,

iii) mean hydraulic gradient,

iv) grid size,

v) grid length to width ratio,

vi) probability density function of permeability,

vii) size of homogeneous and heterogeneous blocks within the basin, and

viii) the weighting factor used to approximate the average permeability between grids.

His results, obtained for a rectangular groundwater basin in which two opposite boundaries were impermeable and the other two were constant heads, indicated the following empirical relationship,

$$\frac{\sigma_d}{\Delta H} = F_d \frac{\sigma_K}{\overline{K}}$$

where, σ_d = standard deviation of the differences in head for homogeneous and heterogeneous solutuions,

ΔH = average drop in head between grids in the direction of flow,

σ_K = basin-wide sample standard deviation of permeability,

$\overline{K}$ = basin-wide sample mean of permeability,

F_d = empirical factor with a value in the range 0.05 to 2.0.

Of the factors considered, (iv)-(viii) had little effect on the above relation. The relation has limited application because only the steady-state flow equation is being solved for a simple groundwater model in which only one aquifer parameter, permeability, is considered random. Also the relation is deduced from a small amount of data, and so further restricts its applicability.

Bittinger[3] investigated the influence that the total input, the aquifer parameters and water management practices have on the return flow in stream aquifer systems. His technique consisted of varying the influencing factors one at a time and analyzing the effects on return flow. He concluded that the return flow response is principally dependent upon the total volume of water added to the aquifer, the width of the aquifer, the location of the application area and the aquifer constant, $(KbT)/(SW^2)$ where, K is permeability, b is saturated thickness, T is time, S is specific yield and W is aquifer width. Water management practices could also significantly effect return flow. Areal variations in permeability and bedrock configuration were found to have an insignificant influence on return flows.

Woods[17] developed a water quality model of a general hydrologic system. The system included a ground water aquifer and he investigated, by changing one variable at a time, the sensitivity of the water quality and system hydrology to changes in physical parameters of the system and in management practices. So far as the ground water aspects of this analysis were concerned, he concluded that the most sensitive term to changes in specific yield was the maximum representative head difference driving flow into drains but this sensitivity decreased with time. Changes in the initial values of this representative head difference had little influence on other system parameters. Changes in the amount of applied irrigation water induced significant effects on all system parameters.

The methods used by Bittinger and Woods to investigate ground water system response involved assigning a small number of different values to the input variables and observing the effect on system response. Such an approach can at best give only a superficial indication of the sensitivity of the system response to changes in the input variables.

Eshett [6] assumed the change in water table elevation in a sub-area of an aquifer to be a linear function of four surface variables, precipitation, pumping, delivered water and artificial spreading. In each sub-area he found maximum likelihood estimates of the coefficients of this linear function assuming the observations of the variables to be normally distributed. Dividing each sub-area into grids, he was able to estimate the net discharge from each grid. He then used these values of net discharge as input to a numerical model of the entire aquifer and solved for water table elevation after a period of time. The regression analysis involved only the surface factors which influence a ground water system and did not consider any of the aquifer parameters as random variables. Also, the interdependence of these surface factors and the fact that large time steps were involved in the study, made it impossible to determine their relative importance in predicting net discharge.

Longenbaugh [10] used stepwise multiple regression to develop a prediction equation for river accretion from applied irrigation water, precipitation, pumping and evapotranspiration. He concluded that the best equation, from a practical point of view, for the aquifer he studied, was,

$$Y = -3595 + 2.3287X_1 + 1064X_2$$

where

Y = river accretion

X_1 = variable measuring volume of ditch diversion,

X_2 = variable measuring precipitation amount.

Pumping volume and consumptive use were found to be non-significant in this aquifer. Longenbaugh and Bittinger [11] reported studies of this same problem using techniques of multivariate analysis. As with Eshett's study, the regression equation used by Longenbaugh and Longenbaugh and Bittinger involved only the surface factors affecting a ground water system. They also encountered problems due to the correlation of the independent variables in the regression equation. It was this correlation which led them to consider multivariate analysis. The results of this approach were not reported in reference [11].

Rorabaugh[15] determined the correlation between changes in water table elevation, precipitation and temperature by graphical means. He used the relationships which he established to predict maximum change in water table elevation during a year knowing the elevation at the beginning of the year. The predictions were acceptably accurate, but, since only the surface factors influencing the ground water system were considered, the results are limited in applicability.

NUMERICAL MODEL

The equation of transient flow in a porous medium may be derived from the mass continuity equation and Darcy's Law and written, (Jacob [8]),

$$\frac{\partial}{\partial x}(Kb\triangle y\,\frac{\partial h}{\partial x})\,\triangle x + \frac{\partial}{\partial y}\,(K\,b\triangle x\,\frac{\partial h}{\partial y})\triangle y = Q + S\triangle x\triangle y\,\frac{\partial h}{\partial T} \qquad (1)$$

where, x, y = space coordinates (L)

K = permeability (L/T)

b = saturated thickness (L)

$\triangle x, \triangle y$ = grid dimensions (L)

h = head (L)

Q = discharge rate from grid (L^3/T)

S = storage coefficient (for confined flow); specific yield (for unconfined flow)

T = time (T) .

For unconfined flow, $b = (h-z)$, where z is the bedrock elevation (L), so equation (1) is non-linear in h. For confined flow b is independent of h, so equation (1) is linear in h.

Dividing the region of flow into grids and using an implicit central finite difference scheme, equation (1), written for one of these grids, becomes:

$$A^T \, h_{i,j-1}^{T+\Delta T} + B^T \, h_{i,j+1}^{T+\Delta T} + C^T \, h_{i-1,j}^{T+\Delta T} + D^T \, h_{i+1,j}^{T+\Delta T}$$

$$- (A + B + C + D + E)^T \, h_{i,j}^{T+\Delta T} = Q^{T+\Delta T/2} - E \, h^T \qquad (2)$$

where,

$$A^T = \left[\left(\frac{\Delta x}{2Kb\Delta y} \right)_{i,j-1} + \left(\frac{\Delta x}{2Kb\Delta y} \right)_{i,j} \right]^{-1}$$

$$B^T = \left[\left(\frac{\Delta x}{2Kb\Delta y} \right)_{i,j+1} + \left(\frac{\Delta x}{2Kb\Delta y} \right)_{i,j} \right]^{-1}$$

$$C^T = \left[\left(\frac{\Delta y}{2Kb\Delta x} \right)_{i-1,j} + \left(\frac{\Delta y}{2Kb\Delta x} \right)_{i,j} \right]^{-1}$$

$$D^T = \left[\left(\frac{\Delta y}{2Kb\Delta x} \right)_{i+1,j} + \left(\frac{\Delta y}{2Kb\Delta x} \right)_{i,j} \right]^{-1}$$

$$E = \left[\frac{(S\Delta x\Delta y)_{i,j}}{\Delta t} \right] \quad .$$

The i, j notation (see Fig. 1) refers to the grid for which a particular equation is written and the superscripts represent the time level of computation.

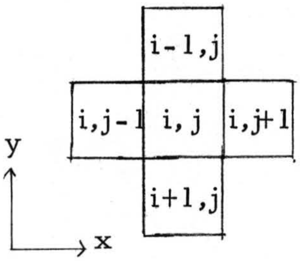


FIGURE 1

Equation (2) is written for every grid in the flow region and the resulting equations are solved simultaneously to give the head in each grid at time (T + ΔT). The development of this model and its applications have been reported by[(1), (2), (4), (5), (12)].

The input variables to the model are h_I, the initial head, K, b(or z), Q and S and the output is h_F, the final head at a given time. The average value of each of the input variables in every grid has to be specified for input. It is therefore required to relate the accuracy of these input variables to the accuracy of h_F, to estimate the parameters of the density functions of the input variables from field data and to relate the accuracy of these estimates to the amount of data used.

THEORETICAL FRAMEWORK

Assume that each of the input variables has a unique non-random value at each point in the aquifer, and that an observation of any one of these input variables at a point in the aquifer cannot be made accurately, but involves a measurement error. This error is considered to be purely random and to be free of bias. For example, if n points (x_i, y_i; i=1, ..., n) are chosen either randomly or by design in the aquifer and at each point an observation of permeability is made, then,

$$k_i = K_i + e_i \quad , \quad i = 1 \ldots n,$$

where, k_i = observed value of permeability at the point (x_i, y_i) (random and observable),

K_i = true value of permeability at the point (x_i, y_i) (non-random and unobservable),

e_i = measurement error at the point (x_i, y_i) (random and unobservable),

n = number of observations.

The errors, e_i, will be considered to be independent of K_i, mutually uncorrelated and normally distributed,(16) with mean zero and variance σ_K^2, that is,

$$e_i \sim N(0, \sigma_K^2), \quad i = 1 \ldots n,$$

and $\mathrm{Cov}(e_i, e_j) = 0$, $i = 1 \ldots n$, $j = 1 \ldots n$, $i \neq j$.

The variance, σ_K^2, of e_i is assumed constant for each observation at each point in the aquifer. It follows that,

$$k_i \sim N(K_i, \sigma_K^2), \quad i = 1 \ldots n \quad .$$

The spatial variation of each input variable is assumed to be expressable as a function of the space coordinates by an equation of the form (using permeability as an example),

$$K = \alpha_0 + \alpha_1 X_1 + \ldots + \alpha_{N-1} X_{N-1} \quad ,$$

where,

α_i = constant coefficient (unknown),

$X_i = X_i(x, y)$, where x, y are space coordinates,

N = number of terms in the equation necessary to closely approximate the spatial variation in K .

Then,

$$k_i = \alpha_0 + \alpha_1 X_{1,i} + \ldots + \alpha_{N-1} X_{N-1,i} + e_i, \quad i = 1 \ldots n,$$

or, in matrix notation,

$$\underline{k} = \underline{\alpha}' \, \underline{X} + \underline{e} \quad . \qquad (3)$$

The coordinates $(x_i, y_i; i = 1, \ldots, n)$ of each observation, $(k_i; i=1, \ldots n)$ will be considered to be observed without error, so that each of the variables $(X_{ji}(x_i, y_i); j=1, \ldots, N-1, i=1, \ldots, n)$ is known exactly, and equation (3) fits the normal theory of the General Linear Hypothesis Model of Full Rank, Model I, Case A* (see Appendix 2). Using this theory, maximum likelihood estimates can be found for $(\alpha_i; i=1, \ldots, N-1)$ and σ_K^2. If these estimates are designated $\hat{\underline{\alpha}}$ and $\hat{\sigma}_K^2$, they are shown in Appendix 2 to be given by,

$$\hat{\underline{\alpha}} = (\underline{X}' \underline{X})^{-1} \underline{X}' \underline{k}$$

where,

$$\hat{\underline{\alpha}} \sim N(\underline{\alpha}, \sigma_K^2 (\underline{X}' \underline{X})^{-1})$$

and,

$$\hat{\sigma}_K^2 = \frac{\underline{k}' \underline{k} - \hat{\underline{\alpha}}' \underline{X}' \underline{k}}{(n-N)}$$

where,

$$(n-N) \frac{\hat{\sigma}_K^2}{\sigma_K^2} \sim \chi^2 (n-N) \quad .$$

The input to the numerical model is the mean value of each variable in each grid. For permeability, the true mean value in a grid, $\overline{K}$, is given by

$$\overline{K} = \frac{1}{\Delta x \Delta y} \int_x^{x+\Delta x} \int_y^{y+\Delta y} (\alpha_0 + \alpha_1 X_1 + \ldots + \alpha_{N-1} X_{N-1}) dx\, dy,$$

where, the grid has coordinates (x, y), $(x, y+\Delta y)$, $(x+\Delta x, y)$, $(x+\Delta x, y+\Delta y)$.

---

*The linear model classification used in this dissertation is the same as that given by Graybill [7].

The estimate of $\overline{K}$ in this grid will be taken to be,

$$\stackrel{\approx}{K} = \frac{1}{\triangle x \triangle y} \int_{x}^{x+\triangle x} \int_{y}^{y+\triangle y} (\hat{\alpha}_0 + \hat{\alpha}_1 X_1 + \cdots + \hat{\alpha}_{N-1} X_{N-1}) dx\, dy \quad (4)$$

and is such that,

$$\stackrel{\approx}{K} \sim N(\overline{K}, \rho_K^2) \quad .$$

The standard deviation, ρ_K, will be taken as the measure of error in the estimate, $\stackrel{\approx}{K}$, of $\overline{K}$. It is shown in Appendix 3 that,

$$\rho_K = \frac{\sigma_K}{\sqrt{n}} f(X_{ji}) \qquad j = 1, \ldots, N-1, \quad i = 1, \ldots, n \quad ,$$

where, $f(X_{ji})$ = a known function of X_{ji}, $j = 1, \ldots, N-1$, $i = 1, \ldots, n$.

The estimate of ρ_K will then be $\tilde{\rho}_K$, where,

$$\tilde{\rho}_K = \frac{\hat{\sigma}_K}{\sqrt{n}} f(X_{ji}) \qquad j = 1, \ldots, N-1, \quad i=1, \ldots, n \qquad (5)$$

Equations analogous to equation (4) for each of the input variables provide a method of estimating the mean value of each of the input variables in every grid from available observations. The accuracy of these estimates is given by equations analogous to equation (5) and is inversely proportional to the square root of the number of observations used to make the estimate.

The solution to the numerical model for every time step involves the inversion of a matrix whose size is equal to the number of grids of the system being considered. This makes it practically impossible to determine directly the relations between the measures of error on the input variables, $\tilde{\rho}_K$, $\tilde{\rho}_b$ (or $\tilde{\rho}_z$), $\tilde{\rho}_{h_I}$, $\tilde{\rho}_Q$, $\tilde{\rho}_S$, and the accuracy of the final head after a number of time steps. In consequence, the following procedure was adopted to investigate the error relations.

EXPERIMENTAL PROCEDURE

In determining the relationship between the errors on the input variables and the error on the output, two general cases were considered: (i) each of the input variables, considered singly, was assumed to contain error and the other variables to be known exactly, (ii) all of the input variables, considered simultaneously, were assumed to contain error. The procedure for determining the relationship was basically the same for both cases, and is described first of all when only permeability, of the input variables, is considered to contain error and secondly when all of the input variables are simultaneously considered to contain error. It consists of the application of the Monte Carlo technique to generate a random sample from the density of h_F, the computation of a tolerance limit to be used as a measure of error on h_F from the error on the input variables.

When only permeability of the input variables contains error, the procedure is as follows:

STEP 1

A randomly generated value of $\tilde{\tilde{K}}$ and $\tilde{\rho}_K$ is assigned to each grid. The other variables are assigned random mean values and zero variances.

STEP 2

In every grid a random value of permeability is generated from its distribution, which is assumed to be normal, $N(\tilde{\tilde{K}}, \tilde{\rho}_K^2)$. Thus, it is implicitly assumed that the differences between $\tilde{\tilde{K}}$ and $\overline{K}$ and between $\tilde{\rho}_K$ and ρ_K do not influence the prediction of the final head and the determination of the accuracy of the prediction. This assumption has no effect on the determination of the error relations between the input and output variables, but is significant in the application of these error relations. This significance is discussed in the section on application of results.

STEP 3

With these random values of permeability and the fixed values of the other input variables, the deterministic model is used to solve for the head after a specified time.

STEP 4

Repeat STEPS 2 and 3 until a random sample of size M of values of head in every grid is generated, that is, $(h_F)_i$, $i=1,\ldots,M$.

STEP 5

In every grid determine the tolerance limits on h_F and the width of the tolerance limit, t. The theory of tolerance limits is developed in Appendix 4.

STEP 6

Randomly generate a new value of $\tilde{\rho}_K$ in every grid and repeat STEPS 2, 3, 4 and 5 until a sample of size m of tolerance limit widths, t, in each grid is obtained. Each width will correspond to a value of $\tilde{\rho}_K$.

STEP 7

Using the theory of the Regression Model, Model III, Case 2*, find a predictive relation between t and $\tilde{\rho}_K$.

The theory of the Regression Model, Model III, Case 2, is described in Appendix 2, and to apply the theory of the model to find the relation between t and $\tilde{\rho}_K$ it is assumed that the joint density of t and $\tilde{\rho}_K$ is given by,

$$f(t, \tilde{\rho}_K) = h(\tilde{\rho}_K) \frac{1}{\sqrt{2\pi}\, \sigma_{t'}} \exp\left\{-\frac{1}{2}\left(\frac{t - G(\tilde{\rho}_K)}{\sigma_{t'}}\right)^2\right\} \tag{6}$$

where,

$G(\tilde{\rho}_K)$ is a linear (in the coefficients α_i) function of $\tilde{\rho}_K$, and $h(\tilde{\rho}_K)$ is the marginal density of $\tilde{\rho}_K$ and does not contain α_i, $i=1\ldots N$ or $\sigma_{t'}$. In this study $G(\tilde{\rho}_K)$ was taken to be a polynomial, so that, $G(\tilde{\rho}_K) = \sum_{i=0}^{N} \alpha_i (\tilde{\rho}_K)^i$.

It follows from equation (6) that the conditional distribution of t given $\tilde{\rho}_K = \tilde{\rho}^*_K$ is f(t'), where,

---

*Following Graybill's [7] classification.

$$f(t') = f(t/\tilde{\rho}_K = \tilde{\rho}_K^*) = \frac{1}{\sqrt{2\pi}\,\sigma_{t'}} \exp\left\{-\frac{1}{2}\left(\frac{t - \sum_{i=0}^{N} \alpha_i (\tilde{\rho}_K^*)^i}{\sigma_{t'}}\right)^2\right\},$$

so that,

$$E(t') = E(t/\tilde{\rho}_K = \tilde{\rho}_K^*) = \sum_{i=0}^{N} \alpha_i (\tilde{\rho}_K^*)^i \quad .$$

The coefficients α_i, $i=1....N$ can now be estimated and the value of N determined such that the "best" predictive relation between $E(t')$ and $\tilde{\rho}_K^*$ is obtained. This was done in the following manner.

In every grid, the linear equation,

$$E(t') = \alpha_0 + \alpha_1 \tilde{\rho}_K^* \quad ,$$

was fitted to the data $(t_j, (\tilde{\rho}_K)_j; j=1....m)$. A test of the hypothesis, $H_0: \alpha_1 = 0$, was made. If this hypothesis was accepted then it was concluded that the data were fitted better by $E(t') =$ constant than by the linear equation being considered. If the hypothesis was rejected it was concluded that the data were better represented by the linear equation. In a similar way the equations,

$$E(t') = \beta_0 + \beta_1 \tilde{\rho}_K^* + \beta_2 (\tilde{\rho}_K^*)^2$$

and

$$E(t') = \gamma_0 + \gamma_1 \tilde{\rho}_K^* + \gamma_2 (\tilde{\rho}_K^*)^2 + \gamma_3 (\tilde{\rho}_K^*)^3 \quad ,$$

were fitted to the data, and the hypotheses, $H_0:\beta_2 = 0$ and $H_0:\gamma_3 = 0$ were tested. These tests determine whether the data are better represented by a quadratic or linear equation and cubic or quadratic equation respectively. If, for instance, the hypothesis, $H_0:\alpha_1 = 0$ were rejected and the hypotheses, $H_0:\beta_2 = 0$, $H_0:\gamma_3 = 0$, were accepted, then it was concluded that the data were "best" represented by a linear equation. In general, before accepting that a polynomial of a given order "best" represents the data it is necessary to accept the null hypotheses on two polynomials of immediately higher order.

The data generated in the above procedure was used to estimate the coefficient of variation of h_F, C_{h_F}, by,

$$\tilde{C}_{h_F} = \frac{\frac{1}{(M-1)} \sum_{i=1}^{M} (h_{F_i} - \mu_{h_F})^2}{\mu_{h_F}} ,$$

$$\mu_{h_F} = \frac{1}{M} \sum_{i=1}^{M} h_{F_i} ,$$

and the coefficient of variation of permeability, C_K, by

$$\tilde{C}_K = \frac{\tilde{\rho}_K}{\approx K} .$$

These were also regarded as measures of error on the input and output variables. They were considered in addition to the previously described measures of error because they are dimensionless. The "best" polynomial relating $\tilde{C}_{h_F}$ and $\tilde{C}_K$ was then determined using exactly the same procedure as the one described above.

When all the input variables are simultaneously considered to contain error the procedure for determining a predictive equation between these errors and the error on output is described below. The notation is adopted that $\tilde{\rho}_1 = \tilde{\rho}_K$, $\tilde{\rho}_2 = \tilde{\rho}_b$ (or $\tilde{\rho}_z$), $\tilde{\rho}_3 = \tilde{\rho}_Q$, $\tilde{\rho}_4 = \tilde{\rho}_{h_I}$, $\tilde{\rho}_5 = \tilde{\rho}_S$.

STEP 1

A randomly generated value of $\overset{\approx}{K}$, $\overset{\approx}{b}$ (or $\overset{\approx}{z}$), $\overset{\approx}{Q}$, $\overset{\approx}{h}_I$, $\overset{\approx}{S}$ and $\tilde{\rho}_i$, $i=1....5$ is assigned to each grid.

STEP 2

In every grid a random value of K, b (or z), Q, h_I and S is generated from its distribution, which is assumed to be normal.

STEP 3

With these values of the input variables, the deterministic model is used to solve for the head after a specified time.

STEP 4

Repeat STEPS 2 and 3 until a random sample of size M of values of head in every grid is generated, that is, $(h_F)_i$, $i=1....M$.

STEP 5

In every grid determine the tolerance limits on h_F and the width of the tolerance limit, t.

STEP 6

Randomly generate a new value of $\tilde{\rho}_i$, $i=1....5$ in every grid and repeat STEPS 2, 3, 4 and 5 until a sample m of tolerance limit widths, t, in each grid is obtained. Each width will correspond to a set of values of $\tilde{\rho}_i$, $i=1....5$.

STEP 7

Using the theory of the Regression Model, Model III, Case 2, find a predictive relation between t and $\tilde{\rho}_i$, $i=1....5$.

To fit the theory of the Regression Model for this multivariate case, it is assumed that the joint density of t and $\tilde{\rho}_i$, $i=1....5$ is given by,

$$f(t, \tilde{\rho}_i, i=1....5) = h(\tilde{\rho}_i, i=1....5) \frac{1}{\sqrt{2\pi}\,\sigma_{t''}} \exp\left\{-\frac{1}{2} \cdot \left(\frac{t-G(\tilde{\rho}_i, i=1....5)}{\sigma_{t''}}\right)^2\right\} \tag{7}$$

where,

$G(\tilde{\rho}_i, i=1....5)$ is a linear (in the coefficients α_i) function of $\tilde{\rho}_i$, $i=1....5$, and, $h(\tilde{\rho}_i, i=1....5)$ is the marginal density of $\tilde{\rho}_i$,

$i=1....5$ and does not contain α_i, $i=1....N$ or $\sigma_{t''}$. In this study $G(\cdot)$ was taken to be a polynomial in $\tilde{\rho}_i$, $i=1....5$.

It follows from equation (7) that the conditional distribution of t given ($\tilde{\rho}_i = \tilde{\rho}_i^*$, $i=1....5$) is f(t''), where,

$$f(t'') = f(t/\tilde{\rho}_i = \tilde{\rho}_i^*, \; i=1....5) =$$

$$\frac{1}{\sqrt{2\pi}\,\sigma_{t''}} \exp\left\{-\frac{1}{2}\left(\frac{t - G(\tilde{\rho}_i^*, \; i=1....5)}{\sigma_{t''}}\right)^2\right\}$$

so that,

$$E(t'') = E(t/\tilde{\rho}_i = \tilde{\rho}_i^*, \; i=1....5) = G(\tilde{\rho}_i^*, \; i=1....5) \; .$$

It is now necessary to estimate the coefficients of the polynomial $G(\cdot)$ and determine which of them are significantly different from zero; that is, determine which terms in the polynomial have to be considered in order to adequately represent the data. This was done when $G(\cdot)$ was assumed to be a polynomial consisting only of linear (in the variables) terms and when it consisted of linear and quadratic (in the variables) terms. These two polynomials can be written explicitly as,

$$G(\cdot) = \sum_{i=0}^{5} \alpha_i \tilde{\rho}_i^* \; , \text{ where, } \tilde{\rho}_0 = 1 \; ,$$

and

$$G(\cdot) = \sum_{i=0}^{5} \alpha_i \tilde{\rho}_i^* + \sum_{i=1}^{5}\left(\sum_{j=1}^{5} \alpha_L \tilde{\rho}_i^* \tilde{\rho}_j^*\right), \; L = 6....20, \; \tilde{\rho}_0^* = 1.$$

For each polynomial, tests of hypotheses were made on the coefficients α_i, both individually and simultaneously, to determine which coefficients were significantly different from zero. These tests are described along with the results.

