COLLECTIVE ADJUSTMENT OF THE PARAMETERS OF THE MATHEMATICAL MODEL OF A LARGE AQUIFER

by

Robert Edmund Lovell

A Thesis Submitted to the Faculty of the

DEPARTMENT OF SYSTEMS ENGINEERING

In Partial Fulfillment of the Requirements
For the Degree of

MASTER OF SCIENCE

In the Graduate College

THE UNIVERSITY OF ARIZONA

1 9 7 1

STATEMENT BY AUTHOR

This thesis has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Robert Edmund Lowell

APPROVAL BY THESIS DIRECTOR

This thesis has been approved on the date shown below:

Lucien Duckstein
LUCIEN DUCKSTEIN
Professor of Systems Engineering

11 may 1971
Date

PREFACE

In a January 1970 memorandum to Dr. Lucien Duckstein, Dr. Chester C. Kisiel, Professor of Hydrology and Water Resources, outlined several problems in hydrology which he felt could benefit from a systems engineering approach. Among them was the development of a "computer program for optimal estimation of aquifer constants in a distributed parameter model" of an aquifer.

The method then being used by the investigators of the Tucson basin aquifer was subjective trial and error. However, a number of papers had recently appeared offering various methods of optimizing the constants, or "calibrating the model" as the procedure is often called. The questions then became: Are any of these methods appropriate in whole or in part? Can they be placed in a form suitable for processing by a digital computer? Were assumptions made in these techniques which would not be applicable in the present case? Can subjective information be introduced somehow into an automated procedure for estimating these constants, or must some different concept be developed? The attempt to answer these questions led to development of the adjustment procedure described here.

I would like to express my thanks here to both Dr. Duckstein and Dr. Kisiel for the time spent in consultation with me in weekly

meetings as the experimental work evolved, and for the opportunity to use them as sounding boards for my ideas as they took form. Both of them were extremely helpful in supplying suggestions and reference material. Also, Joseph S. Gates, Graduate Research Associate at The University of Arizona, was extremely helpful in filling in the voids in my background in hydrologic and geologic matters, and in playing the role of "the hydrologist/geologist team" in my work with the Tucson aquifer. The computer time for my studies was partially supported by a matching grant (B-007-ARIZ) from the Office of Water Resources Research, U. S. Department of the Interior, for studies on the worth of hydrologic data.

TABLE OF CONTENTS

TABLE OF CONTENTS--Continued

LIST OF TABLES

LIST OF ILLUSTRATIONS

ABSTRACT

The problem of evaluating the parameters of the mathematical model of an unconfined aquifer is examined with a view toward development of automated or computer-aided methods. A formulation is presented in which subjective confidence ranges for each of the model parameters are quantified and entered into an objective function as linear penalty functions. Parameters are then adjusted by a procedure which seeks to reduce the model error to acceptable limits.

A digital computer model of the Tucson basin aquifer is adapted and used to illustrate the concepts and demonstrate the method.

CHAPTER 1

INTRODUCTION

The requirements for development of new water resources, and for extension and better management of existing resources, have led many investigators in the past two decades to large scale mathematical modeling of the underground storage and flow of water. Manipulation of these models of underground reservoirs (aquifers) was originally done with electric analog techniques. As digital computers with larger memories and higher computational speeds became available, a number of investigators began adapting the model for digital treatment (Pinder and Bredehoeft 1968). The hybrid computer, combining some of the more useful features of both analog and digital computers, also came into limited use in some applications (Vemuri and Karplus 1969).

However, the accuracy and value of the results obtainable by manipulation of a model are heavily dependent on the quality of the parameters, rather than on the computational methods used for model manipulation. One of the key problems in modeling, then, is the calibration of the model; that is, the adjustment of the parameters (or constants) of the model to obtain a satisfactory match between predicted and historical aquifer response based on available knowledge of hydraulic

stresses imposed on the aquifer in the past. Since the data available for modeling are generally incomplete, and often inaccurate, the inverse problem is not unique. It is possible to adjust the parameters in an infinity of arbitrary ways to obtain the desired fit. But if the model is to serve well as a tool for future aquifer management studies under a great variety of projected future stresses, the parameters ultimately selected must reflect as nearly as feasible the underlying physical structure of the aquifer.

This, then, is the classical system identification problem but with insufficient and inaccurate data, in a situation where ability to gather additional data is severely limited by physical and economic considerations. Observations of input/output relationships for the most part are limited to observing inputs and outputs beyond the control of the investigator. Inputs are prescribed by nature and by users of water. Outputs of water are inputs of stress to the aquifer as far as the modeler is concerned. The output is aquifer response, i.e., change in water level in the case of unconfined aquifers, and change in piezometric head in confined (artesian) aquifers. In general, the gathering of additional data is costly, and experimental manipulation of the physical system itself is infeasible, except at a few test wells which may be under the experimenter's control, and may disturb the system to an unknown extent.

If parameter adjustment is to be done inexpensively, and in a way which is more likely to represent the true underlying structure, something must be done to the original data other than adjust it by trial and error, by randomization, or by using a purely mathematical approach limited only by the absolute physical bounds assignable to the parameters. This requires input of additional information other than that typically inserted into the model.

A potentially valuable source of information usually overlooked is the "educated discretion" of the hydrologist/geologist team which is making the original assignment of parameters to the model. The team reviews geological maps, drilling logs, pumping records, aquifer test data, and all the other original sources of information which may prove useful. During this process they can be expected to develop a reasonably good insight into the acceptable range of perturbation that should be permitted in any automated program for calibrating the particular model. This insight will accumulate from several sources, e.g., understanding of measurement accuracies, insight gained during field collection of data, knowledge of the extent of interpolation and extrapolation of data, the presumed geology of the region, and experience developed in modeling other aquifers.

This collective "educated discretion" or "subjective confidence", if suitably quantified, could play a valuable role in the calibration

program. By incorporating this subjective information into the adjustment procedure, hopefully one can guide the adjustment program in a more natural way toward the actual physical representation. Also these inputs could provide more realistic (narrower) bounds for perturbations of individual parameters. The objective then would be to adjust parameters so that the desired degree of fit between the predicted and historical response surfaces could be met, and so that some measure of confidence-weighted parameter perturbations could be minimized.

This paper is concerned with the general problem of calibrating large unconfined aquifers, and the existing model of the Tucson basin (Arizona) aquifer is used to illustrate the results obtained. Specifically, we are seeking an approach to automated adjustment of aquifer parameters, which gives suitable recognition to subjective inputs from the investigators preparing the model data, as a possible alternative to gathering additional information in the field, and which gives a framework for this input.

Chapter 2 covers the general background of modeling unconfined aquifers for digital computers and some of the parameter adjustment schemes that have been proposed in the past. It also reviews the series of experiments with the model of the Tucson aquifer which were used to develop insight into the problem and to validate the concepts which were to be incorporated in the adjustment algorithm. Those interested only in

considerations leading directly to the final adjustment procedures could profitably omit Chapter 2 and proceed directly to Chapter 3.