A regression model involving only errors in initial heads was also investigated. This is described together with the reasons why it was studied in the section on results.

For each of the regression models, involving tolerance interval widths and estimates of standard deviations, that has been described, an exactly similar model was studied relating estimates of the coefficients of variation of the input variables and the final head.

The results of the investigation of all of the above regression models are presented in the following section, and the computer program used for the investigation is described in Appendix 5.

RESULTS

I. CONFINED FLOW

I.A. Each Input Variable Considered Singly

A confined aquifer, divided into 20 square grids, was used to determine the predictive relations between the errors in the input variables, considered singly, and the error in the output. Six variations of this 20 grid model were considered, which differed in boundary conditions and the "randomness" of the data used. They are described in Appendix 1.

For each of the 20 grids, in each of the 6 models, and for each of the 5 input variables, the "best" polynomial relating errors on h_F to errors on the estimates of the input variables were determined. The results, when the measures of error on h_F are tolerance interval widths, are summarized in Table 1. The entries in the table are the number of grids (out of 20) in which the "best" polynomial was linear, quadratic or cubic.

For some of the models the "best" polynomial was determined between the estimates of the coefficient of variation of the input and output variables. The results are given in Table 2. In this table, α_1 is the "gradient" coefficient in the linear equation. It indicates whether the error on the output variable is less than or greater than the error on the input variable. Since the coefficients of variation are dimensionless, a value of α_1 less than unity indicates that the error on output is less than the error on input and a value greater than one indicates that the output error is greater than the input error.

The following observations can be made from the results given in Tables 1 and 2:

1) the majority of "best" polynomials, for all five input variables, are linear, in both the "tolerance-interval-width" regression model and the "coefficient-of-variation" regression model;

2) for initial head, the mean value of α_1 is .9747, indicating that the error in initial head has an approximate one-to-one relation (slope of regression line approximately unity) with the error in final head;

3) the mean values of α_1 for the other four input variables indicate that the error on final head is two or three orders of magnitude less than the input errors on these variables.

These observations are made from results obtained for comparatively short periods of time and so are valid only for these time periods.

TABLE 1. INPUT VARIABLES CONSIDERED SINGLY. "BEST" POLYNOMIALS FOR "TOLERANCE-INTERVAL-WIDTH" REGRESSION MODEL; CONFINED FLOW.

INPUT VARIABLE CONTAINING ERROR		TYPE OF POLYNOMIAL	MODEL* (i)	(ii)	(iii)	(iv)	(v)	(vi)
	K	LINEAR	16	18	16	18	19	17
		QUADRATIC	3	2	1	0	1	0
		CUBIC	1	0	3	2	0	3
	h_I	LINEAR	16	16	16	17	16	16
		QUADRATIC	1	1	1	1	1	1
		CUBIC	3	3	3	2	3	3
	S	LINEAR	NO	19	19	19	18	18
		QUADRATIC	RUN	1	1	1	1	1
		CUBIC		0	0	0	1	1
	b	LINEAR	19	19	18	18	20	17
		QUADRATIC	1	0	0	1	0	1
		CUBIC	0	1	2	1	0	2
	Q	LINEAR	17	17	17	17	17	17
		QUADRATIC	2	2	2	2	2	2
		CUBIC	1	1	1	1	1	1

Entries in the table are, for each model and each input variable, the number of grids (out of 20) in which the "best" polynomial was linear, quadratic or cubic.

*Each model is described in Appendix 1.

TABLE 2. INPUT VARIABLES CONSIDERED SINGLY. "BEST" POLYNOMIALS FOR "COEFFICIENTS-OF-VARIATION" REGRESSION MODEL; CONFINED FLOW.

		TYPE OF POLYNOMIAL			RANGE ON a_1* IN THE 20 GRIDS	MEAN OF a_1 IN THE 20 GRIDS
		LINEAR	QUADRATIC	CUBIC		
INPUT VARIABLE CONTAINING ERROR	K Model (v)	19	1	0	.001533 .015160	.006447
	h_I Model (iii)	17	0	3	.830 1.093	.9747
	S Model (v)	19	0	1	.000054 .004045	.001644
	b Model (v)	18	2	0	.000103 .007558	.00216
	Q Model (v)	19	1	0	.000378 .002495	.00125

Integer entries in the table are, for each input variable, the number of grids (out of 20) in which the 'best' polynomial was linear, quadratic or cubic.

*a_1 is the "gradient coefficient" in the linear polynomial.

I. B. Input Variables Considered Simultaneously

A confined aquifer, divided into 20 square grids, was used to determine the relations between the errors in the input variables, considered simultaneously, and the output variable. Two basic regression models were considered, a 6-coefficient model and a 21-coefficient model. In terms of tolerance interval width and estimates of standard deviations these can be written,

$$E(t') = \sum_{i=0}^{5} \alpha_i \tilde{\rho}_i^* \quad , \quad \text{where } \tilde{\rho}_0^* = 1 \ ,$$

and

$$E(t') = \sum_{i=0}^{5} \alpha_i \tilde{\rho}_i^* + \sum_{i=1}^{5} \left(\sum_{j=i}^{5} \alpha_L \tilde{\rho}_i^* \tilde{\rho}_j^* \right) , \quad L = 6. \ldots 20, \ \tilde{\rho}_0^* = 1.$$

Similar models relating estimates of coefficients of variation on input and output were also investigated, so that four models in all were studied.

The 6-coefficient model was studied over four time steps covering a 110 day period and the 21-coefficient model over one time step of 20 days.

The boundary conditions and data used to study these models are given in Appendix 1, and the results are summarized in Tables 3, 4, 5, 6 and 7. Table 3 gives the number of grids (out of 20), in which the hypotheses, $H_0:[(\alpha_i=0,\ \alpha_j \text{ unspecified}),\ i=0....5,$ $j=0....5,\ j\neq i]$, $H_0:[(\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_5) = 0,\ \alpha_4 \text{ unspecified}],$

$H_0: [(\alpha_1, \alpha_2, \alpha_3, \alpha_5) = 0, (\alpha_0, \alpha_4)$ unspecified], were rejected at the 95% level in the 6-coefficient regression models. Table 4 indicates whether, for the 6-variable "coefficient-of-variation" regression model, the estimates of the regression coefficients tended to increase or decrease with time and whether they tended to be positive or negative throughout the study period. Table 5, for the 6-variable model, indicates the way in which the sum of squares of deviations from the regression line changed with time, and Table 6 gives typical estimates of the regression coefficients in one grid for the 6-variable "coefficient-of-variation" regression model. Table 7 gives the number of grids (out of 20) in which the hypotheses $H_0(\alpha_i = 0,\ i=0 \dots 20)$ were rejected at the 95% level in the 21-coefficient regression model.

The following observations can be made from the results given in Tables 3-7:

1) from Table 3 it can be seen that only the error in initial head is of significance in predicting the error in final head up to 110 days, and that the predictive relation is linear and, from Table 6, initially one-to-one,

2) from Table 3 the coefficients α_0, α_1, α_2, α_3, α_5 are non-significant from zero both individually and simultaneously,

3) from Tables 4 and 6, the regression coefficient associated with error in initial head, α_4, decreases monotonically with time, but is always positive,

4) from Tables 4 and 6, the constant regression coefficient, α_0, tends to increase with time, but up to 110 days does not become significantly different from zero,

5) from Table 4, the regression coefficients α_1, α_2, α_3, α_5 do not show any discernable trends and are as liable to be negative as positive,

6) from Table 5, the sum of squares of deviations tends to decrease with time up to 110 days,

7) from Table 7, none of the product terms introduced by using the 21-coefficient model is significantly different from zero and the error in initial head remains the only significant input error in predicting the error in final head.

For the 21-variable "coefficients-of-variation" regression model, the hypothesis, H_0:[$(\alpha_0, \ldots \alpha_3, \alpha_5, \ldots \alpha_{20}) = 0$], was tested at the 95% level and accepted in 19 of the 20 grids, indicating that these coefficients are simultaneously non-significant, and that for this regression model, as well as the 6-variable model, the error in final head is linearly related to the error in initial head.

The results for the 21-coefficient regression model given in Table 7 were obtained using 100 data points. A run was made using

TABLE 3. INPUT VARIABLES CONSIDERED SIMULTANEOUSLY; 6-COEFFICIENT REGRESSION MODEL; CONFINED FLOW.

		'TOLERANCE-INTERVAL-WIDTH' REGRESSION MODEL RESULTS AT END OF TIME STEP				'COEFFICIENT-OF-VARIATION' REGRESSION MODEL RESULTS AT END OF TIME STEP			
	REGRESSION COEFFICIENT	1	2	3	4	1	2	3	4
NUMBER OF GRIDS,	a_0	0	0	0	0	0	0	0	0
OUT OF 20, IN	a_1	0	1	0	0	0	1	0	0
WHICH REJECT	a_2	1	0	0	0	1	1	1	1
H_o: (a_i = 0) AT	a_3	2	2	2	2	0	0	0	0
95% LEVEL, AT	a_4	20	20	20	19	19	19	19	19
THE END OF EACH TIME STEP	a_5	0	0	0	1	2	2	2	2
NUMBER OF GRIDS IN WHICH REJECT H_o: (a_0, a_1, a_2, a_3, a_5)=0 AT 95% LEVEL AT END OF EACH TIME STEP		0	0	0	0	1	2	2	2
NUMBER OF GRIDS IN WHICH REJECT H_o: (a_1, a_2, a_3, a_5)=0 AT 95% LEVEL AT END OF EACH TIME STEP		0	0	0	0	2	2	2	2

TABLE 4. CONFINED FLOW. INPUT VARIABLES CONSIDERED SIMULTANEOUSLY. NATURE OF ESTIMATES OF REGRESSION COEFFICIENTS IN 6-VARIABLE, 'COEFFICIENTS-OF-VARIATION' MODEL.

	TIME CHANGE OF ESTIMATES OF REGRESSION COEFFICIENT			NO OF ESTIMATES WHICH WERE ENTIRELY POSITIVE OR NEGATIVE WITH TIME		
	INCREASE	DECREASE	BOTH	POSITIVE	NEGATIVE	BOTH
α_0	13	5	2	12	8	0
α_1	11	7	1	7	12	1
α_2	9	8	3	7	13	0
α_3	9	8	3	9	9	2
α_4	0	20	0	20	0	0
α_5	10	7	3	11	8	1

TABLE 5. CONFINED FLOW. INPUT VARIABLES CONSIDERED SIMULTANEOUSLY. TIME CHANGE OF SUM OF SQUARES OF DEVIATIONS IN 6-COEFFICIENT MODEL.

	INCREASE	DECREASE	MINIMUM	MAXIMUM
'TOLERANCE-INTERVAL-WIDTH' REGRESSION MODEL	2	15	2	1
'COEFFICIENT-OF-VARIATION' REGRESSION MODEL	2	16	1	1

TABLE 6. TYPICAL ESTIMATES OF THE REGRESSION COEFFICIENTS IN THE 6-VARIABLE 'COEFFICIENTS-OF-VARIATION' MODEL. INPUT VARIABLES CONSIDERED SIMULTANEOUSLY; CONFINED FLOW.

REGRESSION COEFFICIENT	RESULTS AT END OF TIME STEP: 1	2	3	4
a_0	.001091	.001566	.002437	.003229
a_1	-.00077	-.00079	-.00091	-.00109
a_2	-.02295	-.02484	-.02784	-.03000
a_3	-.00787	-.00922	-.01137	-.01289
a_4	.9795	.9500	.8938	.8398
a_5	.00956	.00910	.00863	.00864

TABLE 7. CONFINED FLOW. INPUT VARIABLES CONSIDERED SIMULTANEOUSLY. 21-COEFFICIENT REGRESSION MODELS.

	'TOLERANCE-INTERVAL-WIDTH' REGRESSION MODEL	'COEFFICIENT-OF-VARIATION' REGRESSION MODEL
a_0	0	0
a_1	1^1	1
a_2	0	1^1
a_3	1	1^1
a_4	19	19
a_5	0	1^1
a_6	1^1	1^1
a_7	1^1	1^1
a_8	1^1	1^1
a_9	0	1
a_{10}	0	0
a_{11}	0	0
a_{12}	4^3	3^2
a_{13}	1	1^1
a_{14}	0	2^1
a_{15}	2^2	2^1
a_{16}	1^1	2^1
a_{17}	1	2
a_{18}	3^2	3^1
a_{19}	0	1
a_{20}	0	2

Entries in table are number of grids (out of 20) in which hypothesis H_o:$a_i = 0$ was rejected at 95% level.

Superscripts on entries in table are number of times coefficient was negative and significantly different from zero.

only 40 data points but did not give definitive results. Compared to Table 7, the number of times the null hypothesis was rejected on this run was fewer for all of the variables. This indicates that the power of the test of the hypothesis H_0: [$\alpha_i = 0$, $i=0 \ldots . 20$], increases with increasing number of data.

In view of the results obtained with these two regression models, indicating the dominant influence of the error in initial head, a regression model involving only errors in initial head was investigated.

For any one grid it was defined by,

$$E(t_0') = \alpha_0 + \alpha_1(\tilde{\rho}^*_{h_I})_0 + \alpha_2(\tilde{\rho}^*_{h_I})_1 + \alpha_3(\tilde{\rho}^*_{h_I})_2 + \alpha_4(\tilde{\rho}^*_{h_I})_3$$

$$+ \alpha_5(\tilde{\rho}^*_{h_I})_4 \quad ,$$

with a similar model relating the estimates of the coefficients of variation of input and output. The subscripts refer to a pattern of grids as in Figure 2. This model takes into account the influence that errors in initial head in neighboring grids have on the error in final head in any one grid. The data used was the same as for the 6-variable regression model which has just been described. The results are given in Tables 8, 9, 10.

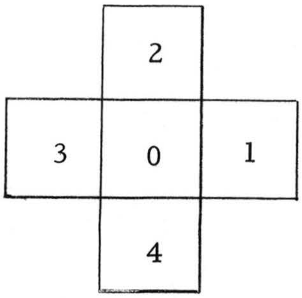


FIGURE 2

The nature of the 20 grid model meant that only grids 7, 8, 9, 12, 13, 14 (see Figure 3) had four neighboring grids as shown in Figure 3. Grids 1, 5, 16, 20 had two neighboring grids and grids 2, 3, 4, 6, 10, 11, 15, 17, 18, 19, had 3 neighboring grids. For this reason, depending on which grid was being considered, the regression model involved 6, 5, or 4 coefficients. Thus, in Tables 8 and 9, α_0 is the constant coefficient, α_1 is the coefficient associated with error in initial head in the grid being considered and α^* consists of all the coefficients associated with errors in initial head in neighboring grids. For all 20 grids, there are therefore, 20 estimates of α_0 and α_1 and 62 estimates of α^*.

1	2	3	4	5
6	7	8	9	10
11	12	13	14	15
16	17	18	19	20

FIGURE 3

Table 8 gives the number of times that the hypotheses, $H_0:(\alpha_i = 0,\ i=0,\ 1,\ \alpha_i \subset \alpha^*)$, $H_0: [\alpha_0, \alpha^*] = 0$, $H_0:[\alpha^*] = 0$, were rejected at the 95% level. Table 9 indicates whether, in the "coefficient-of-variation" model, the estimates of the regression coefficients tended to increase or decrease with time, and whether they tended to be positive or negative throughout the study period. Table 10 indicates the way in which the sum of squares of deviations from the regression line changed with time.

The following observations can be made from the results given in Tables 8, 9, 10:

1) from Table 8 it can be seen that only the error in initial head in the grid being considered is of significance in predicting the error in final head up to 110 days, and that the predictive equation is linear. Initially, it is also approximately a one-to-one relation,

2) from Table 8, the coefficients α_0, α^* are non-significant both individually and simultaneously in predicting the error in final head up to 110 days,

3) from Table 9, the regression coefficient associated with error in initial head, α_1, decreases monotonically with time but is always positive,

4) from Table 9, the constant coefficient, α_0, shows a tendency to increase with time, but up to 110 days does not become significantly different from zero,

5) from Table 9, the coefficients associated with errors in initial head, α^*, do not show any discernable trend and are as liable to be negative as positive,

6) from Table 10, the sum of squares of deviations from the regression line tends to decrease with time up to 110 days.

TABLE 8. CONFINED FLOW. RESULTS FOR REGRESSION MODEL INVOLVING ERRORS ON INITIAL HEAD.

	'TOLERANCE-INTERVAL-WIDTH' REGRESSION MODEL				'COEFFICIENT-OF-VARIATION' REGRESSION MODEL			
COEFF. \ TIME STEP	1	2	3	4	1	2	3	4
α_0	0	0	0	0	0	0	0	0
α_1	18	18	18	18	18	18	18	18
α^*	3^2	3^2	4^3	3^2	0	0	1^1	1^1
NUMBER OF GRIDS IN WHICH REJECT H_o: $[\alpha_o, \alpha^*] = 0$ AT 95% LEVEL	2	2	2	2	0	0	0	0
NUMBER OF GRIDS IN WHICH REJECT H_o: $[\alpha^*] = 0$ AT 95% LEVEL	2	2	2	2	0	0	0	0

Superscripts on entries in table are number of times regression coefficient was negative and significantly different from zero.

α^* consists of all coefficients associated with errors in initial head in neighboring grids.

The entries in the table opposite α^* are the number of times (out of 62) that the hypothesis H_o: $\alpha_i = 0$ was rejected for any of the coefficient in α^*.

The entries in the table opposite α_0, α_1, are the number of grids (out of 20) in which the hypothesis, H_o: $\alpha_i = 0$, $i = 0, 1$, was rejected.

TABLE 9. CONFINED FLOW. NATURE OF ESTIMATES OF REGRESSION COEFFICIENTS IN THE 'COEFFICIENTS-OF-VARIATION' MODEL INVOLVING ERRORS ON INITIAL HEAD.

	TIME CHANGE OF ESTIMATES OF REGRESSION COEFFICIENTS			NO. OF ESTIMATES WHICH WERE ENTIRELY POSITIVE OR NEGATIVE WITH TIME		
	INCREASE	DECREASE	BOTH	POSITIVE	NEGATIVE	BOTH
a_0	12	6	2	8	11	1
a_1	0	20	0	20	0	0
a^*	26	30	6	33	28	1

TABLE 10. CONFINED FLOW. TIME CHANGE OF SUM OF SQUARES OF DEVIATIONS IN REGRESSION MODEL INVOLVING ERRORS ON INITIAL HEAD.

	INCREASE	DECREASE	MINIMUM
'TOLERANCE-INTERVAL-WIDTH' REGRESSION MODEL	2	16	2
'COEFFICIENT-OF-VARIATION' REGRESSION MODEL	1	19	0

A comparison of the sum of squares of deviations in the above model involving only errors in initial head and the previous 6-variable model, shows that the sum of squares was less throughout the study period in the 6-variable model in about 60% of the grids and greater in the other 40%.

II. UNCONFINED FLOW

II. A. Input Variables Considered Simultaneously

A confined aquifer, divided into 20 square grids, was used to determine the relations between the errors in input variables, considered simultaneously, and the output variables. The following regression model was studied,

$$E(t'') = \alpha_0 + \alpha_1 \tilde{\rho}^*_K + \alpha_2 \tilde{\rho}^*_z + \alpha_3 \tilde{\rho}^*_Q + \alpha_4 \tilde{\rho}^*_{h_I} + \alpha_5 \tilde{\rho}^*_S$$

and a similar model relating estimates of coefficients of variation on input and output variables.

Results were obtained for this model with runs of one time step (20 days), 6 time steps (240 days) and 10 time steps (440 days). The boundary conditions and data used in these runs are given in Appendix 1. The boundary conditions were defined to be constant gradients throughout a time step, but were allowed to change randomly with each time step. Impermeable boundaries and constant head boundaries are special cases of such boundary conditions, and so they are considered to be quite general.

The results are summarized on Tables 11-18. Tables 11, 12, 15 give the number of grids in which the hypotheses, H_0: $[(\alpha_i = 0, \alpha_j$ unspecified), i=0.... 5, j=0.... 5, j$\neq$ i] and H_0:$[(\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_5) = 0$, α_4 unspecified] were rejected at the 95% level for the 1, 6 and 10 time step runs respectively. Table 15 also gives similar results for the hypothesis, H_0:$[(\alpha_1, \alpha_2, \alpha_3, \alpha_5) = 0$, (α_0, α_4) unspecified]. Tables 13 and 16 give, for the "coefficient-of-variation" regression model, the number of grids in which the estimates of the regression coefficients increased or decreased with time, and the number of grids in which the estimates of the regression coefficients were either positive or negative throughout the study period for the 6 and 10 time step runs respectively. Table 14 gives, for the "tolerance-interval-width" regression model, the way in which the sum of squares of deviations from the regression line changed during the study period in the 6 time step run. Table 17 gives the way in which the sum of squares of deviations changed in the 10 time step run and Table 18 gives a typical set of estimates of the regression coefficients for one grid in the 10 time step run of the "coefficients-of-variation" model.

The following observations can be made from the results given in Tables 11-18:

1) from Tables 11, 12, 15 it can be seen that the error in initial head is significant in predicting the error in final

head up to 440 days, and that up to about 140 days the relation is linear, and from Table 18 it is initially one-to-one,

2) from Tables 11, 12, 15, the coefficients α_1, α_2, α_3, α_5 are not significantly different from zero both individually and simultaneously up to 440 days,

3) from Tables 13, 16, and 18, the constant coefficient, α_0, tends to increase with time, and from Tables 12 and 15 it can be seen that after about 200 days it can no longer be considered to be not significantly different from zero,

4) from Tables 13, 16 and 18 the regression coefficient associated with the error in initial head, α_4, decreases monotonically with time but is always positive,

5) from Tables 13, 16 and 18 the coefficients α_1, α_2, α_3, α_5, do not show any discernable trends and are as liable to be negative as positive,

6) from Table 14, the sum of squares of deviations tends to decrease with time up to 240 days, and Table 17 indicates that in general this decrease continues up to 440 days, but that in an increasing number of grids a turning point is reached.

As in the study of confined flow, because of the dominant influence of the error in initial head in predicting error in final

TABLE 11. UNCONFINED FLOW. INPUT VARIABLES CONSIDERED SIMULTANEOUSLY. 1 TIME STEP (20 DAYS)

	COEFF. \ TIME STEP	'TOLERANCE-INTERVAL-WIDTH' REGRESSION MODEL 1	'COEFFICIENTS-OF-VARIATION' REGRESSION MODEL 1
NUMBER OF GRIDS (OUT OF 20) IN WHICH REJECT $H_o: a_i = 0$ AT 95% LEVEL	a_0	1	0
	a_1	0	0
	a_2	1	1
	a_3	0	0
	a_4	20	20
	a_5	0	1

NUMBER OF GRIDS IN WHICH REJECT H_o: $[a_0\ a_1\ a_2\ a_3\ a_5] = 0$ AT 95% LEVEL	0	0

TABLE 12a. UNCONFINED FLOW. INPUT VARIABLES CONSIDERED SIMULTANEOUSLY. 6 TIME STEP (240 DAYS)

	COEFF. \ TIME STEP	'TOLERANCE-INTERVAL-WIDTH' REGRESSION MODEL 1	2	3	4	5	6
NUMBER OF GRIDS (OUT OF 20) IN WHICH REJECT $H_o: \alpha_i = 0$ AT 95% LEVEL	α_0	1^1	0	0	0	2	3
	α_1	1	0	1	1	1	2
	α_2	0	1^1	1^1	0	0	0
	α_3	1^1	1^1	2^2	1^1	1^1	1^1
	α_4	20	20	20	20	20	20
	α_5	1	1	1	1	1	1
NUMBER OF GRIDS IN WHICH REJECT $H_o: [\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_5] = 0$ AT 95% LEVEL		0	0	1	2	4	7

Superscripts on entries in table are the number of times the coefficient was negative and significantly different from zero.