Chapter 3 presents the underlying thoughts and methods used in the TRINSUCOSE (TRouble INdex/SUbjective COnfidence/SEnsitivity) algorithm, which is the end product of this study. Chapter 4 reviews the results obtained in application of the algorithm to a portion of the Tucson aquifer. Chapter 5 points out the general conclusions reached, limitations noted, and some possible improvements in and extensions of the method to obtain better results or to treat a wider variety of problems. The more important FORTRAN programs used are presented in Appendix A.

Notation

The notations used for parameters and variables of the model and their fundamental dimensions are as follows:

A, the area of a node (ℓ^2);

h, the underground water level or hydraulic head (ℓ);

HE, the underground water level at the End of the calibration period (ℓ);

HS, the underground water level at the Start of the calibration period (ℓ);

L, the distance between two nodes (ℓ);

Q, the volumetric rate of water discharge, whether by pumping or by imputed pumping at an artificial boundary (ℓ^3/t);

R, the negative of the volumetric rate of the boundary and infiltration recharge of the aquifer (ℓ^3/t);

S, storage coefficient (dimensionless);

T, transmissivity (ℓ^2/t)

t, time (t)

W, the maximum width of the underground flow path between two nodes, i.e., the length of the common boundary between the nodal areas (ℓ).

The notation in the FORTRAN programs of Appendix A conforms in general with the above. Special variable and parameter names in FORTRAN are either self-explanatory or are identified in the individual subprograms by COMMENT cards.

Subscript notation is as follows:

B, the node or nodal area currently under consideration (i.e., the Base node), identified in the computer programs by the FORTRAN subscript (I,J) denoting the row and column, respectively, of the node in the array of square nodes in the Tucson aquifer;

k, a nodal area sharing a common boundary with node B, with corresponding FORTRAN subscripts as follows:

(I-1,J), the node to the north,

(I,J+1), the node to the east,

(I+1,J), the node to the south,

(I,J-1), the node to the west;

kB, the region between the center of node k and the center of node B;

n, any node in the model.

A summation over k indicates a summation over those nodes having a common boundary with B. A summation over n indicates a summation over all nodes in the model.

CHAPTER 2

APPROACH TO THE PROBLEM

The equations of ground water flow have evolved over an extended period. One of the earliest contributions seems to have been that of Darcy (1856) who reported his experiments on water flow through sands of various permeabilities. The relationships originally identified empirically by Darcy have now come to be known as Darcy's Law, and relate velocity components of water flow to the product of pressure gradient and permeability (Hubbert 1969). The common present method of segmenting an aquifer into polygonal nodes for numerical analysis is due to Thiessen (1911). A brief history of the development and computer manipulation of aquifer models, together with extensive historical references, is given in Pinder and Bredehoeft (1968, pp. 1069-1071).

The physical/mathematical relationships for aquifer modeling are now well established both theoretically and experimentally. The remaining problems are the selections of model structures for particular aquifers, assignment and adjustment of parameters, and methods of solution. This paper is particularly concerned with the methods for adjustment of parameters.

Allison (1967, p. 11) summarizes the process of model-building for an aquifer as follows:

a. The area is broken down into a large number of sub-areas, usually by the Thiessen method of polygon construction, around control points referred to as nodes.

b. Initial estimates, based primarily on geologic data, are made on the storage characteristics of each nodal area, and transmissibility between nodes. These estimates provide a basis for writing a set of equations, one for each nodal area within the basin, which form the first rough mathematical model of the basin.

c. Historic hydrologic data are analyzed and estimates made of net recharge to or withdrawals from each nodal area for a period of years, which will be referred to hereafter as the base, or calibration, period. Also, hydrographs of ground water surface fluctuations, based on historic observations of water levels in wells, are prepared for each polygon.

d. An attempt is then made to simulate the behavior of the basin, starting from known values of water surface elevation at each node, feeding in historic values of recharge or discharge, and having the computer solve the initial set of differential equations for each successive year of the base period. The hydrographs of water surface elevations computed for each node are then compared with the observed hydrographs. System parameters are then adjusted where necessary to obtain more faithful reproduction of actual behavior of the basin, and the process is repeated until an acceptable degree of reproduction is obtained.

e. The model is then accepted as built, and is ready for use in predicting the behavior of the basin under any proposed plan for future management.

This procedure appears to be representative of the methods usually employed in modeling large aquifers. Step d is critical since no criteria for adjustment are suggested except to adjust in such a way that mathematical predictions will be better, without much consideration of

physical reasons for the adjustment. The human trial/computer error iteration routine is costly in time and computer expense. Pliska (1968) reports that this method typically requires about 50 calibration cycles over a period of possibly several months. Allison (1967, p. 94) considers a reasonable range of calibration runs to be 20 to 200. Pinder and Bredehoeft (1968, p. 1088) required 37 runs to calibrate a model with 1145 20-foot square nodes extending over a region about 800 x 1000 feet. Gates (1971, p. 3) reports 38 runs were needed to bring the Tucson aquifer model to its present state; that is, to reduce the maximum nodal error from 190.6 feet to 28.6 feet, and the average (absolute) error from 24.4 feet to 5.3 feet.

The concept investigated in this paper is to modify steps c and d by inclusion of gathering of subjective information (if possible, at the time of compilation of the original data), and to modify step d to replace trial and error adjustment with a suitably automated technique for bringing the model into calibration with a minimum of disturbance to the subjective information.

A note of caution must be given on step e: the adjustment of the parameters of a model, even using a technique that represents the "true" underlying physical situation for the calibrated region, does not insure that the model will perform well over "any proposed plan for future management." There may be unidentified geologic structure that cannot

be found by the range of water level variations incurred in the past. For example, continued withdrawal of water from the aquifer below the level where data have previously been available would produce behavior not encountered at the time of calibration and, therefore, not incorporated in the adjustment program. The transmissivity, for example has been shown to be a function of the depth of the water remaining in the aquifer, but is often assumed to be constant in the model. Clearly, any severe change in the depth of water in such a model could invalidate the results. In the extreme case, if the pumping is such that certain regions of the aquifer no longer contain water, the effective boundary will have changed and the original model will not apply at all. Whether or not a model calibrated over a limited range of water levels can be considered valid beyond that range can probably best be judged by those providing the subjective inputs, and should not be assumed a priori.

A discrete parameter model results from the first three steps of Allison's procedure, and the parameters associated with any one node are related by the following differential/difference equation:

$$\sum_{k} (h_k - h_B) T_{kB} W_{kB} / L_{kB} - Q_B - R_B = A_B S_B (dh_B/dt). \qquad \text{Eq. (1)}$$

This equation has a general applicability to isotropic unconfined aquifers and in some special cases can be applied to anisotropic aquifers. The preferred method of solution of families of the above equations when

written around polygonal nodal areas obtained by the Thiessen method seems to be a Gauss-Seidel procedure (Tyson and Weber 1964).

Pinder and Bredehoeft (1968) show the detailed development of similar finite difference equations from Darcy's Law in the more general case of anisotropic aquifers, and extend the discussion to show how it can be specialized to isotropic aquifers. Their model is made up of square nodes of uniform area in order to permit solution by the more efficient alternating direction implicit technique (Bruce et al., 1953).

The Model of the Tucson Basin Aquifer

The digital model of the Tucson aquifer was constructed in 1968-1969 for manipulation originally by the Gauss-Seidel method, and later was modified for manipulation by the alternating direction implicit technique (Moench 1969 and Gates 1970). For model purposes the aquifer was considered to be isotropic, unconfined, with S and T parameters considered constant with h and t. It consisted of 1890 square nodes each 160 acres in area. Since W_{kB} and L_{kB} were both equal to 2640 feet between every adjacent pair of nodes, these quantities cancel each other in Equation (1).

The data were compiled using information from the U. S. Geological Survey analog model, and from the University of Arizona Department of Agricultural Engineering, as well as from the conventional sources described by Allison. At the time the investigations described here began

in early 1970 the calibration had just begun. This calibration was pursued independently of the present work by Gates (1971) using the trial and error techniques.

The base, or calibration, period covered 19 years--from 1947 through 1966. The base period was further broken down into 4 sub-periods as follows:

1947-1952	5 years
1952-1957	5 years
1957-1962	5 years
1962-1966	4 years.

As will be seen below, the data for the shorter time periods proved to be insufficiently independent for actual calibration use.

The point of departure for this study was this uncalibrated model together with subjective confidence information obtained personally from Gates in the form of coarse confidence contours sketched on computer maps of the various parameter arrays. The course of the investigation is presented immediately after a brief review of the present state of automated procedures.

Present State of Automated Adjustment Procedures

The literature reports several approaches to automatic or computer-assisted parameter adjustment both from theoretical and practical viewpoints. Many of these were examined to determine the extent to

which they might be applicable to the present problem. Each of them, however, depended on some assumptions that made them of little use in the large-scale rather generalized case being investigated here.

Haimes, Perrine, and Wismer (1968) treated a special aquifer identification problem by decomposition and multi-level optimization. To make the model analytically tractable they postulated a cluster of wells centered in an infinite aquifer. Since all wells were assumed to have been brought into production and to have been pumped in a uniform way, it was possible to decompose the aquifer into pie-shaped wedges, each containing one well, so that an assumption of no flow across the boundaries could be made. While this procedure may be of theoretical interest, it is not applicable in the present instance since the wells are not clustered in a central region, were not all brought into production at the same time, and have not been pumped in a uniform way.

Coats, Dempsey and Henderson (1968) described a procedure in which upper and lower bounds are placed on each parameter. The parameters were subsequently permitted to take on new values within those bounds by a randomization process. Here, however, we take the position that adjustments should be made in amounts and directions that can be shown to be of positive contribution to calibration, rather than in random patterns. For this reason the method of Coats et al. was not considered further.

Pliska (1968) reported an automated "one run" optimization scheme for large unconfined aquifers. But his method also used randomization techniques and was not considered further.

Vemuri and Karplus (1969) reported the development of a hybrid computation system which permitted near real-time interaction between the hydrologist and the model. The interaction was related largely to adjustment of boundaries of the model rather than obtaining subjective inputs that would be helpful in identifying the parameters at established nodes. For reasons given in Chapter 3 the interaction concept given by Vemuri and Karplus was rejected in this analysis, although the Maximum Principle concept in their report is of theoretical interest.

Emsellem and de Marsily (1969) reported a new method for solving the inverse problem of finding transmissivities by assuming that there are no physical reasons for abrupt discontinuities in their spatial pattern. They optimize against this criterion. They assume, however, that the water level measurements represent a steady state, i.e., that the inflows and outflows of water in each elemental area of the aquifer are in balance so that there is no change in amount of water stored in any element. These special conditions are not met in the general case, and clearly fail for the Tucson aquifer where a pattern of net withdrawals has persisted for many years.

While this is not an exhaustive review of the automated techniques that have been presented, it is representative of recent patterns of thinking about these matters. Since little of direct help was found in these techniques, this investigation was carried in the directions described in the next section.

The Course of the Investigation

There were many steps taken in the investigations leading to the algorithm of Chapter 3 and the results of Chapter 4. These steps are outlined below in the order in which work on them was initiated. Several of these steps gave important insight into model behavior, extent of adjustment necessary, degree of data dependence, and possible approaches to adjustment. Steps which are not described elsewhere in the paper are described immediately following the outline.

1. Development of error criteria and programs.
2. Change in node size.
3. Static water balance tests (see below and Chapter 3).
4. Withholding of data.
5. Imputed value tests.
6. The average transmissivity question.
7. Development of Trouble Index/Node Rating scheme (see Chapter 3).
8. Model truncation.

9. Selection of time step size for model manipulation.
10. Error propagation studies.
11. Preparation and treatment of confidence data (see Chapter 3).
12. The perturbation scheme (see Chapter 3).
13. Parameter adjustment (see Chapters 3 and 4).

Development of Error Criteria and Programs

There appears to be no suitable way to measure parameter error except indirectly by measuring the difference between water levels predicted by model and the measured (or estimated) levels at each node. The typical statistics generated from the spatial distribution of water level errors are mean error, mean absolute deviation, standard deviation, peak errors, and error range. It does not seem appropriate to select any of these as "the" measure to represent the state of calibration of the model. In some way all have to be considered. The mean error can be correlated directly with the static water balance error, while the others are measures of the errors in dynamic predictions. Certainly large peak errors embedded in a field of small deviations would not be suitable. Alternatively, reduction of peak errors while retaining a large root-mean-square or mean absolute deviation would be inappropriate. Chapter 4 shows, however, that the decision on these points need not be made in advance. The TRINSUCOSE algorithm leaves this question open, and automated

repetitions of the algorithm promise that all reasonable error criteria can be met, subjective and physical constraints permitting.

In addition to the measures of "output" error given above, it was felt desirable to have a simple measure of error trends over the areal extent of the model. This would permit study of error propagation through the model, and permit visual pattern recognition of areas not responding or overresponding to the calibration procedures. Accordingly, a program for fitting a least squared error plane to the data was generated. This program also includes (1) individual computation of the deviations from this place at each node, (2) measures of the center of error, (3) computation of the magnitude and direction of the tilt of the plane, and (4) preparation of an error output array. A subprogram was created for this purpose and was used in the error propagation studies described below.