TABLE 12b. UNCONFINED FLOW. INPUT VARIABLES CONSIDERED SIMULTANEOUSLY. 6 TIME STEPS (240 DAYS)

	COEFF. \ TIME STEP	'COEFFICIENT-OF-VARIATION' REGRESSION MODEL 1	2	3	4	5	6
NUMBER OF GRIDS (OUT OF 20) IN WHICH REJECT $H_o: a_i = 0$ AT 95% LEVEL	a_0	0	0	0	1	1	4
	a_1	2	1	0	1	0	0
	a_2	0	0	0	0	0	0
	a_3	0	0	0	0	0	0
	a_4	20	20	20	20	20	20
	a_5	0	1^1	1^1	1^1	1^1	0

	1	2	3	4	5	6
NUMBER OF GRIDS IN WHICH REJECT $H_o: [a_0, a_1, a_2, a_3, a_5] = 0$ AT 95% LEVEL	0	0	1	3	5	8

Superscripts on entries in table are number of times the coefficient was negative and significantly different from zero.

TABLE 13. UNCONFINED FLOW. INPUT VARIABLES CONSIDERED SIMULTANEOUSLY. 6 TIME STEPS. NATURE OF ESTIMATES OF REGRESSION COEFFICIENTS IN 'COEFFICIENTS-OF-VARIATION' MODEL.

	TIME CHANGE OF ESTIMATES OF REGRESSION COEFFICIENTS			NO. OF ESTIMATES WHICH WERE ENTIRELY POSITIVE OR NEGATIVE WITH TIME		
	INCREASE	DECREASE	BOTH	POSITIVE	NEGATIVE	BOTH
a_0	19	0	1	6	2	12
a_1	6	7	7	11	7	2
a_2	8	6	6	10	7	3
a_3	12	3	5	4	8	8
a_4	0	20	0	20	0	0
a_5	5	11	4	11	5	4

TABLE 14. UNCONFINED FLOW. INPUT VARIABLES CONSIDERED SIMULTANEOUSLY. 6 TIME STEPS. TIME CHANGE OF SUM OF SQUARES OF DEVIATIONS.

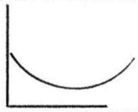
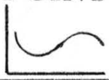
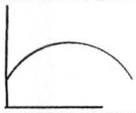
	DECREASE 	MINIMUM 	2 TURNING POINTS 	MAXIMUM 
'TOLERANCE-INTERVAL-WIDTH' REGRESSION MODEL	16	2	1	1

TABLE 15a. UNCONFINED FLOW. INPUT VARIABLES CONSIDERED SIMULTANEOUSLY. 10 TIME STEPS (440 DAYS).

		'TOLERANCE-INTERVAL-WIDTH' REGRESSION MODEL									
	COEFF. \ TIME STEP	1	2	3	4	5	6	7	8	9	10
NUMBER OF GRIDS	α_0	2^1	2^1	1	1	2	3	5	5	6	6
	α_1	1	1^1	1^1	1^1	1^1	1^1	1^1	2^1	1^1	1^1
(OUT OF 20) IN WHICH	α_2	0	0	0	0	0	0	0	0	1^1	1^1
REJECT	α_3	3^1	3^1	2	3^1	3^1	3^1	1^1	1^1	1^1	2^1
$H_o : \alpha_i = 0$	α_4	20	20	20	20	20	20	20	20	18	17
AT 95% LEVEL	α_5	0	0	0	1	1	0	0	0	1^1	1^1
NUMBER OF GRIDS IN WHICH REJECT $H_o : (\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_5) = 0$ AT 95% LEVEL		3	3	3	4	4	6	7	13	15	17
NUMBER OF GRIDS IN WHICH REJECT $H_o : (\alpha_1, \alpha_2, \alpha_3, \alpha_5) = 0$ AT 95% LEVEL		3	3	3	3	1	1	0	0	0	0

Superscripts on entries in table are number of times the coefficients are negative and significantly different from zero.

TABLE 15b. UNCONFINED FLOW. INPUT VARIABLES CONSIDERED SIMULTANEOUSLY. 10 TIME STEPS (440 DAYS).

	COEFF. \ TIME STEP	'COEFFICIENT-OF-VARIATION' REGRESSION MODEL 1	2	3	4	5	6	7	8	9	10
NUMBER OF GRIDS (OUT OF 20) IN WHICH REJECT $H_o : \alpha_i = 0$ AT 95% LEVEL	α_0	0	1	1	1	1	4	6	7	8	9
	α_1	1^{1}	0	0	2^{1}	2^{1}	2^{1}	3^{2}	3^{2}	2^{2}	2^{2}
	α_2	0	0	0	0	0	0	0	0	0	1^{1}
	α_3	1	1	1	2	2	2	2	2	2	1
	α_4	20	20	20	20	20	20	20	20	20	20
	α_5	0	0	0	1^{1}	1^{1}	1^{1}	1^{1}	1^{1}	2^{1}	2^{1}

NUMBER OF GRIDS IN WHICH REJECT $H_o : (\alpha_0, \alpha_1, \alpha_2, \alpha_3, \alpha_5) = 0$ AT 95% LEVEL	1	1	2	3	8	9	13	16	17	18

NUMBER OF GRIDS IN WHICH REJECT $H_o : (\alpha_1, \alpha_2, \alpha_3, \alpha_5) = 0$ AT 95% LEVEL	1	1	1	1	1	0	0	1	1	2

Superscripts on entries in table are number of times the coefficients are negative and significantly different from zero.

TABLE 16. UNCONFINED FLOW. INPUT VARIABLES CONSIDERED SIMULTANEOUSLY. 10 TIME STEPS. NATURE OF ESTIMATES OF REGRESSION COEFFICIENTS IN 'COEFFICIENTS-OF-VARIATION' MODEL.

	TIME CHANGE OF ESTIMATES OF REGRESSION COEFFICIENTS			NO. OF ESTIMATES WHICH WERE ENTIRELY POSITIVE OR NEGATIVE WITH TIME		
	INCREASE	DECREASE	BOTH	POSITIVE	NEGATIVE	BOTH
a_0	16	0	4	10	0	10
a_1	4	8	8	8	4	8
a_2	4	8	8	5	8	7
a_3	7	2	11	10	4	6
a_4	0	20	0	20	0	0
a_5	7	5	8	3	7	10

TABLE 17. UNCONFINED FLOW. INPUT VARIABLES CONSIDERED SIMULTANEOUSLY. 10 TIME STEPS. TIME CHANGE OF SUM OF SQUARES OF DEVIATIONS.

	DECREASE	MINIMUM	2 TURNING POINTS	3 TURNING POINTS
'TOLERANCE-INTERVAL-WIDTH' REGRESSION MODEL	12	4	1	3
'COEFFICIENT-OF-VARIATION' REGRESSION MODEL	13	7		

TABLE 18. UNCONFINED FLOW. INPUT VARIABLES CONSIDERED SIMULTANEOUSLY. 10 TIME STEPS. TYPICAL ESTIMATES OF REGRESSION COEFFICIENTS IN 'COEFFICIENTS-OF-VARIATION' MODEL.

TIME STEP \ COEFF.	α_0	α_1	α_2	α_3	α_4	α_5
1	-.001229	.007663	-.008883	.002526	.922341	.017935
2	-.000920	.007811	-.006873	.002016	.788755	.016479
3	-.000319	.008345	-.005124	.001167	.671388	.013862
4	.000421	.008984	-.003835	.000131	.569848	.010888
5	.001386	.009615	-.002816	-.001228	.466938	.007113
6	.002309	.010023	-.002188	-.002438	.383043	.003723
7	.003138	.010189	-.001767	-.003377	.315063	.000897
8	.003858	.010146	-.001470	-.004039	.260380	-.001425
9	.004464	.009960	-.001264	-.004447	.216661	-.003295
10	.004961	.009684	-.001118	-.004615	.181704	-.004676

head, a regression model involving only errors in initial head was investigated.

For one grid this can be written,

$$E(t_0'') = \alpha_0 + \alpha_1(\tilde{\rho}^*_{h_I})_0 + \alpha_2(\tilde{\rho}^*_{h_I})_1 + \alpha_3(\tilde{\rho}^*_{h_I})_2 + \alpha_4(\tilde{\rho}^*_{h_I})_3 + \alpha_5(\tilde{\rho}^*_{h_I})_4 ,$$

where the subscripts refer to a pattern of grids as in Figure 2. A similar model relating the estimates of the coefficients of variation was also studied. The data used in studying this model was the same as that which has just been described for the 10 time step run. The results are given in Tables 19, 20, 21, where, α^* has the same definition as given for the confined flow model.

Table 19 gives the number of times that the hypotheses $H_0: [(\alpha_i = 0),\ i=0,\ 1,\ \alpha_i \subset \alpha^*]$, $H_0: [(\alpha_0, \alpha^*) = 0]$ and $H_0: (\alpha^* = 0)$, were rejected at the 95% level. Table 20 indicates whether, in the "coefficients-of-variation" model, the estimates of the regression coefficients tended to increase or decrease with time and whether they tended to be positive or negative throughout the study period. Table 21 indicates the way in which the sum of squares of deviations from the regression line changed with time.

The following observations can be made from the results given in Tables 19, 20, 21:

1) from Table 19, the error in initial head in the grid being considered is of significance in predicting the error in

final head up to 440 days, but this significance decreases after about 400 days, and up to 140 days the predictive relation is linear. Initially, it was also approximately a one-to-one relation,

2) from Table 20, α_0 shows a tendency to increase with time, but this is not as marked as in the previous 10 time step model (see Table 16), and it is non-significant up to about 300 days, after which, from Table 19, it cannot be considered to be non-significant from zero. Again, this is not as marked as in the previous 10 time step model (see Table 15),

3) from Table 20, α^* shows a tendency to be positive and to increase with time, and from Table 19, it is non-significant, both individually and simultaneously, up to about 300-350 days, after which it cannot be considered to be non-significant from zero,

4) from Table 20, α_1 decreases monotonically with time and is positive,

5) from Table 21, the sum of squares of deviations from the regression line shows a tendency to decrease with time.

A comparison of the sum of squares of deviations in the above model involving only errors in initial head and the previous 10 time

TABLE 19a. UNCONFINED FLOW. RESULTS FOR REGRESSION MODEL INVOLVING ERRORS ON INITIAL HEAD.

	'TOLERANCE-INTERVAL-WIDTH' REGRESSION MODEL									
TIME STEP / COEFF.	1	2	3	4	5	6	7	8	9	10
α_0	1	1	2^{1}	2^{1}	3^{1}	2	2	4	4	4
α_1	20	20	20	20	20	20	20	20	18	16
α^*	4^{4}	1^{1}	2^{1}	3	3	5^{1}	6^{1}	4	5	6
NUMBER OF GRIDS IN WHICH REJECT $H_o: [\alpha_o, \alpha^*] = 0$ AT 95% LEVEL	1	1	0	1	5	8	10	15	16	17
NUMBER OF GRIDS IN WHICH REJECT $H_o: [\alpha^*] = 0$ AT 95% LEVEL	1	1	1	1	1	2	3	5	5	5

α^* consists of all the coefficients associated with errors in initial head in neighboring grids.

The entries in the table opposite α_0 and α_1 are the number of grids (out of 20) in which the hypothesis $H_o: [\alpha_i = 0, i = 0, 1]$ was rejected.

The entries in the table opposite α^* are the number of times (out of 62) that the hypothesis $H_o: [\alpha_i = 0, \alpha_i \subset \alpha^*]$ was rejected.

Superscripts on entries in the table are the number of times the regression coefficient was negative and significantly different from zero.

TABLE 19b. UNCONFINED FLOW. RESULTS FOR REGRESSION MODEL INVOLVING ERRORS ON INITIAL HEAD

	'COEFFICIENT-OF-VARIATION' REGRESSION MODEL									
TIME STEP / COEFF.	1	2	3	4	5	6	7	8	9	10
a_0	1	1	1	1	2	3	3	6	6	8
a_1	20	20	20	20	20	20	20	20	20	18
a^*	3^{3}	2^{2}	2^{1}	4	4	4	6	9	10	11
NUMBER OF GRIDS IN WHICH REJECT $H_o: [a_0, a^*] = 0$ AT 95% LEVEL	1	0	1	4	9	10	15	15	17	17
NUMBER OF GRIDS IN WHICH REJECT $H_o: [a^*] = 0$ AT 95% LEVEL	1	1	1	1	2	3	3	3	3	4

a^* consists of all the coefficients associated with errors in initial head in neighboring grids.

The entries in the table opposite a_0 and a_1 are the number of grids (out of 20) in which the hypothesis $H_o: [a_i = 0, i = 0_1, 1]$ was rejected.

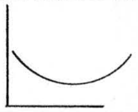
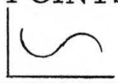
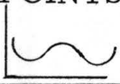
The entries in the table opposite a^* are the number of times (out of 62) that the hypothesis $H_o: [a_i = 0, a_i \subset a^*]$ was rejected.

Superscripts on entries in the table are the number of times the regression coefficient was negative and significantly different from zero.

TABLE 20. UNCONFINED FLOW. NATURE OF ESTIMATES OF REGRESSION COEFFICIENTS IN THE 'COEFFICIENTS-OF-VARIATION' MODEL INVOLVING ERRORS ON INITIAL HEAD.

	TIME CHANGE OF ESTIMATES OF REGRESSION COEFFICIENT			NO. OF ESTIMATES WHICH WERE ENTIRELY POSITIVE OR NEGATIVE WITH TIME		
	INCREASE	DECREASE	BOTH	POSITIVE	NEGATIVE	BOTH
a_0	11	2	7	13	2	5
a_1	0	20	0	19	0	1
a^*	28	7	27	21	9	32

TABLE 21. UNCONFINED FLOW. TIME CHANGE OF SUM OF SQUARES OF DEVIATIONS IN REGRESSION MODEL INVOLVING ERRORS ON INITIAL HEAD.

	DECREASE 	MINIMUM 	2 TURNING POINTS 	3 TURNING POINTS 
'TOLERANCE-INTERVAL-WIDTH' REGRESSION MODEL	12	6	1	1
'COEFFICIENT-OF-VARIATION' REGRESSION MODEL	12	8		

step model shows that for small periods of time the sum of squares is less in the 10 time step model than the initial heads model in about 75% of the grids but after 440 days this is reduced to about 50%. However, it is noteworthy that after 440 days all of the grids (7, 8, 9, 12, 13, 14) which have four neighboring grids have a smaller sum of squares in the model involving only errors in initial head than in the 10 time step model.

DISCUSSION OF RESULTS

In using a numerical model to study confined groundwater aquifers the results of this study clearly indicate that when there is an error in only one input variable, that this error is linearly related to the error in output. For error in initial head, this relation is approximately one-to-one for short time periods, but for the other input variables the error on output is one or two orders of magnitude less than the input error. These relations have been demonstrated only for comparatively short periods of time. How they change over long time periods and whether they hold for unconfined flow have not been investigated, but in view of the similarity of the results of the error relation for both confined and unconfined cases when the input variables were considered simultaneously, it is thought that for short periods of time similar linear relations would hold for unconfined flow. However, extrapolation of these results to longer time periods is considered inappropriate without further study. Such study was not undertaken because the situation in which all the input variables contain errors is far more typical.

When the input variables to the numerical model are all considered to contain error there is a marked similarity and consistency between the results obtained for confined and unconfined flow. Outstanding amongst these is that the error in initial head is the only significant input error so far as predicting the error in final head is concerned. This result is compatable with the results obtained when only one variable contained error. For small periods of time the predictive relation between the error in initial head and error in final head, in any one grid, is linear and initially it is approximately one-to-one. The dependence of the error in final head on the error in initial head, in any one grid, has been shown to decrease with time, while, concurrently, the value of the constant coefficient in the regression equation is increasing and the errors in initial head in neighboring grids tend to become significant. The errors on the input variables K, b (or z), Q and S are always non-significant at the 95% level in predicting the error in final head. Another consistent feature of the results is that the sum of squares of deviations from the regression line has a marked tendency to decrease with time.

The results do, however, show minor variations, both between the confined and unconfined cases and between the different regression models considered, that are worthy of comment.

For unconfined flow the significance of the errors in initial head in neighboring grids is more marked, though this is possibly due to the fact that results were obtained after a longer period of time in the unconfined case, and it is then that these errors become significant.

For unconfined flow, the constant regression coefficient showed a greater tendency to increase with time and become significant at an earlier time in the regression model consisting of all five input variables than the one consisting only of initial heads. This would seem to be due to the greater significance that errors in initial head have in predicting errors in final head. The fact that this difference was not apparent for confined flow is probably due, again, to the comparatively short length of study period.

There is also an indication that the sum of squares of deviations from the regression line ceases to decrease after some time, but results for longer study periods would be needed to determine whether it begins to increase or becomes asymptotic.

An implication of these results, which is an apparent contradiction, is that after a large time period the best predictive equation for the error on final head will be $E(t'') = \text{constant}$, and that any regression coefficients associated with errors in input variables will be non-significant. This implication is suggested by,

(i) error on initial head becomes less important with time (even though it may remain significant in some of the grids),

(ii) errors in K, b (or z), Q, S are never significant,

(iii) the constant coefficient increases in significance with time.

The implication would appear to be a contradiction because if all errors in input variables are zero then the error in final head should be zero for all time. However, the contradiction is only apparent for the following reasons.

After a large period of time the error on the final head in any one grid will be influenced by the errors on all the input variables in all the grids of the model, but the contribution that each one of these input errors makes to the final error will be small. Therefore, the final error becomes a function of a large number of variables each one of which contributes only a small amount. If this function is assumed to be a linear (in the coefficients) function of all the input errors, then it can be viewed as a "flat" surface in n-dimensional space, since all the coefficients will be small and represent gradients. This "flat" surface could be closely approximated by a constant over a large range of values of the input variables. Thus, if a regression equation is fitted to data obtained after a large time only the constant coefficient would appear to be significant. This does not imply that the error on final head is independent of the input errors, only that after a large period of time the contribution of each of these input errors is so small as to be statistically non-significant. Nor does it preclude the case of zero error on final head when input errors are zero.

It is reasonable to assume that the value which the constant takes on is determined far more by the errors in initial head than by the errors on other input variables, even though all of them are statistically non-significant.

The conclusions drawn from this study are strictly only applicable to the models (physical and regression), data and time periods considered. However, the general nature of the physical models used and the consistency and acceptability of the results obtained from the regression models which have been considered suggests that these two criteria are not too restrictive in making general use of the results. So far as the data is concerned, it should be noted that the errors on K, b(or z), Q, S have been shown to be non-significant compared to the error on h_I only when the errors are measured with respect to the true mean values of the variables in a grid and only for the range of errors which has been considered. Thus, gross errors on any one of the input variables or extreme differences between the estimated and true mean values of the input variables would probably significantly influence the estimates of final head and its accuracy. For instance, if discontinuities occur in the spatial change of the true values of one of the input variables, as at fault zones or with abrupt lithological changes, and go undetected, then the results obtained from the application of the procedures of this study would be erroneous. It is therefore incorrect to conclude that any values of K, b (or z), Q, S can be used as input to the model without having any effect on the output, whereas, it is correct to conclude that any values of these variables which are within these ranges of error of the true mean values would not significantly change the prediction of final head and the accuracy of the prediction. Also, the data used in this study is typical of Colorado aquifers and the applicability of the results to different types of data has not been demonstrated. The fact that, in the "coefficients-of-variation" models, dimensionless quantities have been considered, indicates that the results should be applicable to many situations. However, the results may not be applicable where the data which greatly differs from that used in this study. The extrapolation to longer time periods is more hazardous and should only be done with great care.

APPLICATION OF RESULTS

The experimental procedures used in this study and the results obtained have a number of applications and implications in ground water hydrology. Some of these will be indicated and elaborated upon in the following discussion. They fall generally in the areas of data collection, economics of aquifer management, and the relations between accuracy of results and number of field observations available.

Perhaps the most obvious application to which the results and techniques of this study can be applied is to the case in which data is available on the input variables in an aquifer and it is required to predict the water table elevation at some future time and estimate the accuracy of the prediction. This could be done as follows:

1) using equations (4) and (5) compute $\overset{\simeq}{K}$, $\overset{\simeq}{b}$ (or $\overset{\simeq}{z}$), $\overset{\simeq}{Q}$, $\overset{\simeq}{S}$, $\overset{\simeq}{h}_I$ and $\tilde{\rho}_i$, i=1.... 5 in all grids, from available field data,

2) using the Monte Carlo technique, as applied in this study, generate a random sample of h_F for the specified future time,

3) in every grid compute a tolerance interval on h_F. This tolerance interval can be used as the prediction of the final head and the width of the interval as a measure of

the accuracy of the prediction. This procedure does not take into account the influence that the differences between the true mean value of a variable in a grid and the estimate of this value and between $\tilde{\rho}_i$ and ρ_i have on the predicted value of final head and the estimate of the accuracy of the prediction. Using permeability as an example, it is conceivable that $|\approx{K}-\bar{K}|$ could be large and that $\tilde{\rho}_K << \rho_K$. This would result in the prediction of the final head being incorrect, since random values would be generated from $N(\approx{K}, \tilde{\rho}_K^2)$ instead of $N(\bar{K}, \rho_K^2)$ and in assigning a greater accuracy to this prediction than is justified. However, the probability that $|\approx{K}-\bar{K}|$ is large is very small and decreases with increasing sample size, and it is more probable that $\tilde{\rho}_K$ will be larger than ρ_K than smaller. Also, the errors on the variables, K, b(or z), Q and S have been shown to be non-significant, so that the consequences of the above assumption are further diminished.

Having obtained the tolerance interval in every grid from available data it might then be desirable to determine how much more field data, and on which variables, would be required to improve the accuracy of the prediction (that is, decrease the width of the tolerance interval) by a specified amount. This can be accomplished by making use of the fact that the results of this study indicate that, after any period of time, only the errors in initial head, of all the input errors, are significant in predicting the error in final head. Therefore, to reduce the tolerance interval width it is necessary to reduce the error in initial head. This study has established the validity of this predictive relation but has not investigated in detail the time dependence of the coefficients in the predictive equation. Thus, for a particular grid of any aquifer after an arbitrary time, it is not possible to simply specify the desired value of tolerance interval width and then solve the prediction equation for the necessary values of errors on initial heads that will achieve this specified value. For this reason a simple trial and error procedure will be described to arrive at these specified values of errors in initial heads. For small periods of time only the error in initial head in the grid under consideration is of significance in predicting the error in final head, that is;

$$E(t'') = \alpha_0 + \alpha_1 \tilde{\rho}_{h_I} \quad .$$

Thus, to obtain a decrease in the tolerance interval width from $E(t''_a)$ to $E(t''_b)$, where t_a is the computed width corresponding to $(\tilde{\rho}_{h_I})_a$, estimated from the available field data, and t_b is the desired width, it is necessary only to find the value of $\tilde{\rho}_{h_I}$, say $(\tilde{\rho}_{h_I})_b$, that will give the width $E(t''_b)$. This value is easily determined by trial and error. Now, from equation (5),

$$(\tilde{\rho}_{h_I})_a = \frac{\hat{\sigma}_{h_I}}{\sqrt{n_a}} f(X_{ji}) \qquad ,$$

and the term $[\,\hat{\sigma}_{h_I} f(X_{ji})\,]$ is approximately the same for any value of n and is known. Therefore, the total number of observations needed to obtain an error $(\tilde{\rho}_{h_I})_b$ is n_b, where,

$$n_b = \left[\frac{\sqrt{n_a}\,(\tilde{\rho}_{h_I})_a}{(\rho_{h_I})_b} \right]^2 \qquad ,$$

and the extra number of observations needed is $(n_b - n_a)$.