Change in Node Size

Noting that the data from approximately 500 wells had been made to cover 1890 nodes, it seemed that much of the data in the model were interpolations and extrapolations. Therefore, it was felt that a change from 1/4-square mile nodes to 1-square mile nodes, and a corresponding reduction to 512 nodes would provide a model in better match with the available data. This reduction was accomplished by a special computer program which summed, averaged, etc., the various parameters for the nodes as appropriate to the type of parameter. All indications were that

this reduction did not create any anomalies in the patterns, except at some "inside" corners where nodes that were originally on boundaries became interior nodes on reduction. Appropriate reassignment of the boundary recharge data was made at these nodes to correct this.

This node reduction scheme was also used by Gates (1971) in making a more manageable model for his investigations.

Static Water Balance Tests

Analyses of static water balance proved to be of sufficient value that the concept and technique are discussed in Chapter 3. For the present discussion note simply that these tests involve the algebraic summing of all water assigned as inputs and outputs of the aquifer, and the comparison of that number with the amount of water released from storage. Any serious difference is a measure of modeling or data error, quite independent of dynamic considerations of water level contours and ability of individual nodes to deliver or receive the water attributed to them.

The disturbing results of the water balance tests for the 512 1-square mile node model (where positive percentages represent water deliveries to users in excess of that released from storage) were as follows:

For the shorter time periods--

1947-1952	-19.0%
1952-1957	2.3%
1957-1962	16.6%
1962-1966	88.4%

For the full 19-year base period--

1947-1966	18.1%, when based on the Δh data,
1947-1966	- 0.3%, when based on raw h data.

Clearly the time trend makes the short time interval data highly suspect. These results had an important role in the attempt to withhold data described next.

Withholding of Data Attempt

As noted earlier, the 19-year base period was broken down into 4 sub-periods. The original concept of this study was to withhold part of this data, work out an adjustment procedure using the remaining data, and then attempt to verify the calibration obtained by comparison with a computer prediction of the information withheld. Specifically, the data for the 1957-1962 period were to be blended with those of the 1962-1966 period to create a 9-year period, 1957-1966. This, then, would permit a test of the model's ability to interpolate to the 1962 water levels. Further, it was hoped that 1970 or 1971 data would become available which would permit a test of the model's ability to extrapolate water levels.

However, the static water balance tests reported above indicated a serious lack of consistency in the data, which made this whole

withholding process highly suspect. Further, Gates (1970, pp. 10-12) reports that all of the recharge data are averaged over the 19-year period and distributed pro-rata to the shorter intervals. Also, he notes that much of the pumping information is based on averages over some shorter periods, not necessarily coincident with the sub-periods over which the data were distributed. Also noted were the poor closures between maps of cumulative change in water level with maps of the water level. This caused some abrupt discontinuities in the last model period as shown above in the static water balance tests, and below when data were reviewed for computation of new boundary conditions for the truncated model. It seemed that the data in the various time periods were not sufficiently independent to permit withholding of data on a time basis. This approach was therefore abandoned.

Imputed Value Tests

The next step in familiarization with the model was to conduct a series of imputed value tests. Here the concept was to make a node by node analysis of data coherence by computing imputed values of each of the parameters in turn, each time assuming all other parameters were correct. Rearrangements of Equation (1) were used for this purpose. The tests were conducted with the 512 node model using the full 19-year base period, and for the 5-year period (1952-1957) which exhibited the best balance on the static tests.

All the results were unusable. Of the 512 nodes in the 19-year test, the imputed values of storage coefficient (S) of 357 were clearly infeasible (either negative or in excess of 1.0). Only 89 were within the likely range of 0.0 to 0.25. Of these 89, only 28 were within the 0.0 to 0.25 range for both the 19-year and 5-year runs.

Similar results were encountered when transmissivities (T) were imputed. Of the 512 nodes in the 19-year test, at least 180 had physically impossible negative values. Only 89 had values falling between 50% and 200% of the originally assigned values. Of these, only 46 coincided with values falling in the same range on the 5-year test.

The tests of the imputed net pumping/recharge (QR) also showed substantial deviations from results that a model close to calibration would be expected to produce. Over 200 of the nodes showed sign changes between net outflow and net inflow. Fewer than 60 of the nodes had imputed values of QR in the range of 50% to 200% of the originally assigned values.

In retrospect, it appears that some of these results could have been anticipated. For example, those nodes which exhibit very little change in water level would naturally be very insensitive to changes in storage coefficient. Similarly, those nodes which have an underground inflow roughly equal to outflow would exhibit little sensitivity to changes in transmissivity.

Clearly, if some "one-dimensional" quantity is to be developed for each node as a measure of data coherence, some other technique is required. A good answer seems to be the Trouble Index developed in Chapter 3.

Computation of the Average Transmissivity

In the modeling process, values of transmissivity are assigned to each node in the model. In computing the flow through the underground node boundaries it has been customary to compute the effective transmissivity T_{iB} as the arithmetic mean of the assigned T_i and T_B (Moench 1969, Gates 1970, and McMillan 1966, pp. 51-56).

Warren and Price (1965) and Bouwer (1969) have investigated the relative suitabilities of various methods of finding effective transmissivity. They find that theoretical considerations lead to use of the arithmetic mean only for parallel flow through regions of different permeabilities. The harmonic mean is most suitable for series flow through consecutive regions of different transmissivities. In randomly generated transmissivities they suggest that the intermediate results given by the geometric mean are more suitable.

Here the transmissivities are not randomly generated, but are assigned on the basis of the best data available. Further, they enter the model as parameters for strictly series flow. The alternating direction iteration program was therefore changed to use the harmonic mean

($T_{iB} = 2.0T_iT_B/(T_i+T_B)$) with the understanding that the $T_{iB} = 0.0$ when both T_i and T_B are zero.

Model Truncation

After information on Trouble Index and Node Rating became available, a second step in reduction of the model to more manageable size was possible. Here, the concept was to find a reasonably trouble-free double line of nodes across the model. If such a band could be found, then the model could be partitioned along that line for parameter adjustment purposes since the adjustment procedure would make only minor changes at those nodes that were well behaved.

The Trouble Index/Node Rating maps did identify such a band cutting across the southern part of the aquifer. The decision was then made to work only with the smaller southern portion containing 166 nodes. Accordingly, the boundary discharge to the north was computed at the time-averaged hydraulic gradient and attributed to the pumping/recharge (Q and R) parameters of the new boundary in the usual way.

Now, working with 166 nodes in a system with reasonably well-behaved boundaries, the problem was sufficiently large to exhibit a substantial amount of coupled behavior, but not too large to preclude rather extensive computer experimentation.

Selection of Time Step Size for Model Manipulation

The FORTRAN program which had been developed for model manipulation by the alternating direction technique (Gates 1970), was rewritten into a FORTRAN subprogram (SUBROUTINE ALDIR) using the same basic iterative technique but altering the methods of computing internodal transmissivities, as reported above, and to permit selection of time step size by calling arguments in the main program. Two other techniques were incorporated; one to permit selection of uniform time steps, and the other to permit time step size to be increased in a geometric manner.

In manipulating the model of the Tucson basin aquifer it had been customary to double the time step for each succeeding iteration. Tests were conducted to find whether a uniform time step would be preferable. These tests showed a tendency for the results to converge to a stable value (in the same sense as in the following paragraph) slightly faster when uniform time steps were used. Therefore, uniform time steps were used for subsequent iterative tests.

Iterations were made with the 19-year period divided into 1, 2, 4, 8, and 16 uniform time intervals. The standard deviations of the differences in prediction of water levels in these tests were as follows:

between 1 and 2 steps - 11.46 feet,

between 2 and 4 steps - 13.19 feet,

between 4 and 8 steps - .61 feet,

between 8 and 16 steps - .14 feet.

Since there seemed to be very little improvement in using 16 time steps over 8, the figure of 8 time steps was adopted for all future model iterations.

Error Propagation Studies

In order to gain familiarity with model behavior and to develop an idea of how errors propagate or migrate through the system, a series of special tests was conducted. These tests all followed the same general pattern. The model was assumed correct as given except for the water levels at the end of the base period. The model then predicted a number of future water levels for as long as 45 years in some cases. These predicted surfaces were then taken to be the "true" surfaces. Following this, artificial errors were introduced in each of the parameters, one at a time, at one node only, and the model was then caused to compute a new series of response surfaces. The differences between the "true" response surfaces and the response surfaces generated with the artificial error present measured the propagation or migration of the error.

In this manner nine different test series were conducted. The results can be summarized briefly by saying that, when "pseudo-errors" were introduced in the parameters, whether at interior or boundary nodes, the peak errors in the predicted response surfaces remained initially at the nodes where the parameter errors were introduced. In the cases of errors in storage coefficient (S), transmissivity (T), end water level (HE),

and a mixed error of S and T such that S/T remained constant, the "node of peak error" at all times was the node at which the error was introduced.

In the cases of error in beginning water levels (HS) and pumping/recharge (Q+R) the same was true in the early years. However, in the later years of the test (seven to 15 years) the node of peak error would sometimes be found a few nodes away from the node in which the parameter errors were originally introduced, but by the time this peak migration took place, the error was substantially attentuated, usually to less than 10% of its original magnitude.

More details of these test results are contained in Appendix B.

In Chapter 3 it will be seen that a decomposition or decoupling concept is used to permit the first look at the coherence of data at individual nodes. This decoupling is based on the concept that the most important part of the model prediction errors will occur at the node where the parameters are in error. The experiments described above lend important support to this notion and permit the decoupling used for that purpose over a rather wide range of conditions.

This chapter has described the various thoughts and directions taken during the investigation. Those considerations that were retained for use in the final adjustment procedures are presented in the next chapter.

CHAPTER 3

DEVELOPMENT OF THE ADJUSTMENT CONCEPT

Since each of the six quantities associated with each node of the model is subject to uncertainty, it is desirable to develop some knowledge of how these uncertainties influence the design of an adjustment procedure. The following questions were posed:

1. Are the data mutually consistent or coherent? That is, do they represent a physically realizable set? If not, can measures be developed which will express the amount of incoherence in a useful way? Can trouble spots be exposed before operation of the model for prediction of future water surfaces begins?
2. Can a priori confidence ranges be obtained from the hydrologist/geologist team and be inserted into the model calibration procedure in an effective way?
3. How can model sensitivites be expressed so that parameters can be adjusted in some proportion to their contribution to calibration?
4. How can all these things be combined so that parameters can be adjusted in concert rather than sequentially, and

in proportion to the calibration needs? Can this be done in such a way that a measure of the ratio of performance improvement to total weighted perturbation is nearly optimal?

The balance of this chapter addresses itself to these questions.

The Nature of the Uncertainties in the Model

There are two fundamental steps in the initial modeling of an aquifer. Both are subject to uncertainties, but of different kinds.

First, and perhaps most important, is the selection of a basic model structure. This selection will depend heavily on the hydrology/geology of the region and such considerations as the following will enter: Is the aquifer confined (artesian), unconfined, or some spatial mixture of the two? What are the effective boundaries? Are they subject to change with changing water level? Is the aquifer isotropic? Is it leaky? Is it coupled to a surface water system? Can the parameters be considered constant in time and with respect to water levels? What time intervals are appropriate? What node sizes and layouts are suitable?

Clearly an incorrect answer to any of these questions could lead to a very bad model, even though it might somehow be calibrated.

The second concern is the series of uncertainties in the gathering and preparation of data for the selected model structure. It is these uncertainties which are considered in this paper. Here we assume that the

correct model has been selected, and the uncertainties are only in the data assigned to the selected nodes.

In particular, the experimental work here has been done with an unconfined isotropic aquifer with boundaries assumed known. The storage coefficients and transmissivities are assumed to be constant over time and independent of water level. The aquifer is assumed to be non-leaky and coupled to surface water only in ways that can be represented adequately by equivalent pumping and recharge at specific nodes. The model of this aquifer is truncated at seven locations, and the underground inflow-outflow at these artificial boundaries is assumed to be accounted for also by equivalent pumping/recharge rates at the boundary nodes.

The present model also uses square nodes of uniform size arranged in a grid which makes it amenable to computation by the alternating direction technique previously mentioned, in lieu of the accelerated Gauss-Seidel technique often applied to heterogeneous arrangements of polygonal nodes. However, the parameter adjustment methods described herein are equally applicable to either type of node layout.

Data Coherence, Trouble Index, and Node Rating

Once the original data are assigned to the model, it seems appropriate to make a quality inspection. Often gross errors are made in preparation of the data or decisions are made in selecting parameters which, on second look, will appear invalid. The computer can be of

substantial help in this process. First, computer printout of the data as they actually have been prepared for computer use will automatically highlight gross errors. Secondly, without actually going into interative manipulations of the model some measures of data consistency or coherence can be made. Suitable display of these data for review by the data collection team is an important second step. Clearly a program to optimize the parameters could work with data containing gross errors, but better results could surely be obtained if as many of these as feasible were removed before parameter adjustment begins. Essenwanger (1970) makes a very strong case for such review by the compilers of data in the case of preparation of meteorological data for extensive analysis. In hydrology also "a good concept of quality assurance must precede any data analysis to avoid distortion and bias of results by erroneous records."

Two ways were devised for displaying data for quality inspection and are presented in the subsections immediately following. The first is a static look at the overall model and the second a decomposed node by node review of data coherence. In the second method, important information is developed which predicts which nodes are going to perform poorly in the fully coupled iterative routines for prediction of aquifer response. Further, it provides weighting factors for use in perturbation of parameters for model calibration.

Static Coherence of the Overall System

The most obvious measure of coherence of the model data is a simple water volume balance analysis for the overall model. This simply asks the question, "Does the amount of water leaving the model, minus the amount of water entering the model, equal the decrease in the amount of water stored in the aquifer model?" This question may be asked for any of the time periods for which model data are available. Any amount of imbalance detected by this analysis must be attributable to modeling or data error.