For larger periods of time the error in initial head in grids neighboring the one under consideration become significant. Therefore, the procedure described above for small time periods would have to be modified to include these neighboring grids but is otherwide directly applicable.

This determination of the number of additional observations needed to obtain an improved accuracy on the predicted value of head immediately raises the economic question of whether or not the value to be gained from the increase in accuracy is "worth" the cost of obtaining the extra data. This question and similar ones which occur in aquifer management can be approached using the results of this study.

A further implication of the results of this study is in the design and operation of observational data networks. It is clear that greatest emphasis should be given to obtaining water table elevation data and considerably less emphasis to the determination of other aquifer parameters.

The procedures and techniques developed in this study are basically to be used for predictive purposes. However, where historical data is available, as on water table elevations, the combined use of "matching" techniques and this predictive model is possible.

CONCLUSIONS

The following conclusions can be made, subject to the restrictions described in the section on discussion of results.

For confined flow, when only one input variable is considered to contain error, the following conclusions can be made for short time periods:

i) errors on input variables and final head are linearly related,

ii) for initial head this relation is initially one-to-one,

iii) for K, b, S and Q the error on output is one to two orders of magnitude less than the input error.

For both confined and unconfined flow, when the errors on input are considered simultaneously, the following conclusions can be made:

i) the error in initial head is the only significant error in predicting error on final head in any one grid,

ii) for short periods of time these errors on input and output are linearly related and initially are one-to-one,

iii) the importance of the error in initial head in predicting error in final head decreases with time and the errors in initial head in neighboring grids become significant,

iv) after long time periods the error on final head approaches a constant value, where the constant is considered to be a function of all the input errors.

RECOMMENDATIONS

One of the results of this study is that after large periods of time the predicted value of accuracy of final head in any one grid is equal to a constant. The time-dependence of this constant value is worthy of further detailed investigation. Also, the applicability of the results of this study to groundwater systems in which extreme errors occur in the field data or where large and abrupt spatial changes in the values of the input variables are present should be further studied.

The results of this study provide a basis for an economic analysis of ground water systems dealing with the general problem of benefit and cost of data collection.

BIBLIOGRAPHY

1. BIBBY, R., "Flow between the confined aquifer of the Fox Hills Sandstone and the alluvial aquifer, in the North Kiowa-Bijou District, Colorado", M. S. Thesis, Department of Civil Engineering, Colorado State University, Fort Collins, Colorado, March, 1969.

2. BIBBY, R., and D. K. SUNADA, "Mathematical model of leaky aquifer", submitted to ASCE.

3. BITTINGER, M. W., "Simulation and analysis of stream-aquifer systems", Ph.D. Dissertation, Department of Civil Engineering, Utah State Univ., Logan, Utah, 1967.

4. BITTINGER, M. W., H. R. DUKE, R. A. LONGENBAUGH, "Mathematical simulations for better aquifer management", Publication No. 72, International Association of Scientific Hydrology, Symposium of Haifa, pps. 509-519, March, 1967.

5. ESHETT, A., and R. A. LONGENBAUGH, "Mathematical model for transient flow in porous media", Report CER65RAL-AE59, Department of Civil Engineering, Colorado State University, Fort Collins, Colorado, November, 1965.

6. ESHETT, A., "Ground water system analysis", Ph.D. dissertation, Department of Civil Engineering, Colorado State University, Fort Collins, Colorado, August, 1970.

7. GRAYBILL, F. A., "An Introduction to Linear Statistical Models, Vol. 1", McGraw-Hill, 1961.

8. JACOB, C. E., "Flow of ground water", in "Engineering Hydraulics" Ed H. Rouse, J. Wiley, 1964.

9. JOHNSTON, J., "Econometric Methods", McGraw-Hill, 1963.

10. LONGENBAUGH, R. A., "Statistical techniques for predicting river accretions as applied to the South Platte River", M. S. Thesis, Department of Agricultural Engineering, Colorado State University, Fort Collins, Colorado, July, 1962.

11. LONGENBAUGH, R. A. and M. W. BITTINGER, "Ground and surface water relationships studied by statistical techniques", presented at 1963 winter meeting of Amer. Soc. Ag. Engin., Chicago, Ill., December 10-13.

12. LONGENBAUGH, R. A., "Mathematical simulation of a stream-aquifer system", Report CER67-68RAL22, Department of Civil Engineering, Colorado State University, Fort Collins, Colorado, November, 1967.

13. McMILLAN, W. D., "Theoretical analysis of groundwater basin operations", Contribution No. 114, Water Resources Center, Hydraulic Laboratory, University of California, Berkeley, November, 1966.

14. MOOD, A. M. and F. A. GRAYBILL, "Introduction to the theory of statistics", McGraw-Hill, 1963.

15. RORABAUGH, M. I., "Prediction of ground water levels on basis of rainfall and temperature correlations", Trans. American Geophysical Union, V. 37, No. 4, pp. 436-441, August, 1956.

16. SHCHIGOLEV, B. M., "Mathematical analysis of observations", American Elsevier.

17. WOODS, P. C., "Management of hydrologic systems for water quality control", Contribution No. 121, Water Resources Center, University of California, Berkeley, June, 1967.

APPENDIX 1

I. CONFINED FLOW

A. EACH INPUT VARIABLE CONSIDERED SINGLY

The relation between the estimates of standard deviation of each of the variables h_I, K, S, b, Q and the tolerance width on h_F were determined for the following models:

MODEL (i)

The confined aquifer was divided into 20 square grids, with boundaries as shown in Figure A. 1. 1. Values were assigned to all of the input variables as shown in Table A. 1. 1.

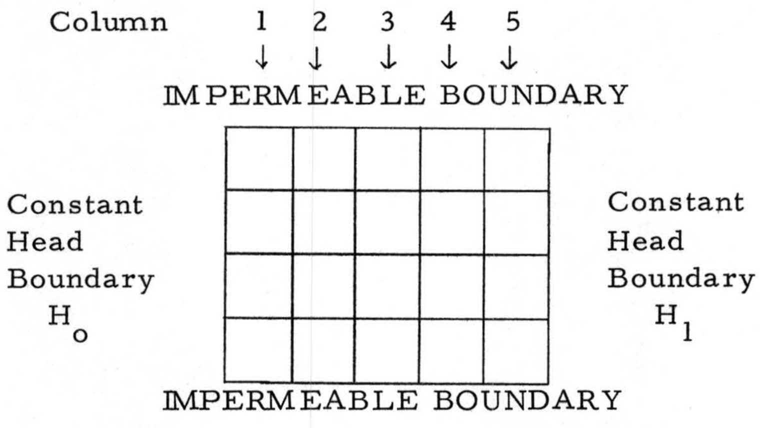


Fig. A. 1. 1. PLAN VIEW OF MODEL (i)

MODEL (ii)

The model and model boundary conditions were the same as for Model (i), but the values of the input variables were randomly generated. Each variable, except the one for which the error relation was being determined, was assigned a maximum, median and minimum value and in each grid a random value of the variable was generated from a triangular distribution based on these three numbers.

MODEL (iii)

This was the same in all respects as Model (ii) except that the mean value of the variable for which the error relation was being determined was, in each grid, generated from a triangular distribution. The standard deviations of K, S, b, Q were computed from these randomly generated mean values.

MODEL (iv)

This model is the same as Model (iii), except that results were obtained at the end of 3 time steps.

MODEL (v)

In Models (i) - (iv) the constant head boundaries H_o, H_1, were equal throughout the run. In this model they are assigned different values. Also in the first four models the values of initial head in each grid had been generated from the same triangular distribution. For this model the triangular distributions were different between columns in the model. This difference between the columns was made compatable with the values assigned to H_o and H_1.

MODEL (vi)

For this model, the 20 grid model was assumed to have constant gradient boundaries (Figure A. 1. 2). At each boundary grid the constant gradient boundary was randomly generated from a triangular distribution.

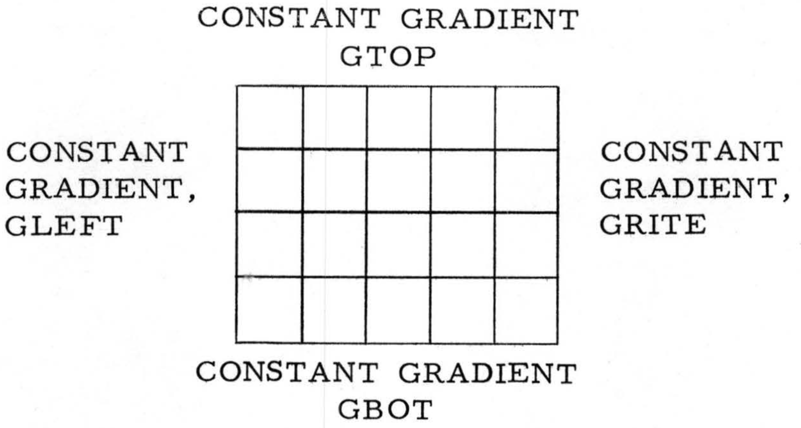


Fig. A. 1. 2. PLAN VIEW OF MODEL (vi)

The objective in studying all the above models was to take into consideration all of the types of boundary conditions which are normally met in ground water systems (but not of course all possible combinations) and also to progress to the stage at which all the data used in the models was being randomly generated. Any results obtained for the latter condition can more reasonably be expected to be generally valid.

The actual data used in the models is given in TABLE A. 1. 1.

		INPUT VARIABLE CONTAINING ERROR:				
		K	h_I	S	b	Q
MODEL (i)	H_o, H_1(FT)	100	200		200	200
	h_I(FT)	100	---		200	200
	b(FT)	50	75		---	75
	Q(FT^3/DAY)	5000	5.10^5		5.10^5	---
	S	.2	.2		.2	.2
	K(FT/DAY)	---	100	NO RUN	100	100
	μ	100	200		75	5.10^5
	ρ	5(5)35*	2(2)16		.625(.625)6.25	$12.10^3(12.10^3)$ 12.10^4
	DX, DY (FT)	1000	10000		10000	10000
	DT(DAY)	50	40		10	10
	NT	1	1		1	1

μ = Mean value of variable containing error
ρ = Standard deviation of variable
DX = DY = Grid dimensions
DT = Time step size

NT = Number of time steps
*5(5)35 = 5, 10, 15, 20, 25, 30, 35
**Maximum, median and minimum values defining triangular distribution.

TABLE A.1.1a DATA USED IN MODELS (i) - (vi)

		INPUT VARIABLE CONTAINING ERROR:				
		K	h_I	S	b	Q
MODEL (ii)	H_o, H_1 (FT)	100	200	200	200	200
	h_I (FT)	55-100-145**	---	100-200-300	100-200-300	100-200-300
	b(FT)	45-50-55	70-75-80	55-75-95	---	55-75-95
	Q(FT^3/DAY)	.1-.2-.3	$0-5.10^5-1.10^6$	$0-5.10^5-1.10^6$	$0-5.10^5-1.10^6$	---
	S	0-5000-10000	.1-.2-.3	---	.1-.2-.3	.1-.2-.3
	K(FT/DAY)	---	1-100-199	1-100-199	1-100-199	1-100-199
	μ	100	200	.2	75	5.10^5
	ρ	3(3)27	2(2)16	.004(.004).04	.625(.625)6.25	$12.10^3(12.10^3)$ 12.10^4
	DX, DY (FT)	1000	10000	10000	10000	10000
	DT(DAY)	50	40	10	10	10
	NT	1	1	1	1	1

TABLE A.1.1.b (cont.)

		INPUT VARIABLE CONTAINING ERROR:				
		K	h_I	S	b	Q
MODEL (iii)	H_o, H_I(FT)					
	h_I(FT)					
	b(FT)	DATA AS MODEL (ii)	DATA AS MODEL (ii)	DATA AS MODEL (ii)	DATA AS MODEL (ii)	DATA AS MODEL (ii)
	Q(FT^3/DAY)					
	S					
	K(FT/DAY)					
	μ	1-100-199	175-200-225	.1-.2-.3	55-75-95	1.10^5-5.10^5-9.10^5
	ρ_{MIN}	$(\mu/3.7)/8.0$	------	$(\mu/5.0)/10.0$	$(100-\mu/4)/10$	$(\mu/5)/10$
	ρ	$\rho_{MIN}(\rho_{MIN})8\rho_{MIN}$	2(2)16	$\rho_{MIN}(\rho_{MIN})10\rho_{MIN}$	$\rho_{MIN}(\rho_{MIN})10\rho_{MIN}$	$\rho_{MIN}(\rho_{MIN})10\rho_{MIN}$
	DX, DY (FT)	1000	10000	10000	10000	10000
	DT(DAY)	1	20	10	10	10
	NT	1	1	1	1	1

TABLE A. 1. 1. c (cont.)

		INPUT VARIABLE CONTAINING ERROR:				
		K	h_I	S	b	Q
MODEL (iv)	H_o, H_1(FT)					
	h_I(FT)					
	b(FT)					
	Q(FT^3/DAY)					
	S					
	K(FT/DAY)	DATA AS	DATA AS	DATA AS	DATA AS	DATA AS
	μ	MODEL (iii)	MODEL (i)	MODEL (iii)	MODEL (iii)	MODEL (iii)
	ρ_{MIN}					
	ρ					
	DX, DY (FT)					
	DT(DAY)					
	NT	3	3	3	3	3

TABLE A. 1. 1. d (cont.)

		INPUT VARIABLE CONTAINING ERROR:				
		K	h_I	S	b	Q
MODEL (v)	H_o (FT)	130	230	260	260	260
	H_1 (FT)	70	170	200	200	200
	h_I (FT)	HMEAN-25 HMEAN HMEAN+25	-----------	HMEAN-100 HMEAN HMEAN+100	HMEAN-100 HMEAN HMEAN+100	HMEAN-100 HMEAN HMEAN+100
	HMEAN (FT)	120(10)80	220(10)180	250(10)210	250(10)210	250(10)210
	b (FT)					
	Q(FT^3/DAY)		DATA AS			
	S		MODEL (iii)			
	K(FT/DAY)					
	μ	DATA AS	HMEAN-25 HMEAN HMEAN+25	DATA AS	DATA AS	DATA AS
	ρ_{MIN}	MODEL (iii)		MODEL (iii)	MODEL (iii)	MODEL (iii)
	ρ					
	DX, DY (FT)		DATA AS			
	DT(DAY)		MODEL (iii)			
	NT					

TABLE A.1.1.e (cont.)

		INPUT VARIABLE CONTAINING ERROR:				
		K	h_I	S	b	Q
MODEL (vi)	HMEAN(FT)					
	h_I(FT)					
	b(FT)					
	Q(FT^3/DAY)					
	S					
	K(FT/DAY)					
	μ					
	ρ_{MIN}	DATA AS	DATA AS	DATA AS	DATA AS	DATA AS
	ρ	MODEL (iii)	MODEL (iii)	MODEL (v)	MODEL (v)	MODEL (v)
	DX, DY(FT)					
	DT(DAY)					
	NT					
	GTOP(FT)	0-5-10	0-25-50	0-25-50	0-25-50	0-25-50
	GBOT(FT)	-10--5-0	-50--25-0	-50--25-0	-50--25-0	-50--25-0
	GRITE(FT)	-20--10-0	-100--50-0	-100--50-0	-100--50-0	-100--50-0
	GLEFT(FT)	0-10-20	0-50-100	0-50-100	0-50-100	0-50-100

TABLE A. 1. 1. f (cont.)

B. INPUT VARIABLES CONSIDERED SIMULTANEOUSLY

The model and data used to study the prediction equation, $E(t') = \alpha_0 + \alpha_1\tilde{\rho}^*_K + \alpha_2\tilde{\rho}^*_b + \alpha_3\tilde{\rho}^*_Q + \alpha_4\tilde{\rho}^*_{h_I} + \alpha_5\tilde{\rho}^*_S$, are given in Figure A. 1. 3 and Table A. 1. 2.

The standard deviations of the input variables were randomly generated from uniform distributions and the maximum and minimum values assigned to these distributions are given in Table A. 1. 2.

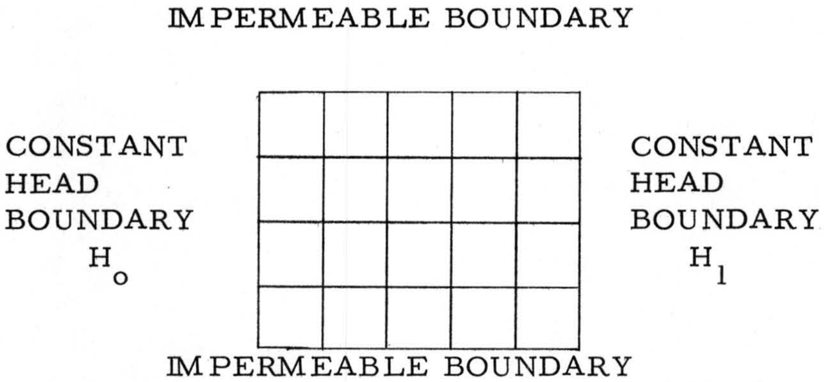


FIGURE A. 1. 3. PLAN VIEW MODEL USED TO STUDY FIRST REGRESSION MODEL.

The model and data used to study the second regression model consisting of the first and second terms of the Taylor Expansion are given in Figure A. 1. 4 and Table A. 1. 3.

The mean values of each of the input variables in each grid were randomly generated from triangular distributions and the standard deviations from uniform distributions. The constant gradients in each of the boundary grids were also generated from triangular distributions.

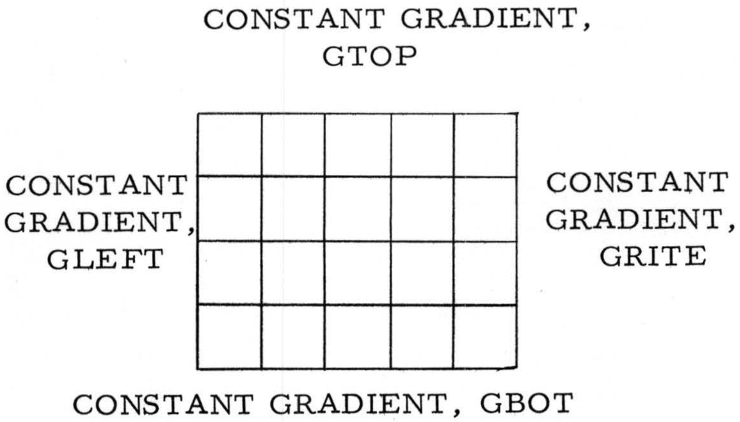


FIGURE A. 1. 4. PLAN VIEW OF MODEL USED TO STUDY SECOND REGRESSION MODEL.

H_o, H_1(FT)		200	S	μ	.2
h_I	μ	200		ρ	.004 .04
(FT)	ρ	1* 10	K (FT/DAY)	μ	100
b	μ	75		ρ	1 20
(FT)	ρ	1 10	DX, DY(FT)		10000
Q(FT^3/	μ	5.10^5	NT		4
DAY)	ρ	0 1.10^5	DT(DAY)		10, 20, 40, 40

*Maximum and Minimum values defining uniform distribution.

TABLE A. 1. 2. DATA USED TO STUDY FIRST REGRESSION MODEL

h_I	μ	175* 200 225	K	μ	50 100 150
(FT)	ρ	1 $\lvert \mu_{h_I} - 140 \rvert / 5$	(FT/ DAY)	ρ	1 $\mu_K/5$
b	μ	60 75 90	DX, DY(FT)		10000
(FT)	ρ	1** 8	NT		1
Q	μ	0 5.10^5 1.10^6	DT(DAY)		20
(FT^3/ DAY)	ρ	0 $\mu_Q/5$	GTOP(FT)		-80 0 80
S	μ	.1 .2 .3	GBOT(FT)		-80 0 80
	ρ	.004 .04	GRITE(FT)		-80 0 80
			GLEFT(FT)		-80 0 80

**Maximum and minimum values defining uniform distribution.

*Maximum, median and minimum values defining triangular distribution.

TABLE A. 1. 3. DATA USED TO STUDY SECOND REGRESSION MODEL

II. UNCONFINED FLOW

A. INPUT VARIABLES CONSIDERED SIMULTANEOUSLY

The model and data used to study the predictive equation, $E(t'') = \alpha_0 + \alpha_1\tilde{\rho}_K + \alpha_2\tilde{\rho}_Z + \alpha_3\tilde{\rho}_Q + \alpha_4\tilde{\rho}_{h_I} + \alpha_5\tilde{\rho}_S$, for unconfined flow is given in Figure A. 1. 5 and Table A. 1. 4.

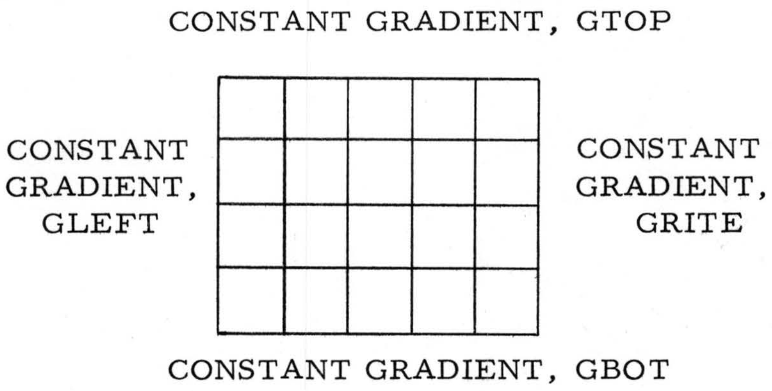


FIGURE A. 1. 5.

The mean values of each of the input variables in each grid were randomly generated from triangular distributions and the standard deviations from uniform distributions. The constant gradient in each boundary grid were also generated from triangular distributions. The numbers given in Table A. 1. 4 defined these distributions, and remained the same for the 1, 6 and 10 time-step runs. However, the random samples generated from the distributions were different for each run.

Also, the value of the constant gradient was allowed to change with each time step. The value of the constant gradient for any time step other than the first was randomly generated from a triangular distribution whose median value was the value of the gradient in the preceding time step and whose maximum and minimum values were the median value plus five and minus five respectively.

h_I (FT)	μ	275	300	325
	ρ	1	10	
Z (FT)	μ	50	75	100
	ρ	1	10	
Q (FT^3/DAY)	μ	0	5.10^5	1.10^6
	ρ	0	$\mu_Q/5$	
S	μ	.1	.2	.3
	ρ	.004	.02	
K (FT/DAY)	μ	50	100	150
	ρ	1	$\mu_K/5$	
DX, DY (FT)		10000		
NT		1, 6, 10		
DT(DAYS)		20, 40, 40, 40, 50, 50, 50, 50, 50, 50		
GTOP (FT)		-80	0	80
GBOT (FT)		-80	0	80
GRITE (FT)		-80	0	80
GLEFT (FT)		-80	0	80

TABLE A. 1.4.

APPENDIX 2

GENERAL LINEAR HYPOTHESIS MODEL OF FULL RANK, MODEL I, CASE A, AND REGRESSION MODEL, MODEL III, CASE 2.

The description of these two linear models is taken from Graybill [7].

GENERAL LINEAR HYPOTHESIS MODEL OF FULL RANK, MODEL I, CASE A.

Consider a random variable y which has a density function,

$$f(y : x_1 \ldots\ldots x_p, \beta_1 \ldots\ldots \beta_p) \quad ,$$

where, x_i = known, non-random variables,

β_i = unknown parameters.

Assume that,

(i) $$E(y) = \sum_{i=1}^{p} \beta_i x_i$$

(ii) $$\mathrm{Var}\,(y) = \sigma^2$$

(iii) σ^2 is independent of β_i and x_i .