This simple water volume balance test takes no account of the dynamics of water movement within the aquifer, and, therefore, asks nothing about whether the shape of the underground water "surface" is feasible in any dynamic sense. Note, however, that the shape of the water surface is used in one method of determining transmissivities in an aquifer (Emsellem and de Marsily 1969).

The program for performing this analysis makes the following computation:

$$\text{ERROR} = \sum_{n} (Q_n + R_n - A_n S_n (HE_n - HS_n)) . \qquad \text{Eq. (2)}$$

The FORTRAN subprogram used for the present investigation of static coherence is SUBROUTINE STATIC presented in Appendix A.

While this computation gives some idea of static coherence of data, it provides no insight into how the individual parameters might be adjusted. Its usefulness is limited to its ability to identify such extremely poor water balance as to cast doubt on the validity of the entire model.

For reasonably close water balance--say "net outflow" and "decrease in storage" agree within 20%--the investigator may simply proceed to the next steps. The static water balance error is directly correlatable with the mean error provided by the error analysis sections of the iterative routines, and is automatically adjusted by the TRINSUCOSE algorithm as the other measures of modeling error are reduced.

Coherence of Data for Individual Nodes

A "first look" at the coherence of data for each node was also desirable. But a water volume balance at individual nodes required some method for inclusion of underground water flow between nodes. According to the discrete equation, Equation (1), the instantaneous rate of underground flow through the polygon boundaries of a node is proportional to the hydraulic gradient between the nodes--a gradient customarily measured in feet of water. In the model the gradient parameters were available at the beginning and end of the time periods to be used for model calibration. For many of the nodes the gradients at the beginning and end

of the calibration period were approximately the same. This led to the idea of assuming linearity of gradient with time over the calibration period for those that are not the same.

Certainly the assumption of time linearity of hydraulic gradients between nodes is subject to some objection. Clearly, the pumping history at one node could be such that the water level was drawn down to the present level early in the calibration period and remained in that vicinity for the balance of the time. Similarly, a well which was developed later in the calibration period would disturb the water level in a way that would cause the average gradient to be understated. But, as a way of getting an initial feel for the coherence of data at individual nodes, it does provide a useful tool.

Based on the linearity assumption, a FORTRAN subroutine entitled RATING (reproduced in Appendix A) was developed to produce the following information:

1. Imbalance. It produces an array of water imbalance information at each node measured in feet per year of excess water (+) or shortage of water (-). This quantity, when multiplied by the area of the node, produces the volumetric rate of the excess or shortage of water in that node.
2. Trouble Index. A numerical "trouble index" was needed to quantify by magnitude and direction the amount of incoherence

at each node. Also, this Trouble Index becomes useful later as a weighting factor in parameter perturbation.

3. Node Rating. It was also found convenient to rank order the trouble indices for easier visual identification on a computer output map of data incoherence. In this way the less coherent nodes could be called to the attention of the hydrologist for reconsideration before they are admitted to the automatic adjustment process. Accordingly, a Node Rating section was included in the subprogram.

The Trouble Index, as explained below, is roughly proportional to the cause of data incoherence, while the Imbalance is an estimate of the effect of that incoherence.

One property desired in the Trouble Index was that in some way it be proportional to a combination of parameter changes needed to correct the data incoherence. Many of the parameters are not useful for this purpose. For example, the adjustment of storage coefficient (S) would have no meaning at a node where the water levels at the beginning and end of the time period were the same. In that case any storage coefficient would be equally valid. In other cases physically impossible storage coefficients would be required, i.e., less than 0 or greater than 1.

To adjust the transmissivity (T) alone to compensate for the imbalance is often equally unprofitable. Some nodes cannot be made

coherent by adjustment of transmissivity alone because physically unrealizable negative values would be required.

Adjustments of pumping (Q) and recharge (R) alone face similar objections. The Q and R are really input/output to the model in that sense and are not really model parameters. Further, to cure the imbalance, it might be necessary to attribute pumping to a node where there is no pumping, or to attribute recharge to a node where recharge is not occurring.

The remaining alternative is to look to imputed changes in water levels. An imputed change in HE_B (the water level in the base node at the end of the calibration period) has many desirable characteristics and seems to overcome most of the previous objections. It has the following properties:

1. It always exists. If the water imbalance shows more water is required in a node, then a negative "Trouble Index" will always bring in the required amount. This negative Trouble Index--as an adjustment to the height of water surface at the end--would always impute larger water inflow (or less outflow) at all boundaries, and also impute more water withdrawn from (or less placed in) storage. A positive Trouble Index always decreases imputed water inflows (or increases outflows) and imputes less withdrawal from storage (or greater increase

in storage). Hence, a Trouble Index can always be found which will "correct" the imbalance.

2. The Trouble Index includes the effect of the transmissivity assigned to the node by taking on proportionately higher values in nodes of low transmissivity in order to correct an equivalent amount of imbalance.
3. It includes the effect of storage coefficient since it imputes a change in Δh.
4. Experience with the Tucson aquifer has shown that it always takes on values in a working range.
5. The values are unique.

Note that a perturbation of the starting height of water surface (HS), everything else remaining constant, would not have the same properties since the changes in underground flow of water would be in the opposite sense of the water delivery from storage. A Trouble Index based on HS, therefore, could be relatively insensitive and unstable.

Clearly the perturbation of both the beginning and ending water levels together would obscure the contribution of S since Δh would remain unchanged.

Figure 1 on the following page illustrates the physical meaning of the parameters and Trouble Index for a typical node.