If a random sample, y_j, $j=1\ldots\ldots n$, is taken from this density, that is

$$y_j = \sum_{i=1}^{p} \beta_i x_{ji} + e_j \quad , \qquad j=1 \;.\,.\,.\,.\,.\; n \;, \qquad (A.2.1)$$

where,

e_j = random error,

and if the random errors, e_j, j=1 ..... n, are uncorrelated, then

$$E(e_j) = 0$$

$$Var\,(e_j) = \sigma^2 \quad .$$

Equation (A. 2. 1) can be written in matrix notation as,

$$\underline{Y} = \underline{X}\,\underline{\beta} + \underline{e} \quad .$$

If the random sample is taken in such a way that the x_{ji} are specified (either randomly or by design) and then an observation, y_j, is made and if the rank of $\underline{X}$ is p ($p \leq n$), then, this is the General Linear Hypothesis Model of Full Rank. Case A of this model is when $\underline{e} \sim N\,(O, \; \sigma^2 \underline{I})$.

REGRESSION MODEL, MODEL III, CASE 2.

Consider the random variables y, x_1 ..... x_p which have the density function,

$$f(y\,,\, x_1 \,.\,.\,.\,.\,.\, x_p) = h(x_1 \,.\,.\,.\,.\,.\, x_p)\,\frac{1}{\sqrt{2\pi}\,\sigma} \quad \exp\left\{-\frac{1}{2}\left(\frac{y - G(x_1 \ldots x_p)}{\sigma}\right)^2\right\}$$

where

$G(x_1 \ldots x_p)$ is a linear (in the coefficients β_i)

function of x_i,

$h(x_1 \ldots . x_p)$ is the marginal density of $(x_1 \ldots .. x_p)$ and does not contain the parameters β_i or σ^2 .

It follows that the conditional density of y given $(x_1 = X_1, x_2 = X_2, \ldots, x_p = X_p)$ is normal, that is,

$$f(y / x_1 = X_1, \ldots, x_p = X_p) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left\{ -\frac{1}{2} \left(\frac{y - G(X_1 \ldots . X_p)}{\sigma} \right)^2 \right\}$$

and,

$$E(y / x_1 = X_1, \ldots, x_p = X_p) = G(X_1 \ldots . X_p) \quad .$$

If a random sample, $(y_j, X_{1j}, \ldots X_{pj};\ j=1 \ldots . n)$, is taken from this density, that is,

$$y_j = G(X_{1j}, \ldots X_{pj}) + e_j, \qquad j=1 \ldots n,$$

where, e_j = random error,

and if the random errors, e_j, $j=1 \ldots . n$, are uncorrelated, then,

$$E(e_j) = 0$$

$$\mathrm{Var}\,(e_j) = \sigma^2 \quad .$$

This is the Regression Model, Model III, Case 2.

POINT ESTIMATES OF $\underline{\beta}$ AND σ^2

For both of the models maximum likelihood estimates can be obtained for $\underline{\beta}$ and σ^2. For Model I, Case A, the likelihood function is:

$$L(\underline{e} : \underline{\beta}, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{-\frac{\underline{e}'\underline{e}}{2\sigma^2}\right\}$$

$$= \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{-\frac{(\underline{Y} - \underline{X}\underline{\beta})'(\underline{Y} - \underline{X}\underline{\beta})}{2\sigma^2}\right\}$$

i.e. $$\text{Log}[L(\cdot)] = -\frac{n}{2}\text{Log}(2\pi\sigma^2) - \frac{1}{2\sigma^2}(\underline{Y} - \underline{X}\underline{\beta})'(\underline{Y} - \underline{X}\underline{\beta})$$

i.e. $$\frac{\partial}{\partial \underline{\beta}}[\text{Log}[L(\cdot)]] = \frac{1}{\sigma^2}(\underline{X}'\underline{Y} - \underline{X}'\underline{X}\hat{\underline{\beta}}) = 0$$

i.e. $$\frac{\partial}{\partial \sigma^2}[\text{Log}[L(\cdot)]] = -\frac{n}{2\hat{\sigma}^2} + \frac{(\underline{Y} - \underline{X}\hat{\underline{\beta}})'(\underline{Y} - \underline{X}\hat{\underline{\beta}})}{2\hat{\sigma}^4} = 0$$

i.e. $$\hat{\underline{\beta}} = (\underline{X}'\underline{X})^{-1}\underline{X}'\underline{Y} , \qquad \text{(A.2.2)}$$

since $(\underline{X}'\underline{X})$ has an inverse,

and $$\hat{\sigma}^2 = \frac{(\underline{Y} - \underline{X}\hat{\underline{\beta}})'(\underline{Y} - \underline{X}\hat{\underline{\beta}})}{(n-p)} , \qquad \text{(A.2.3)}$$

making a correction for bias in equation (A.2.3).

These estimates have the following properties,

(i) consistent and efficient

(ii) unbiased

(iii) sufficient

(iv) complete

(v) minimum variance unbiased

(vi) $\hat{\underline{\beta}} \sim N(\underline{\beta} , \sigma^2 (\underline{X}'\underline{X})^{-1})$

(vii) $(n-p)\dfrac{\hat{\sigma}^2}{\sigma^2} \sim \chi^2(n-p)$

(viii) $\hat{\underline{\beta}}$ and $\hat{\sigma}^2$ are independent.

For Model III, Case 2, the maximum likelihood estimates of $\underline{\beta}$ and σ^2 can be derived in similar manner. They are exactly the same as the ones derived above, but have the following properties:

(i) consistent and efficient,

(ii) unbiased,

(iii) sufficient.

CONFIDENCE INTERVAL ESTIMATES OF β_i AND σ^2

For Model I, Case A. a confidence interval can be put on σ^2 using the fact that,

$$(n-p)\frac{\hat{\sigma}^2}{\sigma^2} \sim \chi^2(n-p) \quad .$$

This means that two constants, α_0 and α_1, can be found such that,

$$P\{\alpha_0 \leq (n-p)\frac{\hat{\sigma}^2}{\sigma^2} \leq \alpha_1\} = (1-\alpha) \ ,$$

$$P\{\frac{\hat{\sigma}^2(n-p)}{\alpha_1} \leq \sigma^2 \leq \frac{\hat{\sigma}^2(n-p)}{\alpha_0}\} = (1-\alpha) \ ,$$

which is the $(1-\alpha)$ confidence interval on σ^2.

A confidence interval can be obtained for β_i by using the facts that, $\left[\dfrac{\hat{\beta}_i - \beta_i}{\sigma\sqrt{c_i}}\right] \sim N(0, 1)$, where c_i is the i^{th} diagonal element of $(\underline{X}'\underline{X})^{-1}$, and

$$(n-p)\frac{\hat{\sigma}^2}{\sigma^2} \sim \chi^2(n-p) \ ,$$

and that these two statistics are independent. Thus,

$$\left[\frac{\hat{\beta}_i - \beta_i}{\hat{\sigma}\ \sqrt{c_i}}\right] \sim t(n-p)$$

and the constant $t_{\alpha/2}$ can be found such that,

$$P\left\{ - t_{\alpha/2} \leq \frac{\hat{\beta}_i - \beta_i}{\hat{\sigma}\ \sqrt{c_i}} \leq t_{\alpha/2}\right\} = 1 - \alpha$$

i.e. $$P\left\{ (\hat{\beta}_i - t_{\alpha/2}\ \hat{\sigma}\sqrt{c_i}) \leq \beta_i \leq (\hat{\beta}_i + t_{\alpha/2}\ \hat{\sigma}\ \sqrt{c_i})\right\} = 1 - \alpha$$

which is a $(1 - \alpha)$ confidence interval on β_i.

It can be shown (Graybill [7] p. 204, Johnston[9]) that the above confidence intervals on β_i and σ^2 also apply in Model III, Case 2.

TESTING THE HYPOTHESIS, $H_0: \underline{\gamma}_1 = \underline{\gamma}_1^*$

If Model I, Case A, $\underline{Y} = \underline{X}\,\underline{\beta} + \underline{e}$, is partitioned so that,

$$\underline{Y} = \underline{X}_1\ \underline{\gamma}_1 + \underline{X}_2\ \underline{\gamma}_2 + \underline{e}$$

where $\underline{\gamma}_1$ has dimension $(r \times 1)$ then the likelihood ratio rest of the hypothesis, $H_0: \underline{\gamma}_1 = \underline{\gamma}_1^*$, can be found by making use of the fact that the statistic u, defined by,

$$u = \frac{(n-p)}{r}\ \frac{\underline{\hat{\beta}}'\ \underline{X}'\ \underline{Y} - \underline{\tilde{\gamma}}_2\ \underline{X}'_2\ \underline{Y}}{\underline{Y}'\underline{Y} - \underline{\hat{\beta}}'\ \underline{X}'\ \underline{Y}}\ ,$$

has a non-central F-distribution, $F'(r, n-p, \lambda)$, and that if and only if the null hypothesis is true u has a central F-distribution,

$F(r, n-p)$. With probability of type I error α, the hypothesis H_0: $\underline{\gamma}_1 = \underline{\gamma}_1^*$ is rejected if $u > F_\alpha(r, n-p)$.

Again, it can be shown that this test can be applied in Model III, Case 2, with the same probability of type I error.

APPENDIX 3

PROOF THAT: $\mathrm{VAR}\,(\tilde{\bar{K}}) \propto 1/n$

It has been shown that in a grid $[(x, y), (x + \Delta x, y), (x, y + \Delta y), (x + \Delta x, y + \Delta y)]$, the estimate of the mean value of permeability, $\tilde{\bar{K}}$, is normally distributed, viz,

$$\tilde{\bar{K}} \sim N\,(\bar{K},\ \rho_K^2)$$

where, $\bar{K}$ = true mean value of permeability in the grid, and

$$\rho_K^2 = \mathrm{Var}\,(\tilde{\bar{K}})$$

$$= \mathrm{Var}\left\{\frac{1}{\Delta x\,\Delta y}\int_x^{x+\Delta x}\int_y^{y+\Delta y}\sum_{i=0}^{N}\hat{\alpha}_i X_i\,dx\,dy\right\},$$

where $X_0 = 1$,

$$= \mathrm{Var}\left\{\sum_{i=0}^{N}\hat{\alpha}_i\int_x^{x+\Delta x}\int_y^{y+\Delta y}\frac{X_i}{\Delta x\,\Delta y}\,dx\,dy\right\}$$

$$= \mathrm{Var}\left\{\sum_{i=0}^{N}\hat{\alpha}_i J_i\right\}$$

where $$J_i = \int_x^{x+\Delta x}\int_y^{y+\Delta y}\frac{X_i}{\Delta x\,\Delta y}\,dx\,dy, \quad i=0\ \ldots\ N.$$

Therefore, $$\rho_K^2 = \sum_{i=0}^{N}\sum_{j=0}^{N} J_i J_j\,\mathrm{Cov}\,(\hat{\alpha}_i, \hat{\alpha}_j) \qquad . \qquad (A.\,3.\,1)$$

Now, the covariance matrix of $\hat{\alpha}$ is $\sigma_K^2\,(\underline{X}'\underline{X})^{-1}$,

where,

$$\underline{X} = \begin{bmatrix} 1 & X_{11} & X_{12} & & X_{1N} \\ 1 & X_{21} & X_{22} & & X_{2N} \\ \cdot & \cdot & & & \cdot \\ \cdot & & \ddots & & \cdot \\ \cdot & & & & \cdot \\ 1 & X_{n1} & \cdots & \cdots & X_{nN} \end{bmatrix}$$

and

$$\underline{X}'\underline{X} = \begin{bmatrix} n & \sum_{i=1}^{n} X_{i1} & \cdots & \sum_{i=1}^{n} X_{iN} \\ \sum_{i=1}^{n} X_{i1} & \sum_{i=1}^{n} X_{i1}^{2} & & \cdot \\ \cdot & & \ddots & \cdot \\ \cdot & & & \cdot \\ \sum_{i=1}^{n} X_{iN} & \cdots & \cdots & \sum_{i=1}^{n} X_{iN}^{2} \end{bmatrix}$$

The general term of this matrix is,

$$\sum_{i=1}^{n} X_{ij} X_{ik} , \qquad j=0 \ldots N, \; k=0 \ldots N .$$

This can be rewritten,

$$n \sum_{i=1}^{n} \left(\frac{X_{ij} X_{ik}}{n}\right) ,$$

so that the term $\sum_{i=1}^{n} \left(\frac{X_{ij} X_{ik}}{n}\right)$ will always have the same order of magnitude for any value of n.

Thus,

$$\underline{X}'\underline{X} = n \begin{bmatrix} 1 & \sum_{i=1}^{n} \frac{X_{i1}}{n} & \cdots\cdots\cdots & \sum_{i=1}^{n} \frac{X_{iN}}{n} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ \sum_{i=1}^{n} \frac{X_{iN}}{n} & \cdots\cdots\cdots & & \sum_{i=1}^{n} \frac{X_{iN}^2}{n} \end{bmatrix} = n\,\underline{Y}$$

and $(\underline{X}'\underline{X})^{-1} = \frac{1}{n}\,\underline{Y}^{-1}$.

Therefore, $\mathrm{Cov}\,(\alpha_i, \alpha_j) = \frac{\sigma_K^2}{n} F(X_{ij})$,

where, $F(X_{ij})$ is a function of the space coordinates and has the same order of magnitude for any value of n.

In equation (A. 3. 1) J_i and J_j are constants,

thus,

$$\rho_K^2 = \frac{\sigma_K^2}{n} \sum_{i=0}^{N} \sum_{j=0}^{N} J_i\, J_j\, F(X_{ij}) ,$$

i. e. $$\rho_K^2 = \frac{\sigma_K^2}{n} F'(X_{ij})$$

i. e. $$\rho_K = \frac{\sigma_K}{\sqrt{n}} F''(X_{ij})$$

i. e. $$\rho_K \propto \frac{1}{\sqrt{n}} .$$

APPENDIX 4

TOLERANCE LIMITS*

The development of tolerance limits is based on a simple property of order statistics, namely that the distribution of the area under the density function between any two ordered observations is independent of the form of the density function. This is stated in Theorem. 1.

<u>Theorem. 1.</u>

If z is a continuous random variable with density $f_z(z)$, $-\infty < z < \infty$, and $X_1 \ldots . X_N$ is an ordered sample from this distribution, and

$$u_i = \int_{-\infty}^{X_i} f(z)\, dz = F(X_i) \quad , \quad i=1 \ldots N,$$

then, the joint density of u_i, given by

$$g(u_1 \ldots u_N) = N! \quad , \quad 0 < u_1 < u_2 < \ldots < u_N < 1$$

is independent of $f_z(z)$.

From $g(u_1 \ldots u_N)$, the distribution of the area under $f_z(z)$ between any pair of ordered observations can be found, viz.,

---

*This description is based on that given by Mood and Graybill [14].

Theorem. 2.

Let the random variable $Y_{i,j}$ be the area under $f_z(z)$ between X_i and X_j $(i < j)$, then, the density of $Y_{i,j}$ is,

$$f_{Y_{i,j}}(y) = \frac{N!}{(j-i-1)!(N-j+i)!} (y)^{j-i-1} (1-y)^{N-j+i} , \quad 0 < y < 1 .$$

Tolerance limits are defined to be L_1 and L_2 such that,

$$P\left\{\int_{L_1}^{L_2} f_z(z)\, d\, z > \beta\right\} = 1 - \alpha$$

L_1 and L_2 are functions of the ordered sample from the density $f_z(z)$, and the density of $\int_{L_1}^{L_2} f_z(z)\, dz$ is given by Theorem 2.

Thus, if $Y = \int_{L_1}^{L_2} f_z(z)\, dz$, then $P\{ Y > \beta \} = 1 - \alpha$

i.e. $$\int_{\beta}^{1} \frac{N!}{\beta(j-i-1)!(N-j+1)!} (y)^{j-i-1} (1-y)^{N-j+1}\, dy = 1-\alpha \quad .$$

From this equation, if any three of the four variables, α, β, $[L_1, L_2]$, N, are specified the other can be determined. For example, if $(1-\alpha) = .9$, $\beta = .9$, $N = 38$, then, $L_1 = X_1$ and $L_2 = X_N$.

APPENDIX 5

COMPUTER PROGRAM

The computer program as described here is written to obtain a finite difference solution to the unconfined flow equation when the region of flow is rectangular and the boundary conditions are constant gradients. Simple modifications can be made for other flow regions and boundary conditions. The regression models,

$$E(t'') = \alpha_0 + \alpha_1 \tilde{\rho}_K + \alpha_2 \tilde{\rho}_z + \alpha_3 \tilde{\rho}_{h_I} + \alpha_4 \tilde{\rho}_Q + \alpha_5 \tilde{\rho}_S$$

$$E(c'') = \alpha_0 + \alpha_1 \tilde{c}_K + \alpha_2 \tilde{c}_z + \alpha_3 \tilde{c}_{h_I} + \alpha_4 \tilde{c}_Q + \alpha_5 \tilde{c}_S$$

are analyzed by the program.

DESCRIPTION OF PROGRAM SUBROUTINES

Subroutine READATA

The following variables are read in by this subroutine;

NBETA, NTSTEP, NVAR, NRUN, NROW, NCOL,STUDENT

TIM

KBC,

FKMIN, FKMED, FKMAX

HMIN, HMED, HMAX

ZMIN, ZMED, ZMAX,

QMIN, QMED, QMAX,

PHIMIN, PHIMED, PHIMAX,

DX, DY,

GTMIN, GTMED, GTMAX,

GBMIN, GBMED, GBMAX,

GRMIN, GRMED, GRMAX,

GLMIN, GLMED, GLMAX.

These variables are defined in the program. The data is also written out by the subroutine.

Subroutine RANDOM

This subroutine generates random values of the mean of the normal distributions of h_I, K, Q, S and z from triangular distributions defined by the minimum, median and maximum values of each variable. Random values of the standard deviations of the normal distributions of these input variables are generated from uniform distributions defined by the upper and lower values of these variables. These upper and lower values are defined in the subroutine. Random values of the constant gradient boundary conditions, GTOP, GBOT, GRITE, GLEFT, are also generated from triangular distributions defined by the minimum, median and maximum values of the variables.

Subroutine RANDY

This subroutine generates random values of h_I, z, Q, K, S from their normal distributions.

Subroutine AMATRIX

This subroutine computes the values of the elements of the matrix equation which results from writing the finite difference equation for each of the interior grids of the flow region.

Subroutine RBSOLV

This subroutine solves the matrix equation set up in subroutine AMATRIX by Gauss Elimination. Savings in computer storage and execution time are effected by condensing the matrix from being square with dimension (NROW-2)*(NCOL-2) to being rectangular with (NROW-2)*(NCOL-2) rows and NROW columns. The dimensions of this reduced matrix make it desirable to define NCOL $\geq$ NROW.

Subroutine MINMAX

This computes the values of tolerance interval width and coefficient of variation in each grid for the end of each time step.

Subroutine POLY

This subroutine computes the maximum likelihood estimates of the regression coefficients, confidence intervals on the coefficients and tests the hypotheses $H_o : [\alpha_0, \alpha_1, \alpha_2, \alpha_4, \alpha_5] = 0$ and $H_o : [\alpha_1 \alpha_2 \alpha_4 \alpha_5] = 0$. The results of these computations are written out.

Data Preparation

Before using the program it is necessary to define the dimensions of the variables to suit the problem being studied. This involves the DIMENSION statements of the main program and subroutine POLY and the COMMON statements. Also, tolerance interval widths are computed in this program from a random sample of size 38. If this sample size is changed adjustments have to be made in subroutine MINMAX. Changes would also have to be made to study different regression models.

Data cards are read in as follows (for NCOL=7, NROW=6, NTSTEP$\leq$ 16)

CARD 1 NBETA, NTSTEP, NVAR, NRUN, NROW, NCOL, STUDENT FORMAT (6 110, F10·3)

CARD 2 TIM (I) , I = 1, NTSTEP

FORMAT (16F5·1)

CARDS 3-8 KBC (I, J), J=1, NCOL, I=1, NROW

FORMAT (7 I 10)

CARDS 9-14 FKMIN (I, J), J=1, NCOL, I=1, NROW

FORMAT (7 F 10·2)

CARDS 15-20 FKMED (I , J), J=1, NCOL, I=1, NROW

FORMAT (7 F 10·2)

CARDS 21-26 FKMAX (I,J), J=1, NCOL, I=1, NROW

FORMAT (7 F 10·2)

CARDS 27-32 HMIN (I, J), J=1, NCOL, I=1, NROW

FORMAT (7 F 10·2)

CARDS 33-38 HMED (I, J), J=1, NCOL, I=1, NROW

FORMAT (7 F 10·2)

CARDS 39-44 HMAX (I, J), J=1, NCOL, I=1, NROW

FORMAT (7 F 10·2)

CARDS 45-50 ZMIN (I, J), J=1, NCOL, I=1, NROW

FORMAT (7 F 10·2)

CARDS 51-56 ZMED (I, J), J=1, NCOL, I=1, NROW

FORMAT (7 F 10·2)

CARDS 57-62 ZMAX (I, J), J=1, NCOL, I=1, NROW

FORMAT (7 F 10·2)

CARDS 63-68 QMIN (I, J), J=1, NCOL, I=1, NROW

FORMAT (7 F 10·2)

CARDS 69-74 QMED (I, J), J=1, NCOL, I=1, NROW

FORMAT (7 F 10·2)

CARDS 75-80 QMAX (I, J), J=1, NCOL, I=1, NROW

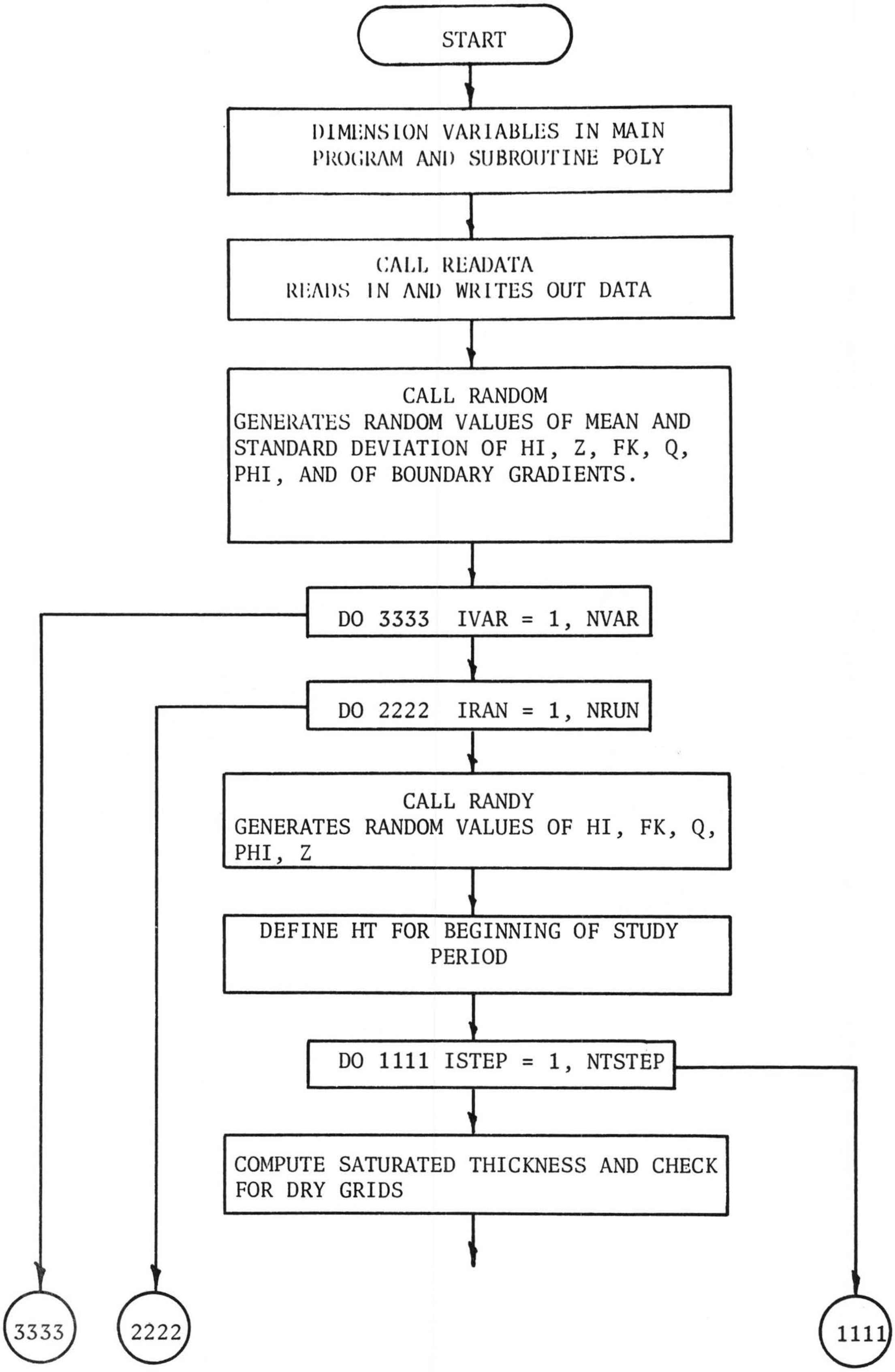
FORMAT (7 F 10·2)

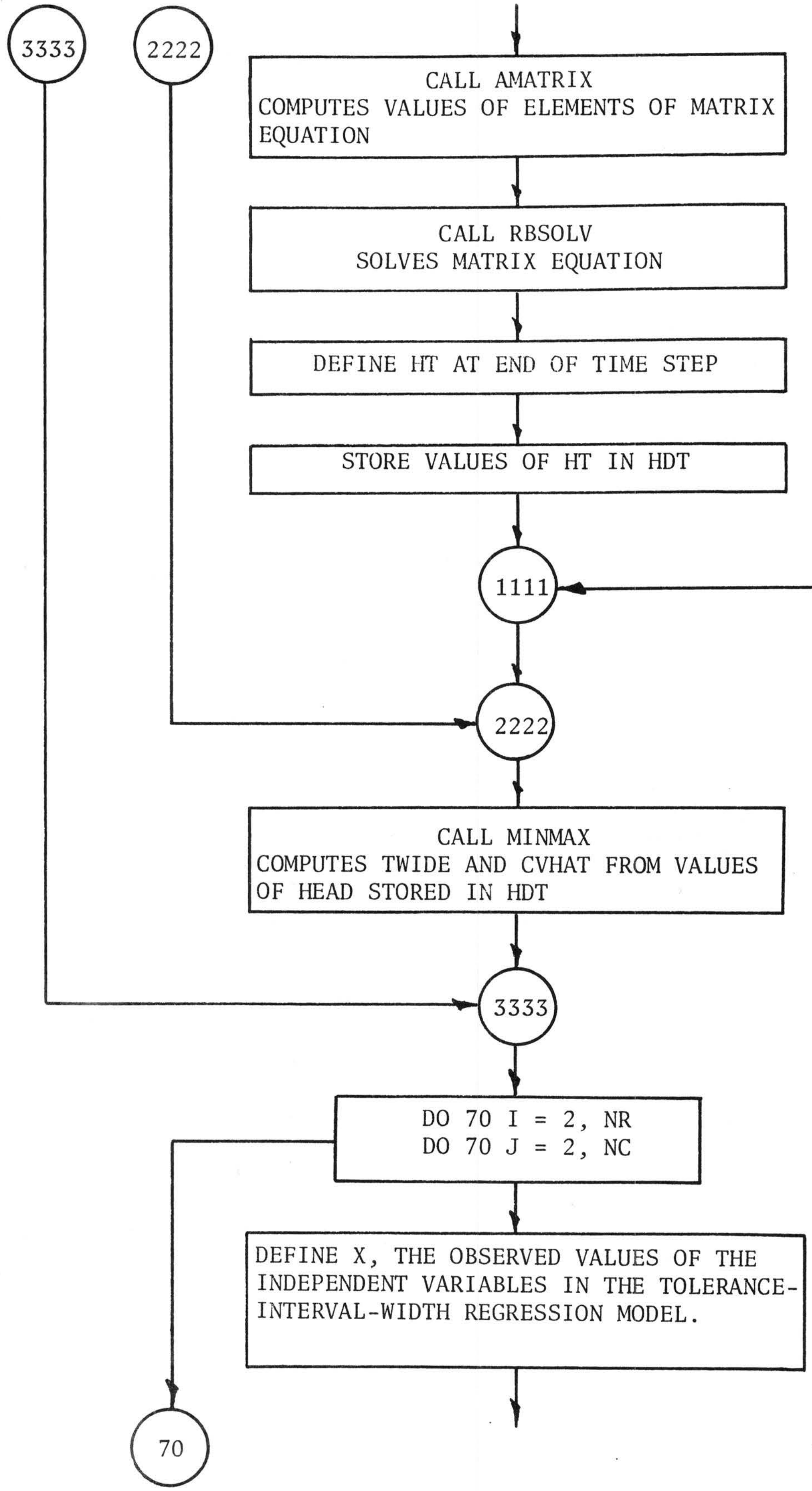
CARDS 81-86 PHIMIN (I, J), J=1, NCOL, I=1, NROW

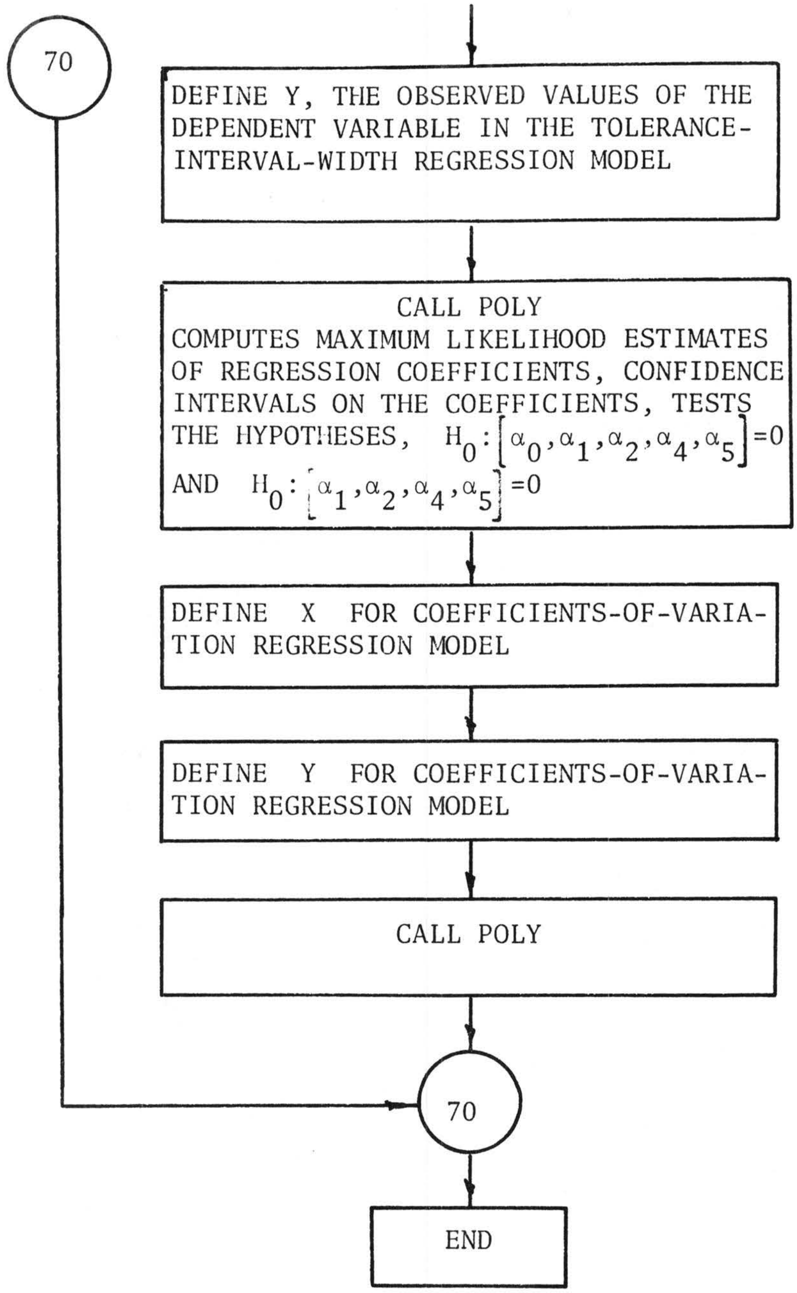
FORMAT (7 F 10.2)

CARDS 87-92	PHIMED (I, J), J=1, NCOL, I=1, NROW
	FORMAT (7 F 10·9)
CARDS 93-98	PHIMAX (I, J), J=1, NCOL, I=1, NROW
	FORMAT (7 F 10·9)
CARDS 99-104	DX (I, J), J=1, NCOL, I=1, NROW
	FORMAT (7 F 10·2)
CARDS 105-110	DY (I, J), J=1, NCOL, I=1, NROW
	FORMAT (7 F 10·2)
CARD 111	GTMIN (J), J=1, NCOL
	FORMAT (7 F 10·2)
CARD 112	GTMED (J), J=1, NCOL
	FORMAT (7 F 10·2)
CARD 113	GTMAX (J), J=1, NCOL
	FORMAT (7 F 10·2)
CARD 114	GBMIN (J), J=1, NCOL
	FORMAT (7 F 10·2)
CARD 115	GBMED (J), J=1, NCOL
	FORMAT (7 F 10·2)
CARD 116	GBMAX (J), J=1, NCOL
	FORMAT (7 F 10·2)
CARD 117	GRMIN (I), I=1, NROW
	FORMAT (7 F 10·2)
CARD 118	GRMED (I), I=1, NROW
	FORMAT (7 F 10·2)

CARD 119	GRMAX (I), I=1, NROW
	FORMAT (7 F 10·2)
CARD 120	GLMIN (I), I=1, NROW
	FORMAT (7 F 10·2)
CARD 121	GLMED (I), I=1, NROW
	FORMAT (7 F 10·2)
CARD 122	GLMAX (I), I=1, NROW
	FORMAT (7 F 10·2)


START
DIMENSION VARIABLES IN MAIN PROGRAM AND SUBROUTINE POLY
CALL READATA
READS IN AND WRITES OUT DATA
CALL RANDOM
GENERATES RANDOM VALUES OF MEAN AND STANDARD DEVIATION OF HI, Z, FK, Q, PHI, AND OF BOUNDARY GRADIENTS.
DO 3333 IVAR = 1, NVAR
DO 2222 IRAN = 1, NRUN
CALL RANDY
GENERATES RANDOM VALUES OF HI, FK, Q, PHI, Z
DEFINE HT FOR BEGINNING OF STUDY PERIOD
DO 1111 ISTEP = 1, NTSTEP
COMPUTE SATURATED THICKNESS AND CHECK FOR DRY GRIDS
3333
2222
1111


3333
2222
1111
CALL AMATRIX
COMPUTES VALUES OF ELEMENTS OF MATRIX EQUATION
CALL RBSOLV
SOLVES MATRIX EQUATION
DEFINE HT AT END OF TIME STEP
STORE VALUES OF HT IN HDT
1111
2222
CALL MINMAX
COMPUTES TWIDE AND CVHAT FROM VALUES OF HEAD STORED IN HDT
3333
DO 70 I = 2, NR
DO 70 J = 2, NC
DEFINE X, THE OBSERVED VALUES OF THE INDEPENDENT VARIABLES IN THE TOLERANCE-INTERVAL-WIDTH REGRESSION MODEL.
70