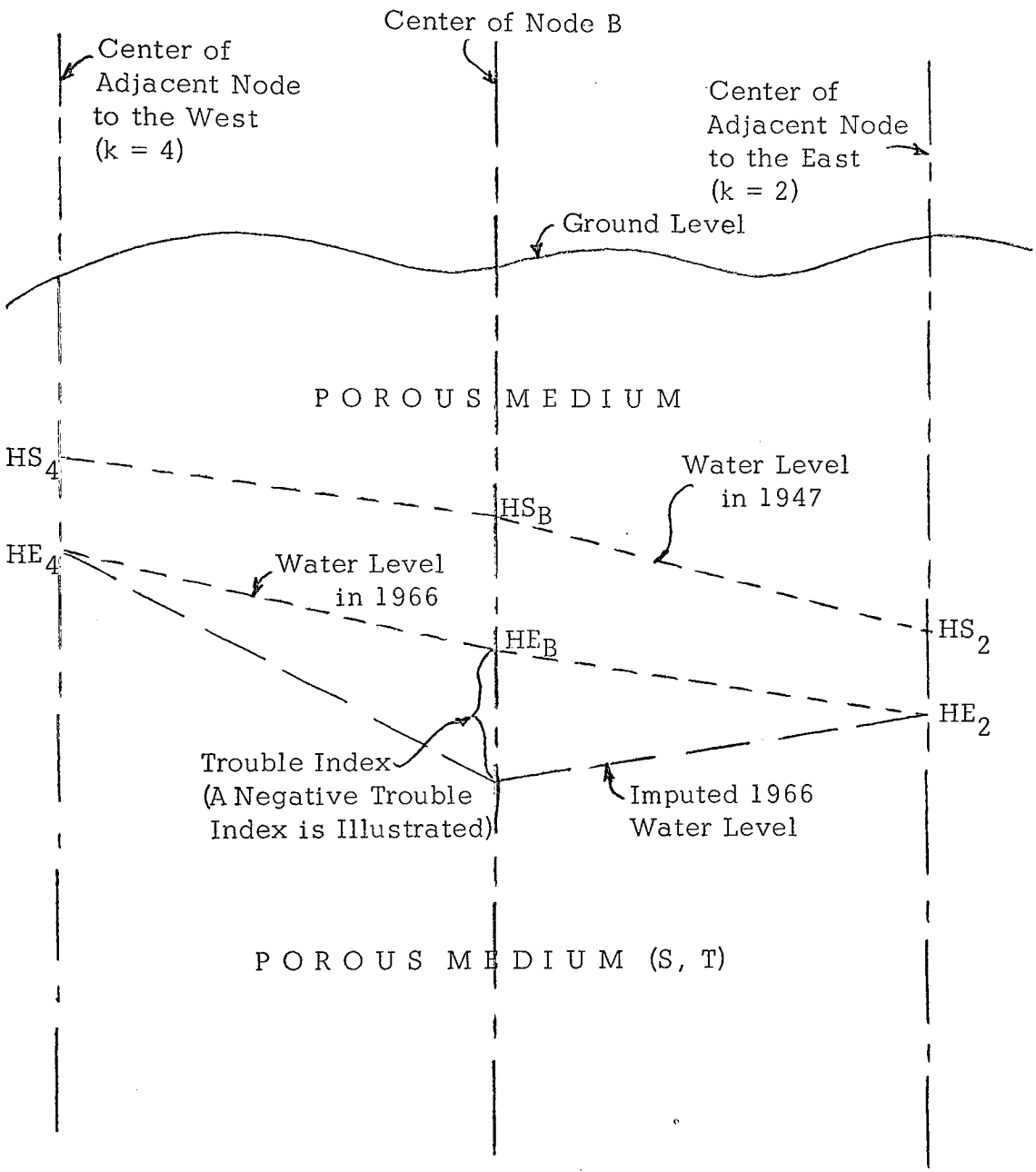

Fig. 1. The Physical Relationships Associated with the Trouble Index

Once having computed the Trouble Index, a few additional FORTRAN statements were used to assign alphabetic grades to the nodes. In this way a Node Rating map could be prepared in a form which permitted the incoherent nodes to be spotted easily, highlighted, and viewed in their spatial relationships. The following rating scheme was adopted:

A, Trouble Index less than 3 feet,

B, Trouble Index between 3 and 6 feet,

C, Trouble Index between 6 and 10 feet,

D, Trouble Index between 10 and 25 feet,

E, Trouble Index greater than 25 feet.

On the output maps an upward pointing arrow preceded each positive rating and a downward pointing arrow preceded each negative rating. Figure 2 on page 40 is a reprint of a portion of the node rating map of the nodes actually used in these investigations.

The Subjective Inputs (Confidence Maps)

When an aquifer modeling project begins, data are collected from many sources which may include geological studies of the area, hydrologic surveys, water level measurements, aquifer tests, aerial photographs, drilling logs and pumping records, utility bills; use is made of virtually any source that will permit the investigator to select a model structure, select node patterns, and assign parameters to the nodes. A detailed description of this process as applied to aquifers in California

NODE RATING. ($0 \leq A \leq 3 < B \leq 6 < C \leq 10 < D \leq 25 < E$)

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
2	.	.	.	.	↑B	↑A	↑A	↑B	↓B	↑A	↑A	↑C	.	.
3	.	.	.	.	↑A	↑A	↑A	↑B	↓B	↑A	↑A	↑A	.	.
4	.	.	.	.	↑C	↑A	↑A	↑C	↓E	↓A	↓B	↓A	.	.
5	.	.	.	.	↑D	↑A	↓C	↓A	↓D	↓C	↓A	↓A	.	.
6	.	.	.	.	↑E	↑D	↑B	↑D	↓A	↓B	↑B	↑A	.	.
7	.	.	.	.	↑E	↑D	↑A	↓B	↓B	↓A	↑A	↑B	.	.
8	.	.	.	.	↑E	↑D	↑A	↓A	↓A	↓D	↑B	↑D	.	.
9	.	.	.	.	↑E	↑D	↑A	↑B	↓B	↓C	↓A	↑D	.	.
10	.	.	.	.	↑E	↓C	↓C	↑A	↓A	↓B	↑B	↑B	↑D	.
11	.	.	.	.	↑E	↑B	↓B	↑A	↓D	↑C	↓B	↓A	↑B	.
12	.	.	.	.	↑E	↑C	↓A	↑A	↓A	↑A	↑A	↓A	↑D	.
13	.	.	.	.	↑E	↑C	↑B	↓C	↓D	↑B	↑A	↓A	↑A	.
14	.	.	.	.	↑E	↑D	↑B	↓C	↓C	↑A	↑A	↓A	↑D	.
15	.	.	.	.	↑D	↑C	↑D	↓B	↑B	↑A	↓A	↑B	.	.
16	.	.	.	.	↑A	↑C	↑A	↓D	↑A	↑A	↑A	.	.	.
17	.	.	.	.	↑A	↑C	↓C	↓B	↓A	↓B	.	.	.	.
18	.	.	.	↑E	↑C	↑A	↓D	↑A	↑D	.	.	.	.	.
19	.	.	.	↑D	↑B	↓B	↓B	↑A	↑B	.	.	.	.	.
20	.	.	.	↑D	↑B	↓B	↑A	↑B	.	.	.	.	.	.
21	.	.	↑D	↓A	↓A	↑A	↑B	.	.	.	.	.	.	.
22	.	.	↑D	↑A	↓C	↑D	.	.	.	.	.	.	.	.
23	.	↑E	↑D	↑A	↑A	↑B	.	.	.	.	.	.	.	.
24	.	↑E	↓C	↓E	↑B	↑C	.	.	.	.	.	.	.	.

Fig. 2. A Node Rating Map

is described by Allison (1967). Sometimes "data" are taken from previous models and do not represent original data in the true sense (Gates 1971, p. 1).

As the data gathering process proceeds, the data gathering group develops some feeling for the validity of the data (parameters) assigned to the model nodes. The calibration method advocated here depends very heavily on the quantification of these feelings in suitable form. The sources of these feelings are many. Certainly all of the following will contribute:

1. understanding of instrument and measurement errors,
2. knowledge or estimates of the underlying geological features,
3. awareness of the extent and assumptions made in interpolation and extrapolation of data, and
4. accumulated experience with aquifer models.