70
DEFINE Y, THE OBSERVED VALUES OF THE DEPENDENT VARIABLE IN THE TOLERANCE-INTERVAL-WIDTH REGRESSION MODEL
CALL POLY
COMPUTES MAXIMUM LIKELIHOOD ESTIMATES OF REGRESSION COEFFICIENTS, CONFIDENCE INTERVALS ON THE COEFFICIENTS, TESTS THE HYPOTHESES, $H_0: [\alpha_0, \alpha_1, \alpha_2, \alpha_4, \alpha_5] = 0$ AND $H_0: [\alpha_1, \alpha_2, \alpha_4, \alpha_5] = 0$
DEFINE X FOR COEFFICIENTS-OF-VARIATION REGRESSION MODEL
DEFINE Y FOR COEFFICIENTS-OF-VARIATION REGRESSION MODEL
CALL POLY
70
END

```
      PROGRAM STATFLO
C THIS PROGRAM SOLVES UNCONFINED FLOW PROBLEM WHEN HI, Z, PHI, Q, FK, ARE RANDOM
C INPUT VARIABLES WITH NORMAL DISTRIBUTIONS.
C GRID SYSTEM MUST BE RECTANGULAR
C BOUNDARY CONDITIONS MUST BE CONSTANT GRADIENTS.
C IN THE REGRESSION MODELS,
C                          T=A0+A1*SK+A2*SZ+A3*SH+A4*SQ+A5*SP
C                          C=A0+A1*CK+A2*CZ+A3*CH+A4*CQ+A5*CP
C WHERE,
C       SK,CK=STAN. DEV.,COEFF OF VAR. OF PERMEABILITY
C       SZ,CZ=STAN. DEV.,COEFF OF VAR. OF BEDROCK ELEV.
C       SH,CH=STAN. DEV.,COEFF OF VAR. OF INITIAL HEAD
C       SQ,CQ=STAN. DEV.,COEFF OF VAR. OF DISCHARGE
C       SP,CP=STAN. DEV.,COEFF OF VAR. OF SPECIFIC YIELD
C THE MAX. LIKELIHOOD ESTIMATES OF THE REGRESSION COEFFS. ARE CALCLUATED AND
C THE HYPOTHESES (AI)=0, (A0,A1,A2,A4,A5)=0, (A1,A2,A4,A5)=0,ARE TESTED.


C DIMENSION THE FOLLOWING VARIABLES,
C AA(IP,NROW) = MATRIX OF COEFFS. FROM FINITE DIFFERENCE EQUATIONS.
C RHS(IP) =     VECTOR ''    ''    ''    ''    ''    ''
C A(NBETA-1)
C Y(NTSTEP,NVAR)
C WHERE,
C       NROW = NO. OF ROWS OF GRID SYSTEM
C       NCOL = NO.OF COLS. OF GRID SYSTEM(NCOL GREATER THAN OR EQUAL TO NROW)
C       IP = (NROW-2)*(NCOL-2)
C       NTSTEP = NO. OF TIME STEPS
C       NVAR = NO. OF POINTS TO BE GENERATED FOR REGRESSION ANALYSIS
C       NBETA = NO. OF REGRESSION COEFFS..


      DIMENSION AA(20,6), RHS(20), A(5), Y(10,15)


C CDTIME = TIME AT BEGINNING OF A TIME STEP
C CTIME = TIME AT END OF A TIME STEP
C NRUN = NO. OF OBSERVATIOVS TO BE GENERATED FOR COMPUTATION OF TOLERANCE INTERV
C DX(NROW,NCOL) = GRID SIZE IN X-DIRECTION(FEET)
C DY(NROW,NCOL) = GRID SIZE IN Y-DIRECTION(FEET)
C KBC(NROW,NCOL) = PARAMETER CHARACTERISING BOUNDARY CONDITIONS
C KBC = 1, IF BOUNDARY GRID WITH CONST HEAD OR CONST GRADIENT
C KBC = 2, IF IMPERM BOUNDARY GRID
C KBC = 3, IF INTERIOR GRID
C XNORM(IP*3+2*NROW*NCOL) = STORES STAN. NORMAL RANDOM NOS..
C FK(NROW,NCOL) = PERMEABILITY IN EACH GRID(FEET/DAY)
C B(NROW,NCOL) = SATURATED THICKNESS IN EACH GRID(FEET)
C Q(NROW,NCOL) = DISCHARGE IN EACH GRID(FEET**3/DAY)
C HI(NROW,NCOL) = INITIAL HEAD IN EACH GRID (FEET)
C HT(NROW,NCOL) = HEAD AT CURRENT TIME IN EACH GRID(FEET)
C PHI(NROW,NCOL) = SPECIFIC YIELD IN EACH GRID (-)
C Z(NROW,NCOL) = BEDROCK ELEVATION IN EACH GRID (FEET)
C FKMEAN(NROW,NCOL) = MEAN VALUE OF PERMEABILITY IN EACH GRID (FEET/DAY)
C FKMIN(NROW,NCOL) = MINIMUM VALUE OF FKMEAN IN EACH GRID(FT/DAY)
C FKMED(NROW,NCOL) = MEDIAN  VALUE OF FKMEAN IN EACH GRID(FT/DAY)
C FKMAX(NROW,NCOL) = MAXIMUM VALUE OF FKMEAN IN EACH GRID(FT/DAY)
C FKLOW(NROW,NCOL) = MINIMUM VALUE OF FKVAR IN EACH GRID(FEET/DAY)
C  FKUP(NROW,NCOL) = MAXIMUM VALUE OF FKVAR IN EACH GRID(FEET/DAY)
C ANALAGOUS VARIABLES FOR Q,HI,PHI,Z HAVE SAME DEFINITIONS
C FKVAR(NROW,NCOL,NVAR) = STANDARD DEVIATION OF PERMEABILITY IN EACH GRID(FT/DAY
C ZVAR(NROW,NCOL,NVAR) = STAN. DEV. OF BEDROCK IN EACH GRID(FEET)
C HIVAR(IP,NVAR) = STAN.DEV.OF INITIAL HEAD    IN EACH GRID(FEET)
C QVAR(IP,NVAR) = STAN. DEV. OF DISCHARGE IN EACH GRID(FT**3/DAY)
C PHIVAR(IP,NVAR) = STAN. DEV. OF SPECIFIC YIEID IN EACH GRID(-)
C HOT(IP,NRUN,NTSTEP) = STOTES VALUES OF HEAD IN EACH GRID FOR ALL NRUN RUNS AT
C                       END OF EACH TIME STEP
C STUDENT = VALUE OF STUDENTS T AT 95 PER CENT LEVEL FOR (NVAR-NBETA) D. OF F.
C TWIDE(IP,NVAR,NTSTEP) = STORES VALUES OF TOLERANCE INTERVAL WIDTH IN EACH GRID
C                         AT END OF EACH TIME STEP
C CVHAT(IP,NVAR,NTSTEP) = STOREA VALUES OF SAMPLE COEFFS. OF VARIATION IN EACH
C                         GRID AT END OF EACH TIME STEP
C X(NBETA-1,NVAR)
```

```
C TIM(NTSTEP) = LENGTH OF EACH TIME STEP
C GTOP(NCOL,NTSTEP+1) = GRADIENTS FOR 'TOP' OF GRID SYSTEM
C GBOT(NCOL,NTSTEP+1) = GRADIENTS FOR 'BOTTOM' OF GRID SYSTEM
C GRITE(NROW,NTSTEP+1) GRADIETTS FOR 'RIGHT' OF GRID SYSTEM
C GLEFT(NROW,NTSTEP+1) GRADIENTS FOR 'LEFT'  OF GRID SYSTEM
C GTMIN,GTMED,GTMAX(NCOL) = MINIMUM, MEDIAN, MAXIMUM VALUES OF GTOP
C SIMILARLY FOR GBOT, GRITE, GLEFT


      COMMON CTIME, CDTIME, NTSTEP, NROW, NCOL, NVAR, NRUN
     1 , DX(6,7), DY(6,7), KBC(6,7), XNORM(150)
     2 , FK(6,7), B(6,7), Q(6,7), HI(6,7), PHI(6,7), HT(6,7), Z(6,7)
     3 , FKMEAN(6,7), ZMEAN(6,7), QMEAN(6,7), HIMEAN(6,7), PHIMEN(6,7)
     4 , FKVAR(6,7,15), ZVAR(6,7,15)
     * , QVAR(20,15), HIVAR(20,15), PHIVAR(20,15)
     5 , FKMIN(6,7),ZMIN(6,7),QMIN(6,7),HMIN(6,7),PHIMIN(6,7)
     6 , FKMED(6,7), ZMED(6,7), QMED(6,7), HMED(6,7), PHIMED(6,7)
     7 , FKMAX(6,7), ZMAX(6,7), QMAX(6,7), HMAX(6,7), PHIMAX(6,7)
     8 , FKLOW(6,7), ZLOW(6,7), QLOW(6,7), HLOW(6,7), PHILOW(6,7)
     9 , FKUP(6,7), ZUP(6,7), QUP(6,7), HUP(6,7), PHIUP(6,7)
     0 , HDT(20,38,10), TWIDE(20,15,10), CVHAT(20,15,10)
     1 , STUDENT, NBETA
     4 , X(5,15)
     6 , TIM(10)
     7 , GTMIN(7), GTMED(7), GTMAX(7), GBMIN(7), GBMED(7), GBMAX(7)
     8 , GRMIN(7), GRMED(7), GRMAX(7), GLMIN(7), GLMED(7), GLMAX(7)
     9 , GTOP(7,11), GBOT(7,11), GRITE(6,11), GLEFT(6,11)
C CALL READATA TO READ IN AND WRITE OUT INITIAL DATA
      CALL READATA
C CALL RANDOM TO GENERATE RANDOM VALUES OFMEAN AND STANDARD DEV. OF INPUT VARIAB
      CALL RANDOM
      NC = NCOL - 1
      NR = NROW - 1
      IP = (NROW - 2)*(NCOL - 2)
C EACH TIME THRU LOOP 3333 COMPUTES NVAR VALUES OF TWIDE AND CVHAT AT END OF
C EACH TIME STEP CORRESPONDING TO NVAR VALUES OF STAN. DEV. AND COEFFS. OF VAR.
C ON HI, Z, FK, PHI, Q.
      DO 3333 IVAR = 1,NVAR
C EACH TIME THRU LOOP 2222 COMPUTES NRUN RANDOM VALUES OF HEAD AT END OF EACH
C TIME STEP
      DO 2222 IRAN = 1,NRUN
C CALL RANDY TO GENERATE RANDOM VALUES OF HI, FK, Q, PHI, Z.
      CALL RANDY(IVAR)
C SET HT = HI FOR INTERIOR GRIDS
      DO 20 I = 2,NR
      DO 20 J = 2,NC
   20 HT(I,J) = HI(I,J)
C DEFINE HT FOR BOUNDARY GRIDS AS HT FOR ADJOINING INTERIOR GRID + CONSTANT GRAD
      DO 21 I = 2,NR
      HT(I,1) = HT(I,2) + GLEFT(I,1)
   21 HT(I,NCOL) = HT(I,NC) + GRITE(I,1)
      DO 22 J = 2,NC
      HT(1,J) = HT(2,J) + GTOP(J,1)
   22 HT(NROW,J) = HT(NR,J) + GBOT(J,1)
C EACH TIME THRU LOOP 1111 COMPUTES HEAD AT END OF ONE TIME STEP
      CTIME = 0.0
      DO 1111 ISTEP = 1,NTSTEP
C COMPUTE SATURATED THICKNESS IN EACH GRID AND CHECK IF NON-NEGATIVE. IF NEGATIV
C WRITE OUT LOCATION OF DRY GRID AND STOP EXECUTION.
      DO 50 I = 1,NROW
      DO 50 J = 1,NCOL
      B(I,J) = HT(I,J) - Z(I,J)
      KCHECK = B(I,J)/1000000.0 + 2.0
      GO TO (51,50),KCHECK
   51 WRITE(6,52) I,J
   52 FORMAT(1HO,*DRY GRID *,2I3)
      STOP
   50 CONTINUE
      CDTIME = CTIME
      CTIME = CTIME + TIM(ISTEP)
C CALL AMATRIX TO SET UP MATRIX (AA) AND VECTOR (RHS)
      CALL AMATRIX(IRAN,RHS,AA)
```

```
C CALL RMSOLV TO SOLVE MATRIX EQU. (AA)*(HT) = (RHS)
      CALL RBSOLV(AA,IP,NROW,RHS)
C REDEFINE HT AS VALUES OF HEAD AT END OF TIME STEP
      L = 0
      DO 1 J = 2,NC
      DO 1 I = 2,NR
      L = L + 1
    1 HT(I,J) = RHS(L)
C REDEFINE HT IN BOUNDARY GRIDS AS HT IN ADJOINING GRID PLUS CONSTANT GRADIENT
      IS = ISTEP + 1
      DO 23 I = 2,NR
      HT(I,1) = HT(I,2) + GLEFT(I,IS)
   23 HT(I,NCOL) = HT(I,NC) + GRITE(I,IS)
      DO 24 J = 2,NC
      HT(1,J) = HT(2,J) + GTOP(I,IS)
   24 HT(NROW,J) = HT(NR,J) + GBOT(I,IS)
C STORE VALUES OF HT IN HDT FOR INTERIOR GRIDS AT END OF EACH TIME SYEP
      IGRID = 0
      DO 7 I = 2,NR
      DO 7 J = 2,NC
      IGRID = IGRID + 1
      HDT(IGRID,IRAN,ISTEP) = HT(I,J)
    7 CONTINUE
 1111 CONTINUE
 2222 CONTINUE
C CALL MINMAX TO COMPUTE TOL. INT. WIDTHS AND COEFFS. OF VAR. IN EACH GRID AT
C END EACH TIME STEP  FROM RANDOM SAMPLES OF HEAD STORED IN HDT
      CALL MINMAX(HDT,NRUN,NROW,NCOL,TWIDE,IVAR,IP,NVAR,NTSTEP,CVHAT)
 3333 CONTINUE
C LOOP 70 COMPUTES ESTIMATES OF REGRESSION COEFFS., CONFIDENCE INTERVALS ON
C THESE ESTIMATES AND TESTS OF HYPOTHESES FOR BOTH 'TOLERANCE-INTERVAL-WIDTH'
C AND 'COEFFS-OF-VAR' REGRESSION MODELS, FOR ONE GRID
      NBETA1 = NBETA + NTSTEP
      N1 = NBETA - 1
      IGRID = 0
      DO 70 I = 2,NR
      DO 70 J = 2,NC
      IGRID = IGRID + 1
      WRITE(6,72) IGRID
   72 FORMAT(1H0,*RESULTS ROR GRID*,I3)
C LOOP 321 STORES OBSERVED VALUES OF INDEPENDENT VARIABLES OF REGRESSION EQU.
C FOR 'TOL-INT-WIDTH' MODEL IN (X)
      DO 321 K = 1,NVAR
      X(1,K) = FKVAR(I,J,K)
      X(2,K) = ZVAR(I,J,K)
      X(3,K) = HIVAR(IGRID,K)
      X(4,K) =  QVAR(IGRID,K)
      X(5,K) =PHIVAR(IGRID,K)
  321 CONTINUE
C LOOP 71 STORES OBSERVED VALUES AFTER EVERY TIME STEP OF DEPENDENT VARIABLE
C (TWIDE) OF REGRESSION EQU.. IN (Y)
      DO 71 K = 1,NVAR
      DO 71 L = 1,NTSTEP
      Y(L,K) = TWIDE(IGRID,K,L)
   71 CONTINUE
C CALL POLY TO PERFORM REGRESSION ANALYSIS FOR 'TOL-INT-WIDTH' MODEL
      CALL POLY(Y,X,NVAR,NBETA,NBETA1,STUDENT,IGRID,N1,IP,NTSTEP)
C LOOP 74 STORES OBSERVED VALUES OF INDEP. VARIABLES OF REGRESSION EQU., FOR
C COEFFS. OF VARIABLE REGRESSION MODEL, IN (X)
      A(1) = FKMEAN(I,J)
      A(2) = ZMEAN(I,J)
      A(3) = HIMEAN(I,J)
      A(4) = QMEAN(I,J)
      A(5) = PHIMEN(I,J)
      DO 74 K = 1,NVAR
      DO 75 L = 1,N1
   75 X(L,K) = X(L,K)/A(L)
   74 CONTINUE
C LOOP 77 STORES OBSERVED VALUES AFTER EVERY TIME STEP OF DEPENDENT VARIABLES
C (CVHAT) OF REGRESSION EQU.. IN Y
      DO 77 K = 1,NVAR
      DO 77 L = 1,NTSTEP
```

```
      Y(L,K) = CVHAT(IGRID,K,L)
   77 CONTINUE
C CALL POLY TO PERFORM REGRESSION ANALYSIS OF 'COEFF-OF-VAR' MODEL
      CALL POLY(Y,X,NVAR,NBETA,NBETA1,STUDENT,IGRID,N1,IP,NTSTEP)
   70 CONTINUE
      END
      SUBROUTINE RANDY(IVAR)
C THIS SUBROUTINE GENERATES RANDOM VALUES OF HI, Z, Q, PHI, FK FROM THEIR NORMAL
C DISTRIBUTIONS IN EACH GRID.  THESE VALUES ARE HELD CONSTANT FOR ENTIRE STUDY
C PERIOD
      COMMON CTIME, CDTIME, NTSTEP, NROW, NCOL, NVAR, NRUN
     1 , DX(6,7), DY(6,7), KBC(6,7), XNORM(150)
     2 , FK(6,7), B(6,7), Q(6,7), HI(6,7), PHI(6,7), HT(6,7), Z(6,7)
     3 , FKMEAN(6,7), ZMEAN(6,7), QMEAN(6,7), HIMEAN(6,7), PHIMEN(6,7)
     4 , FKVAR(6,7,15), ZVAR(6,7,15)
     * , QVAR(20,15), HIVAR(20,15), PHIVAR(20,15)
     5 , FKMIN(6,7),ZMIN(6,7),QMIN(6,7),HMIN(6,7),PHIMIN(6,7)
     6 , FKMED(6,7), ZMED(6,7), QMED(6,7), HMED(6,7), PHIMED(6,7)
     7 , FKMAX(6,7), ZMAX(6,7), QMAX(6,7), HMAX(6,7), PHIMAX(6,7)
     8 , FKLOW(6,7), ZLOW(6,7), QLOW(6,7), HLOW(6,7), PHILOW(6,7)
     9 , FKUP(6,7), ZUP(6,7), QUP(6,7), HUP(6,7), PHIUP(6,7)
     0 , HDT(20,38,10), TWIDE(20,15,10), CVHAT(20,15,10)
     1 , STUDENT, NBETA
     4 , X(5,15)
     6 , TIM(10)
     7 , GTMIN(7), GTMED(7), GTMAX(7), GBMIN(7), GBMED(7), GBMAX(7)
     8 , GRMIN(7), GRMED(7), GRMAX(7), GLMIN(7), GLMED(7), GLMAX(7)
     9 , GTOP(7,11), GBOT(7,11), GRITE(6,11), GLEFT(6,11)
C N = TOTAL NO. OF STAN. NORMAL RANDOM NUMBERS REQUIRED
      N = (NROW - 2)*(NCOL - 2)*3 + 2*NROW*NCOL
      NR = NROW - 1
      NC = NCOL - 1
C IN THIS LOOP GENERATE INDEP. STAN. NORMAL NUMBERS - FOR ALGORITHM SEE 'HANDBOO
C -K OF MATH.FNS.' NATIONAL BUREAU OF STANDARDS, PAGE 953
      DO 1 I = 1,N,2
      R1 = RANF(0)
      R2 = RANF(0)
      AL = -ALOG(R1)
      X1 = 1.414213562373*SQRT(AL)
      AL = 6.2831853073*R2
      XNORM(I) = X1*COS(AL)
    1 XNORM(I+1) = X1*SIN(AL)
C TRANSFORM STAN. NORMAL RANDOM NUMBERS TO GIVE NORMAL RANDOM NUMBERS FOR HI, Q,
C PHI, FK, Z.
C FOR HI, Q, PHI ONLY NEED TO GENERATE N/S. FOR INTERIOR GRIDS.FOR FK, Z NEED TO
C GENERATE NOS. FOR ALL GRIDS.
      K = 0
      IGRID = 0
      DO 4 I = 2,NR
      DO 4 J = 2,NC
      IGRID = IGRID + 1
      K = K + 1
      HI(I,J) = XNORM(K)* HIVAR(IGRID,IVAR) +  HIMEAN(I,J)
      K = K + 1
      Q(I,J)  = XNORM(K)*  QVAR(IGRID,IVAR) +   QMEAN(I,J)
      K = K + 1
      PHI(I,J)= XNORM(K)*PHIVAR(IGRID,IVAR) +  PHIMEN(I,J)
    4 CONTINUE
      DO 5 I = 1,NROW
      DO 5 J = 1,NCOL
      K = K + 1
      FK(I,J) = XNORM(K)*FKVAR(I,J,IVAR) + FKMEAN(I,J)
      K = K + 1
    5 Z(I,J) = XNORM(K)*ZVAR(I,J,IVAR) + ZMEAN(I,J)
      RETURN
      END
      SUBROUTINE READATA

C THIS SUBROUTINE READS IN ALL THE DATA

      COMMON CTIME, CDTIME, NTSTEP, NROW, NCOL, NVAR, NRUN
     1 , DX(6,7), DY(6,7), KBC(6,7), XNORM(150)
```

```
     2 , FK(6,7), B(6,7), Q(6,7), HI(6,7), PHI(6,7), HT(6,7), Z(6,7)
     3 , FKMEAN(6,7), ZMEAN(6,7), QMEAN(6,7), HIMEAN(6,7), PHIMEN(6,7)
     4 , FKVAR(6,7,15), ZVAR(6,7,15)
     * , QVAR(20,15), HIVAR(20,15), PHIVAR(20,15)
     5 , FKMIN(6,7),ZMIN(6,7),QMIN(6,7),HMIN(6,7),PHIMIN(6,7)
     6 , FKMED(6,7), ZMED(6,7), QMED(6,7), HMED(6,7), PHIMED(6,7)
     7 , FKMAX(6,7), ZMAX(6,7), QMAX(6,7), HMAX(6,7), PHIMAX(6,7)
     8 , FKLOW(6,7), ZLOW(6,7), QLOW(6,7), HLOW(6,7), PHILOW(6,7)
     9 , FKUP(6,7), ZUP(6,7), QUP(6,7), HUP(6,7), PHIUP(6,7)
     0 , HDT(20,38,10), TWIDE(20,15,10), CVHAT(20,15,10)
     1 , STUDENT, NBETA
     4 , X(5,15)
     6 , TIM(10)
     7 , GTMIN(7), GTMED(7), GTMAX(7), GBMIN(7), GBMED(7), GBMAX(7)
     8 , GRMIN(7), GRMED(7), GRMAX(7), GLMIN(7), GLMED(7), GLMAX(7)
     9 , GTOP(7,11), GBOT(7,11), GRITE(6,11), GLEFT(6,11)
      READ(5,1) NBETA, NTSTEP, NVAR, NRUN, NROW, NCOL, STUDENT
    1 FORMAT(6I10,F10.3)
      WRITE(6,6)
    6 FORMAT(1H0,*  STUDENT                          NTSTEP                        NVAR
     9                NRUN                         NROW                     NCOL
     9    NBETA*)
      WRITE(6,7) STUDENT, NTSTEP, NVAR, NRUN, NROW, NCOL, NBETA
    7 FORMAT(1H ,F7.4,6I20)
      READ(5,2) (TIM(I),I=1,NTSTEP)
    2 FORMAT(16F5.1)
      WRITE(6,8)
    8 FORMAT(1H0,55X,*LENGTH OF TIME STEP(DAYS)*)
      WRITE(6,9) ((I,TIM(I)),I=1,NTSTEP)
    9 FORMAT(1H ,65X,I3,F6.2)
      READ(5,3) ((KBC(I,J),J=1,NCOL),I=1,NROW)
    3 FORMAT(7I10)
      WRITE(6,10)
   10 FORMAT(1H0,60X,*KBC*)
      WRITE(6,11) ((KBC(I,J),J=1,NCOL),I=1,NROW)
   11 FORMAT(1H ,7I18)
      READ(5,4) (( FKMIN(I,J),J=1,NCOL),I=1,NROW)
      READ(5,4) (( FKMED(I,J),J=1,NCOL),I=1,NROW)
      READ(5,4) (( FKMAX(I,J),J=1,NCOL),I=1,NROW)
      READ(5,4) ((   HMIN(I,J),J=1,NCOL),I=1,NROW)
      READ(5,4) ((   HMED(I,J),J=1,NCOL),I=1,NROW)
      READ(5,4) ((   HMAX(I,J),J=1,NCOL),I=1,NROW)
      READ(5,4) ((   ZMIN(I,J),J=1,NCOL),I=1,NROW)
      READ(5,4) ((   ZMED(I,J),J=1,NCOL),I=1,NROW)
      READ(5,4) ((   ZMAX(I,J),J=1,NCOL),I=1,NROW)
      READ(5,4) ((   QMIN(I,J),J=1,NCOL),I=1,NROW)
      READ(5,4) ((   QMED(I,J),J=1,NCOL),I=1,NROW)
      READ(5,4) ((   QMAX(I,J),J=1,NCOL),I=1,NROW)
      READ(5,5) ((PHIMIN(I,J),J=1,NCOL),I=1,NROW)
      READ(5,5) ((PHIMED(I,J),J=1,NCOL),I=1,NROW)
      READ(5,5) ((PHIMAX(I,J),J=1,NCOL),I=1,NROW)
    5 FORMAT(7F10.9)
      READ(5,4) ((     DX(I,J),J=1,NCOL),I=1,NROW)
      READ(5,4) ((     DY(I,J),J=1,NCOL),I=1,NROW)
      READ(5,4) (GTMIN(J),J=1,NCOL)
      READ(5,4) (GTMED(J),J=1,NCOL)
      READ(5,4) (GTMAX(J),J=1,NCOL)
      READ(5,4) (GBMIN(J),J=1,NCOL)
      READ(5,4) (GBMED(J),J=1,NCOL)
      READ(5,4) (GBMAX(J),J=1,NCOL)
      READ(5,4) (GRMIN(I),I=1,NROW)
      READ(5,4) (GRMED(I),I=1,NROW)
      READ(5,4) (GRMAX(I),I=1,NROW)
      READ(5,4) (GLMIN(I),I=1,NROW)
      READ(5,4) (GLMED(I),I=1,NROW)
      READ(5,4) (GLMAX(I),I=1,NROW)
    4 FORMAT(7F10.2)
      WRITE(6,12)
   12 FORMAT(1H0,60X,* FKMIN*)
      WRITE(6,13) (( FKMIN(I,J),J=1,NCOL),I=1,NROW)
   13 FORMAT(1H ,7F18.5)
```

```
      WRITE(6,14)
   14 FORMAT(1H0,60X,* FKMED*)
      WRITE(6,13) (( FKMED(I,J),J=1,NCOL),I=1,NROW)
      WRITE(6,15)
   15 FORMAT(1H0,60X,* FKMAX*)
      WRITE(6,13) (( FKMAX(I,J),J=1,NCOL),I=1,NROW)
      WRITE(6,16)
   16 FORMAT(1H0,60X,*  HMIN*)
      WRITE(6,13) ((  HMIN(I,J),J=1,NCOL),I=1,NROW)
      WRITE(6,17)
   17 FORMAT(1H0,60X,*  HMED*)
      WRITE(6,13) ((  HMED(I,J),J=1,NCOL),I=1,NROW)
      WRITE(6,18)
   18 FORMAT(1H0,60X,*  HMAX*)
      WRITE(6,13) ((  HMAX(I,J),J=1,NCOL),I=1,NROW)
      WRITE(6,19)
   19 FORMAT(1H0,60X,*  ZMIN*)
      WRITE(6,13) ((  ZMIN(I,J),J=1,NCOL),I=1,NROW)
      WRITE(6,20)
   20 FORMAT(1H0,60X,*  ZMED*)
      WRITE(6,13) ((  ZMED(I,J),J=1,NCOL),I=1,NROW)
      WRITE(6,21)
   21 FORMAT(1H0,60X,*  ZMAX*)
      WRITE(6,13) ((  ZMAX(I,J),J=1,NCOL),I=1,NROW)
      WRITE(6,22)
   22 FORMAT(1H0,60X,*QMIN  *)
      WRITE(6,13) ((  QMIN(I,J),J=1,NCOL),I=1,NROW)
      WRITE(6,23)
   23 FORMAT(1H0,60X,*  QMED*)
      WRITE(6,13) ((  QMED(I,J),J=1,NCOL),I=1,NROW)
      WRITE(6,24)
   24 FORMAT(1H0,60X,*  QMAX*)
      WRITE(6,13) ((  QMAX(I,J),J=1,NCOL),I=1,NROW)
      WRITE(6,25)
   25 FORMAT(1H0,60X,*    DX*)
      WRITE(6,13) ((    DX(I,J),J=1,NCOL),I=1,NROW)
      WRITE(6,26)
   26 FORMAT(1H0,60X,*    DY*)
      WRITE(6,13) ((    DY(I,J),J=1,NCOL),I=1,NROW)
      WRITE(6,27)
   27 FORMAT(1H0,60X,*PHIMIN*)
      WRITE(6,13) ((PHIMIN(I,J),J=1,NCOL),I=1,NROW)
      WRITE(6,28)
   28 FORMAT(1H0,60X,*PHIMED*)
      WRITE(6,13) ((PHIMED(I,J),J=1,NCOL),I=1,NROW)
      WRITE(6,29)
   29 FORMAT(1H0,60X,*PHIMAX*)
      WRITE(6,13) ((PHIMAX(I,J),J=1,NCOL),I=1,NROW)
      WRITE(6,30)
   30 FORMAT(1H0,60X,*GTMIN*)
      WRITE(6,13) (GTMIN(J),J=1,NCOL)
      WRITE(6,31)
   31 FORMAT(1H0,60X,*GTMED*)
      WRITE(6,13) (GTMED(J),J=1,NCOL)
      WRITE(6,32)
   32 FORMAT(1H0,60X,*GTMAX*)
      WRITE(6,13) (GTMAX(J),J=1,NCOL)
      WRITE(6,33)
   33 FORMAT(1H0,60X,*GBMIN*)
      WRITE(6,13) (GBMIN(J),J=1,NCOL)
      WRITE(6,34)
   34 FORMAT(1H0,60X,*GBMED*)
      WRITE(6,13) (GBMED(J),J=1,NCOL)
      WRITE(6,35)
   35 FORMAT(1H0,60X,*GBMAX*)
      WRITE(6,13) (GBMAX(J),J=1,NCOL)
      WRITE(6,36)
   36 FORMAT(1H0,60X,*GRMIN*)
      WRITE(6,13) (GRMIN(I),I=1,NROW)
      WRITE(6,37)
   37 FORMAT(1H0,60X,*GRMED*)
      WRITE(6,13) (GRMED(I),I=1,NROW)
      WRITE(6,38)
```

```
   38 FORMAT(1H0,60X,*GRMAX*)
      WRITE(6,13) (GRMAX(I),I=1,NROW)
      WRITE(6,39)
   39 FORMAT(1H0,60X,*GLMIN*)
      WRITE(6,13) (GLMIN(I),I=1,NROW)
      WRITE(6,40)
   40 FORMAT(1H0,60X,*GLMED*)
      WRITE(6,13) (GLMED(I),I=1,NROW)
      WRITE(6,41)
   41 FORMAT(1H0,60X,*GLMAX*)
      WRITE(6,13) (GLMAX(I),I=1,NROW)
      RETURN
      END
      SUBROUTINE RANDOM
C THIS SUBROUTINE GENERATES RANDOM VALUES OF PHIMEN, QMEAN, HIMEAN, FKMEAN, ZMEA
C -N, GTOP, GBOT, GRITE, GLEFT FROM TRIANGULAR DISTRIBUTIONS DEFINED BY THEIR MI
C -NIMUM, MEDIAN, MAXIMUM VALUES.
C HIVAR, PHIVAR, QVAR, FKVAR, ZVAR, FROM UNIFORM DISTRIBUTIONS DEFINED BY THEIR
C MINMUM AND MAXIMUM VALUES.
C UNIFORM RANDOM NOS. ON INTERVAL (0,1) ARE OBTAINED FROM LIBRARY SUBROUTINE BY
C CALLING RANF(0). SUBROUTINE RANSET SPECIFIES THE SEED FOR THIS RANDOM NUMBER
C GENERATE.
C VALUES ARE ALSO ASSIGNED TO PHILOW, PHIUP,HLOW, HUP, QLOW, QUP, ZLOW, ZUP, FKL
C -OW, FKUP.
      COMMON CTIME, CDTIME, NTSTEP, NROW, NCOL, NVAR, NRUN
     1 , DX(6,7), DY(6,7), KBC(6,7), XNORM(150)
     2 , FK(6,7), B(6,7), Q(6,7), HI(6,7), PHI(6,7), HT(6,7), Z(6,7)
     3 , FKMEAN(6,7), ZMEAN(6,7), QMEAN(6,7), HIMEAN(6,7), PHIMEN(6,7)
     4 , FKVAR(6,7,15), ZVAR(6,7,15)
     * , QVAR(20,15), HIVAR(20,15), PHIVAR(20,15)
     5 , FKMIN(6,7),ZMIN(6,7),QMIN(6,7),HMIN(6,7),PHIMIN(6,7)
     6 , FKMED(6,7), ZMED(6,7), QMED(6,7), HMED(6,7), PHIMED(6,7)
     7 , FKMAX(6,7), ZMAX(6,7), QMAX(6,7), HMAX(6,7), PHIMAX(6,7)
     8 , FKLOW(6,7), ZLOW(6,7), QLOW(6,7), HLOW(6,7), PHILOW(6,7)
     9 , FKUP(6,7), ZUP(6,7), QUP(6,7), HUP(6,7), PHIUP(6,7)
     0 , HDT(20,38,10), TWIDE(20,15,10), CVHAT(20,15,10)
     1 , STUDENT, NBETA
     4 , X(5,15)
     6 , TIM(10)
     7 , GTMIN(7), GTMED(7), GTMAX(7), GBMIN(7), GBMED(7), GBMAX(7)
     8 , GRMIN(7), GRMED(7), GRMAX(7), GLMIN(7), GLMED(7), GLMAX(7)
     9 , GTOP(7,11), GBOT(7,11), GRITE(6,11), GLEFT(6,11)
      CALL RANSET(354871083)
      NR = NROW - 1
      NC = NCOL - 1
      DO 100 I = 1,NROW
      DO 100 J = 1,NCOL
      PHILOW(I,J) = 0.004
      PHIUP(I,J) = 0.02
      HLOW(I,J) = 1.0
      HUP(I,J) = 10.0
      QLOW(I,J) = 0.0
      ZLOW(I,J) = 1.0
      ZUP(I,J) = 10.0
      FKLOW(I,J) = 1.0
  100 CONTINUE
      DO 777 I = 2,NR
      DO 777 J = 2,NC
      T = PHIMIN(I,J)
      U = PHIMED(I,J)
      V = PHIMAX(I,J)
      RN = RANF(0)
      G = T + SQRT(RN*(U - T)*(V - T))
      IF(G.LE.U) GO TO 2
      G = V - SQRT(RN*(U - V)*(V-T)-V*U+T*U-V*T+V*V)
    2 PHIMEN(I,J) = G
      T = HMIN(I,J)
      U = HMED(I,J)
      V = HMAX(I,J)
      RN = RANF(0)
      G = T + SQRT(RN*(U - T)*(V - T))
      IF(G.LE.U) GO TO 3
```

```
      G = V - SQRT(RN*(U - V)*(V-T)-V*U+T*U-V*T+V*V)
    3 HIMEAN(I,J) = G
      T = QMIN(I,J)
      U = QMED(I,J)
      V = QMAX(I,J)
      RN = RANF(0)
      G = T + SQRT(RN*(U - T)*(V - T))
      IF(G.LE.U) GO TO 4
      G = V - SQRT(RN*(U - V)*(V-T)-V*U+T*U-V*T+V*V)
    4 QMEAN(I,J) = G
      QUP(I,J) = QMEAN(I,J)/5.0
  777 CONTINUE
      DO 888 I = 1,NROW
      DO 888 J = 1,NCOL
      T = ZMIN(I,J)
      U = ZMED(I,J)
      V = ZMAX(I,J)
      RN = RANF(0)
      G = T + SQRT(RN*(U - T)*(V - T))
      IF(G.LE.U) GO TO 5
      G = V - SQRT(RN*(U - V)*(V-T)-V*U+T*U-V*T+V*V)
    5 ZMEAN(I,J) = G
      T = FKMIN(I,J)
      U = FKMED(I,J)
      V = FKMAX(I,J)
      RN = RANF(0)
      G = T + SQRT(RN*(U - T)*(V - T))
      IF(G.LE.U) GO TO 1
      G = V - SQRT(RN*(U - V)*(V-T)-V*U+T*U-V*T+V*V)
    1 FKMEAN(I,J) = G
      FKUP(I,J) = FKMEAN(I,J)/5.0
  888 CONTINUE
      WRITE(6,52)
   52 FORMAT(1H0,*             HMEAN                    FKMEAN                    ZME
     9AN                QMEAN                    PHIMEAN*)
      DO 53 I = 2,NR
      DO 53 J = 2,NC
      WRITE(6,54) HIMEAN(I,J),FKMEAN(I,J),ZMEAN(I,J),QMEAN(I,J),PHIMEN(I
     9,J)
   53 CONTINUE
   54 FORMAT(1H ,5F20.6)
      DO 10 J = 1,NCOL
      T = GTMIN(J)
      U = GTMED(J)
      V = GTMAX(J)
      RN = RANF(0)
      G = T + SQRT(RN*(U - T)*(V - T))
      IF(G.LE.U) GO TO 11
      G = V - SQRT(RN*(U - V)*(V-T)-V*U+T*U-V*T+V*V)
   11 GTOP(J,1) = G
      T = GBMIN(J)
      U = GBMED(J)
      V = GBMAX(J)
      RN = RANF(0)
      G = T + SQRT(RN*(U - T)*(V - T))
      IF(G.LE.U) GO TO 10
      G = V - SQRT(RN*(U - V)*(V-T)-V*U+T*U-V*T+V*V)
   10 GBOT(J,1) = G
      DO 12 I = 1,NROW
      T = GRMIN(I)
      U = GRMED(I)
      V = GRMAX(I)
      RN = RANF(0)
      G = T + SQRT(RN*(U - T)*(V - T))
      IF(G.LE.U) GO TO 13
      G = V - SQRT(RN*(U - V)*(V-T)-V*U+T*U-V*T+V*V)
   13 GRITE(I,1) = G
      T = GLMIN(I)
      U = GLMED(I)
      V = GLMAX(I)
      RN = RANF(0)
      G = T + SQRT(RN*(U - T)*(V - T))
```

```
      IF(G.LE.U) GO TO 12
      G = V - SQRT(RN*(U - V)*(V-T)-V*U+T*U-V*T+V*V)
   12 GLEFT(I,1) = G
      DO 15 K = 1,NTSTEP
      DO 16 J = 1,NCOL
      T = GTOP(J,K) - 5.0
      U = GTOP(J,K)
      V = GTOP(J,K) + 5.0
      RN = RANF(0)
      G = T + SQRT(RN*(U - T)*(V - T))
      IF(G.LE.U) GO TO 17
      G = V - SQRT(RN*(U - V)*(V-T)-V*U+T*U-V*T+V*V)
   17 GTOP(J,K+1) = G
      T = GBOT(J,K) - 5.0
      U = GBOT(J,K)
      V = GBOT(J,K) + 5.0
      RN = RANF(0)
      G = T + SQRT(RN*(U - T)*(V - T))
      IF(G.LE.U) GO TO 16
      G = V - SQRT(RN*(U - V)*(V-T)-V*U+T*U-V*T+V*V)
   16 GBOT(J,K+1) = G
      DO 18 I = 1,NROW
      T = GRITE(I,K) - 5.0
      U = GRITE(I,K)
      V = GRITE(I,K) + 5.0
      RN = RANF(0)
      G = T + SQRT(RN*(U - T)*(V - T))
      IF(G.LE.U) GO TO 19
      G = V - SQRT(RN*(U - V)*(V-T)-V*U+T*U-V*T+V*V)
   19 GRITE(I,K+1) = G
      T = GLEFT(I,K) - 5.0
      U = GLEFT(I,K)
      V = GLEFT(I,K) + 5.0
      RN = RANF(0)
      G = T + SQRT(RN*(U - T)*(V - T))
      IF(G.LE.U) GO TO 18
      G = V - SQRT(RN*(U - V)*(V-T)-V*U+T*U-V*T+V*V)
   18 GLEFT(I,K+1) = G
   15 CONTINUE
      IGRID = 0
      DO 40 I = 2,NR
      DO 40 J = 2,NC
      IGRID = IGRID + 1
      DO 40 K = 1,NVAR
      RN = RANF(0)
      HIVAR(IGRID,K) = RN*(HUP       - HLOW) + HLOW
      RN = RANF(0)
      PHIVAR(IGRID,K) = RN*(PHIUP       -PHILOW) + PHILOW
      RN = RANF(0)
   40 QVAR(IGRID,K) = RN*(QUP(I,J) - QLOW) + QLOW
      DO 41 I = 1,NROW
      DO 41 J = 1,NCOL
      DO 41 K = 1,NVAR
      RN = RANF(0)
      FKVAR(I,J,K) = RN*(FKUP(I,J) - FKLOW) + FKLOW
      RN = RANF(0)
   41 ZVAR(I,J,K) = RN*(ZUP - ZLOW) + ZLOW
      RETURN
      END
      SUBROUTINE MINMAX(X,NRUN,NROW,NCOL,T,IVAR,IP,NVAR,NTSTEP,C)
      DIMENSION F1(38),X(IP,IRUN,NTSTEP),T(IP,NVAR,NTSTEP),C(IP,NVAR,NTSTEP)
     9TEP)
      NR = NROW - 1
      NC = NCOL - 1
      RUN = NRUN
      IGRID = 0
      DO 200 IROW = 2,NR
      DO 200 ICOL = 2,NC
      IGRID = IGRID + 1
      DO 200 J = 1,NTSTEP
      S = 0.0
      SS = 0.0
```

```
      DO 100 I = 1,NRUN
      G = X(IGRID,I,J)
      S = S + G
      SS = SS + G*G
  100 F1(I) = G(IGRID,I,J)
      B1 = AMAX1(F1(1),F1(2),F1(3),F1(4),F1(5),F1(6),F1(7),F1(8),F1(9),
     9F1(10),F1(11),F1(12),F1(13),F1(14),F1(15),F1(16),F1(17),F1(18),F1(19),F1(
     919),F1(20),F1(21),F1(22),F1(23),F1(24),F1(25),F1(26),F1(27),F1(28)
     9,F1(29),F1(30),F1(31),F1(32),F1(33),F1(34),F1(35),F1(36),F1(37),
     9F1(38))
      S1 = AMIN1(F1(1),F1(2),F1(3),F1(4),F1(5),F1(6),F1(7),F1(8),F1(9),
     9F1(10),F1(11),F1(12),F1(13),F1(14),F1(15),F1(16),F1(17),F1(18),F1(19),F1(
     919),F1(20),F1(21),F1(22),F1(23),F1(24),F1(25),F1(26),F1(27),F1(28)
     9,F1(29),F1(30),F1(31),F1(32),F1(33),F1(34),F1(35),F1(36),F1(37),
     9F1(38))
      T(IGRID,IVAR,J) = B1 - S1
      SDHAT = SQRT((SS - S*S/RUN)/(RUN - 1.0))
      SMEAN = S/RUN
      CVHAT(IGRID,IVAR,J) = SDHAT/SMEAN
  200 CONTINUE
      RETURN
      END
      SUBROUTINE RBSOLV (C,N,M,V)
C THIS SUBROUTINE SOLVES MATRIX EQUATION SET UP BY SUBROUTINE AMATRIX TO GIVE
C VALUES OF HEAD AT END OF TIME STEP.SOLUTION TECHIQUE IS GAUSS ELIMINATION.
      DIMENSION C(N,M),V(N)
      K=M-1
      LR=M-3
      DO 60 L=1,LR
      IM=LR+1-L
      DO 60 I=1,IM
      DO 50 J=2,K
   50 C(L,J-1)=C(L,J)
   60 C(L,K)=0.0
      C(1,K)=C(1,M)
      C(1,M)=0.0
      DO 70 I=2,K
       J=I-2
   70 C(N-J,M)=0.0
      C(N,K)=0.0
      IM=N-1
      LP=K
      DO 220 I=1,IM
      NPIV=I
      LS=I+1
      LL=LP-1
      DO 100 L=LS,LP
      IF(ABS(C(L,1)).GT.ABS(C(NPIV,1))) NPIV=L
  100 CONTINUE
      IF(NPIV.LE.I) 130,110
  110 DO 120 J=1,M
      TEMP=C(I,J)
      C(I,J)=C(NPIV,J)
  120 C(NPIV,J)=TEMP
      TEMP=V(I)
      V(I)=V(NPIV)
      V(NPIV)=TEMP
  130 V(I)=V(I)/C(I,1)
      DO 140 J=2,M
  140 C(I,J)=C(I,J)/C(I,1)
      IF(I.GE.N-LR) LL=LL+1
      DO 190 L=LS,LL
      TEMP=C(L,1)
       V(L)=V(L)-TEMP*V(I)
      DO 160 J=2,M
  160 C(L,J-1)=C(L,J)-TEMP*C(I,J)
      IF(L.EQ.LS) 170,180
  170 C(L,M)=0.0
      GO TO 190
  180 C(L,K)=0.0                                                        RBSOL180
  190 CONTINUE
```

```
      IF(I.GE.N-LR) GO TO 220
      TEMP=C(LP,1)
       V(LP)=V(LP)-TEMP*V(I)
      C(LP,1) =-TEMP*C(I,2)
      DO 210 J=2,K
  210 C(LP,J)=C(LP,J)-TEMP*C(I,J+1)
      IF(LP.LT.N) LP=LP+1
  220 CONTINUE
      V(N)=V(N)/C(N,1)
      LP=2
      DO 250 I=1,IM
      L=N-I
      DO 240 J=2,LP
      LL=L+J-1
  240 V(L) = V(L) - C(L,J)*V(LL)
      IF(LP.LT.M-1) LP=LP+1
  250 CONTINUE
      RETURN
      END
      SUBROUTINE AMATRIX(IRAN , RHS , A)
      DIMENSION A(20,6), RHS(20), RKH(6,7)
      COMMON CTIME, CDTIME, NTSTEP, NROW, NCOL, NVAR, NRUN
     1 , DX(6,7), DY(6,7), KBC(6,7), XNORM(150)
     2 , FK(6,7), B(6,7), Q(6,7), HI(6,7), PHI(6,7), HT(6,7), Z(6,7)
     3 , FKMEAN(6,7), ZMEAN(6,7), QMEAN(6,7), HIMEAN(6,7), PHIMEN(6,7)
     4 , FKVAR(6,7,15), ZVAR(6,7,15)
     * , QVAR(20,15), HIVAR(20,15), PHIVAR(20,15)
     5 , FKMIN(6,7),ZMIN(6,7),QMIN(6,7),HMIN(6,7),PHIMIN(6,7)
     6 , FKMED(6,7), ZMED(6,7), QMED(6,7), HMED(6,7), PHIMED(6,7)
     7 , FKMAX(6,7), ZMAX(6,7), QMAX(6,7), HMAX(6,7), PHIMAX(6,7)
     8 , FKLOW(6,7), ZLOW(6,7), QLOW(6,7), HLOW(6,7), PHILOW(6,7)
     9 , FKUP(6,7), ZUP(6,7), QUP(6,7), HUP(6,7), PHIUP(6,7)
     0 , HDT(20,38,10), TWIDE(20,15,10), CVHAT(20,15,10)
     1 , STUDENT, NBETA
     4 , X(5,15)
     6 , TIM(10)
     7 , GTMIN(7), GTMED(7), GTMAX(7), GBMIN(7), GBMED(7), GBMAX(7)
     8 , GRMIN(7), GRMED(7), GRMAX(7), GLMIN(7), GLMED(7), GLMAX(7)
     9 , GTOP(7,11), GBOT(7,11), GRITE(6,11), GLEFT(6,11)
C
C*************************************************************************
C
C SETS UP MATRIX FOR CONSTANT GRADIENTS BOUNDARY CONDITIONS ONLY
C
C*************************************************************************
C
      NRA= (NROW - 2)*(NCOL - 2)
      NCA = NROW
      NC = NCOL - 1
      NR = NROW - 1
      NCA1 = 1
      NCA2 = NROW - 3
      NCA3 = NROW - 2
      NCA4 = NROW - 1
      NCA5 = NROW
      DO 90 J =1,NCA
      DO 90 I =1,NRA
   90 A(I,J) = 0.0
      DO 91 I =1,NROW
      DO 91 J =1,NCOL
   91 RKH(I,J) = 0.0
      IA = 0
      DO 100 J = 2,NC
      DO 100 I = 2,NR
      K1 = KBC(I,J+1)
      K2 = KBC(I-1,J)
      K3 = KBC(I,J-1)
      K4 = KBC(I+1,J)
      IA = IA + 1
      PX= 2.0*FK(I,J)*DY(I,J)*B(I,J)/DX(I,J)
      PY= 2.0*FK(I,J)*DX(I,J)*B(I,J)/DY(I,J)
      P1 = 2.0*FK(I,J+1)*DY(I,J+1)*B(I,J+1)/DX(I,J+1)
      P1 = (PX*P1)/(PX + P1)
```

```
      P2 = 2.0*FK(I-1,J)*DX(I-1,J)*B(I-1,J)/DY(I-1,J)
      P2 = (PY*P2)/(PY + P2)
      P3 = 2.0*FK(I,J-1)*DY(I,J-1)*B(I,J-1)/DX(I,J-1)
      P3 = (PX*P3)/(PX + P3)
      P4 = 2.0*FK(I+1,J)*DX(I+1,J)*B(I+1,J)/DY(I+1,J)
      P4 = (PY*P4)/(PY + P4)
      PRHS = PHI(I,J)*DX(I,J)*DY(I,J)/DT
      GO TO (1,2,3),K1
    1 RKH(I,J)=P1*HT(I,J+1)
      A(IA,NCA3) = P1
      GO TO 2
    3 A(IA,NCA5)=P1
    2 GO TO (4,5,6),K2
    4 RKH(I,J)=P2*HT(I-1,J) + RKH(I,J)
      A(IA,NCA3) = P2+ A(IA,NCA3)
      GO TO 5
    6 A(IA,NCA2)=P2
    5 GO TO (7,8,9),K3
    7 RKH(I,J)=P3*HT(I,J-1) + RKH(I,J)
      A(IA,NCA3) = P3+ A(IA,NCA3)
      GO TO 8
    9 A(IA,NCA1)=P3
    8 GO TO (10,11,12),K4
   10 RKH(I,J)=P4*HT(I+1,J) + RKH(I,J)
      A(IA,NCA3) = P4+ A(IA,NCA3)
      GO TO 11
   12 A(IA,NCA4)=P4
   11 A(IA,NCA3) = -A(IA,NCA3) -(A(IA,NCA1) + A(IA,NCA2) + A(IA,NCA4) +
     9 A(IA,NCA5) + PRHS)
      RHS(IA) = Q(I,J) - PRHS*HT(I,J) - RKH(I,J)
  100 CONTINUE
      RETURN
      END
      SUBROUTINE POLY(Y,X,NVAR,NBETA,NBETA1,STUDENT,IGRID,N1,IP,NTSTEP)
C THIS SUBROUTINE ESTS. COEFFS., COMPUTES CONF. INT. ON EACH REGRESSION COEFFS.
C TESTS HYPOS.  (A0,A1,A2,A4,A5)=0, (A1,A2,A4,A5)=0
C FOLLOWING VARIABLES HAVE TO BE DIMENSIONED,
C SIGMA2(NTSTEP) = EST. OF VARIANCE OF COND. DIST.
C CONFUP(NTSTEP ,NBETA) = UPPER CONF. INT. ON REGRESSION COEFF.
C CONFLOW(NTSTEP,NBETA) = LOWER CONF. INT. ON REGRESSION COEFF.
C FSTAT(NTSTEP) = F-STAT FOR TESTS OFHYPOS
C E(NBETA,NBETA1) = MATRIX (XX) AND (XY) FOR EACH TIME STEP IN NORMAL EQUATIONS
C (XX)*(BETA) = (XY)
C RHS(NTSTEP,NBETA) = VECTOR (XY) FROM NORMAL EQUS. FOR EACH TINE STEP. (RHS) IS
C SAME AS COLS. N2-NBETA1 OF (E)
C RBET(NTSTEP) = (BETA)*(XY)
C RGAMMA(NTSTEP) = EST. OF REGRESSION COEFFS. IN REDUCED MODEL
C BETA(NTSTEP,NBETA) = MAX. LIKELIHOOD EST. OF REGRESSION COEFFS.
C F(2,I+NTSTEP) = MATRIX OF COEFFS. IN REDUCED MODEL
      DIMENSION X(N1,NVAR), Y(NTSTEP, NVAR)
     1 , SIGMA2(10), CONFUP(10,6), CONFLOW(10,6), FSTAT(10), E(6,16)
     2 , RHS(10,6), RBET(10), RGAMMA(10), BETA(10,6), F(2,12)
C NHI = NO. OF REGRESSION COEFFS. ASSOCIATED WITH ERROR IN INITISL HEAD IN
C REGRESSION MODEL. FOR REGRESSION EQUS. IN THIS PROGRAM , THE COEFF. IS (A3) IN
C BOTH MODELS. THEREFORE NHI = 4
      NHI = 4
      N2 = NBETA + 1
C SET ELEMENTS OF (E) EQUAL ZERO
      DO 2 K = 1,NBETA
      DO 2 L = 1,NBETA1
    2 E(K,L) = 0.0
C LOOPS 20,21 DEFINE ELEMENTS OF (E) FOR COLS. 1-NBETA
      E(1,1) = NVAR
      DO 20 L = 2,NBETA
      V = 0.0
      DO 3 K = 1,NVAR
      LL = L - 1
    3 V = V + X(LL,K)
      E(1,L) = V
   20 E(L,1) = V
      DO 21 L = 2,NBETA
      DO 21 M = L,NBETA
```

```
      V = 0.0
      DO 4 K = 1,NVAR
    4 V = V + X(L-1,K)*X(M-1,K)
      E(L,M) = V
   21 E(M,L) = V
C LOOPS 6,7 DEFINE ELEMENTS OF MATRIX (E) ROR COLS. N2-NBETA1
      DO 6 L = N2,NBETA1
      DO 6 K = 1,NVAR
      LL = L - NBETA
    6 E(1,L) = E(1,L) + Y(LL,K)
      DO 7 J = N2,NBETA1
      JJ = J - NBETA
      DO 7 L = 2,NBETA
      DO 7 K = 1,NVAR
    7 E(L,J) = E(L,J) + Y(JJ,K)*X(L-1,K)
C LOOP 8 STORES COLS. N2-NBETA1 OF (E) IN (RHS)
      DO 8 K = N2,NBETA1
      KK = K - NBETA
      DO 8 L = 1,NBETA
    8 RHS(KK,L) = E(L,K)
C DEFINE COLS. 1-2 OF (F) TO BE USED TO TEST HYPO THAT (A1,A2,A4,A5)=0
      E44 = E(NHI,NHI)
      F(1,1) = E(1,1)
      F(1,2) = E(1,NHI)
      F(2,1) = F(1,2)
      F(2,2) = E44
C LOOP 30 DEFINES COLS. 3-NBET OF (F)
      NBET = NTSTEP + 2
      II = 1
      DO 30 I = 1,2
      DO 31 J = 3,NBET
      JJ = J + NHI
   31 F(I,J) = E(II,JJ)
   30 II = II + NHI - 1
C LIBRARY SUBROUTIME MATRIX CONPUTES INVERSE OF COLS. 1-NBETA OF (E), AND SOLUTI
C -ON CORRESPONDING TO EACH VECTOR OF (E) STORED IN COLS. N2-NBETA1. THIS
C INVERSE AND SOLUTION VECTORS ARE RETURNED IN (E). DETERMINANT OF MATRIX BEING
C INVERTED IS RETURNED IN DET.
      CALL MATRIX(10,NBETA,NBETA1,2,E,NBETA ,DET)
C LOOP 9 STORES SOLUTION VECTORS IN (BETA). THESE SOLUTION VECTORS ARE MAX. LIKE
C -LIHOOD ESTS. OF REGRESSION COEFFS. . EACH VECTOR CORRESPONDS TO ONE TIME STEP
      DO 9 I = N2,NBETA1
      II = I - NBETA
      DO 9 L = 1,NBETA
    9 BETA(II,L) = E(L,I)
      AN = NVAR - NBETA
C LOOP 10 COMPUTES,
C                     (YY)= SUM OF SQUARES OF DEPENDENT VARIABLE.
C                     (RBET)=(BETA)*(XY)
C                     (SIGMA2)= EST. OF VARIANCE OF COND. DIST.
      DO 10 I = 1,NTSTEP
      YY = 0.0
      DO 11 J = 1,NVAR
   11 YY = YY + Y(I,J)**2
      RBETA = 0.0
      DO 12 J = 1,NBETA
   12 RBETA = RBETA + RHS(I,J)*BETA(I,J)
      SIGMA2(I) = (YY - RBETA)/AN
      RBET(I) = RBETA
   10 CONTINUE
C LOOP 13 COMPUTES CONF. INT. ON (BETA).
C IF INT. CONTAINS ZERO , IT IS EQUIVALENT TO ACCEPTING HYPO THAT REGRESSION COE
C -FFS. IS EQUAL TO ZERO.
      DO 13 I = 1,NTSTEP
      DO 13 J = 1,NBETA
      TROOT = STUDENT*SQRT(E(J,J)*SIGMA2(I))
      CONFUP(I,J) = BETA(I,J) + TROOT
   13 CONFLOW(I,J) = BETA(I,J) - TROOT
      DO 14 I = 1,NTSTEP
      WRITE(6,16)
   16 FORMAT(1H )
      DO 14 J = 1,NBETA
```

```
      WRITE(6,15) J,CONFLOW(I,J),BETA(I,J),CONFUP(I,J)
   14 CONTINUE
   15 FORMAT(1H ,20X,I5,3F25.10)
C
C TO TEST HYPO (A0,A1,A2,A4,A5)=0
C
      WRITE(6,19)
   19 FORMAT(1H )
      AN1 = N1
      DO 17 I = 1,NTSTEP
      RGAMMA(I)=RHS(I,NHI)**2/E44
      FSTAT(I) = ((RBET(I) - RGAMMA(I))/AN1)/SIGMA2(I)
      ERROR = SIGMA2(I)*AN
   17 WRITE(6,18) I, FSTAT(I), ERROR
   18 FORMAT(1H ,I10,*    FSTAT = *,F20.4,*    ERROR = *,F20.4)
C TO TEST HYPO (A1,A2,A4,A5)=0
      WRITE(6,34)
   34 FORMAT(1H )
      AN1 = NBETA - 2
      CALL MATRIX(10,2,NBET,2,F,2,DET)
      DO 32 I = 1,NTSTEP
      RGAMMA(I) = F(1,I+2)*RHS(I+NBETA,1) + F(2,I+2)*RHS(I+NBETA,NHI)
      FSTAT(I) = ((RBET(I) - RGAMMA(I))/AN1)/SIGMA2(I)
   32 WRITE(6,33) I, FSTAT(I)
   33 FORMAT(1H ,I10,*   FSTAT = *,F20.4)
      RETURN
      END
```

Typed and Reproduced by
TYPE-INK Fort Collins