The quantification of these feelings can take the form of subjective confidence limits for the data assigned to each node. For example, the hydrologist may be confident that the data assigned to NODE (I,J) have a 90% chance of falling within the following limits:

Parameter	Confidence Bounds
Transmissivity	$T \times 1/2 \leq T_{est.} \leq T \times 2.0$
Storage Coefficient	$S \times 2/3 \leq S_{est.} \leq S \times 3/2$
Pumping	$Q \times 1/2 \leq Q_{est.} \leq Q \times 2.0$

Natural Recharge	$R \times 1/3 \leq R_{est.} \leq R \times 3.0$
Water Level at Start	$HS - 30' \leq HS_{est.} \leq HS + 30'$
Water Level at End	$HE - 25' \leq HE_{est.} \leq HE + 25'$

This information can then be put on grids, maps or in computer arrays prior to the beginning of the calibration runs.

It has been convenient to call these inputs subjective inputs even though they may contain some objective elements, and even though some of the assigned nodal parameters are also very subjective. The grids or arrays prepared from this information are here called confidence maps.

The preparation of this information need not be as tedious as it might first appear. In the most detailed case a confidence factor or increment would be attributed to each parameter at each node as the data were originally tabulated. More likely, however (and this was the case during the present investigation), the various parameter maps will be scanned by the hydrologist and confidence "contours" drawn around various clusters of nodes. In the present model in the most extreme case (transmissivity) only three levels of confidence were assigned, and in the simplest case (storage coefficient) equal confidence was assigned to all nodes. Information presented in this manner can be easily converted to the form needed for computer input.

The primary reason for getting this information a priori is, of course, to have it available for the parameter adjustment program. A

second important reason to extract this information from the hydrologist before manipulating the model is to be sure that it is free of feedback information based on examination of first iterative results. In this way an attempt is made to avoid too much interaction between the investigator and the model. This departs from the concept of Vemuri and Karplus (1969) where a hybrid computer is used and dynamic interaction between the model and the modeler is not only encouraged but is essential to the adjustment process. It would appear that hydrologists are not much better at interacting with the model on a pure hydrologic basis than systems analysts. Gates (1971), a hydrologist, reports some of his adjustments for the model of the Tucson basin aquifer were made on the basis of mathematical rather than hydrologic considerations. In this type of adjustment hydrologic considerations enter only to inhibit certain adjustments rather than to suggest or guide them.

Here the alternate approach is used where subjective input based only on hydrologic considerations is obtained, and it is applied in the adjustment process in a quantitative way. The confidence maps are used to establish weighting factors for parameter perturbation so that no subjective confidence bound is exceeded. The actual adjustments are made so that the sum of confidence-weighted perturbations is near minimum.

The Sensitivity of the Model

A completely rigorous sensitivity analysis of the model would require an evaluation of the change in the predicted water level at every node in the model for a small change in each parameter at each node. For the 1890-node model of the Tucson aquifer (6 parameters per node), the sensitivities at 1890 points to perturbations of each of 6 x 1890 = 11,340 parameters would need to be computed. Then, from this information, small adjustments could be made in accordance with some optimization scheme to improve model response. As the adjustment proceeded, the optimization scheme would require periodic re-evaluation of the sensitivities. Clearly such a scheme would be prohibitive in use of computer time, even for models with a moderate number of nodes.

Again appeal is made to the decomposition or decoupling concept introduced above under Data Coherence. The sensitivity question is reduced to "What value of the selected parameter is required at that node (if all other parameters remain constant) to produce a change in water balance at that node of 0.01 feet/year of water?" The amount of change of parameter imputed by the imbalance change is then expressed as a factor, or incremental change, whichever is appropriate for that parameter, and in keeping with the manner in which the confidences are expressed.

Some of the values imputed by these calculations will be physically unrealizable; transmissivities less than zero, storage coefficients

greater than 1.0, etc. In these cases the program assigns a vanishingly small sensitivity to that parameter. Because of the actual physical coupling between nodes, the sensitivities to certain parameters must be considered tentative. An adjustment in withdrawal rate (Q), recharge rate (R), and storage coefficient (S) at one node will clearly not propagate any change to other nodes. The same is not true, however, for transmissivity (T) or for the water levels HS and HE. An adjustment of T, HS, or HE at one node may improve the "decoupled" balance for that node but have an undesirable influence on the adjacent nodes. For this reason perturbations must be kept small, and when indicated, must be recomputed.

Development of the TRINSUCOSE Algorithm

On the basis of the error migration studies described in Chapter 2 it was felt that the errors in parameters would produce their largest distortions of the response surface in the immediate neighborhood of the node containing the errors. Errors are attenuated substantially as they migrate to distant points. The complete node decoupling concept used in generating the Trouble Index was, therefore, considered suitable for use in the parameter adjustment process.

This localized method of parameter adjustment clearly cannot lead to "the" mathematically optimal assignment of parameter values, since some of the couplings which are ignored will be of consequence.

However, as shown by the results of Chapter 4, a very large amount of improvement in prediction can be obtained with a relatively modest application of the amount of perturbation permitted within the confidence-bounded parameter ranges.

In keeping with these concepts, the adjustment considerations described above were combined into the TRINSUCOSE adjustment algorithm. The steps were (1) computation of sensitivities, (2) treatment of confidence information, (3) computation and application of weighting factors, and (4) adjustment of parameters and analysis of results. Each of these is described in a subsection below.

Computation of Sensitivities

Rearrangement of the basic differential/difference equation from Chapter 2, and incorporation of a second finite difference to replace the differential portion of that equation, permits the direct computation of the amount (rate) of imbalance of the water volume on a node by node basis. The rewritten equation is as follows:

$$\text{IMBALANCE}_B = \sum_k (h_k - h_B) T_{kB} - Q_B - R_B - A_B S_B (HE_B - HS_B). \qquad \text{Eq. (3)}$$

In the corresponding computer program the dimensions are changed so the imbalance is presented in feet/year of water for each node.

The first term on the right represents the rate of underground inflow into node B from all adjacent nodes, the second and third terms, the

rates of inflow not related to underground flow, i.e., pumping or, recharge by percolation or by simulated boundary flow, and the fourth term, the rate of increase in storage. If all are in balance, the IMBALANCE at node B will be zero. If not, the IMBALANCE at node B will be positive if the incoherence of the data imputes a buildup of water, and will be negative for a disappearance of water at that node. The water level difference (h_K-h_B) appearing in the summation is to be taken as the average water level difference between the beginning and end of the calibration period. This is in keeping with the linearity concepts expressed above in the description of the Trouble Index.

By arbitrarily setting the imbalance at each node to a value of 0.01 feet/year higher than that computed by the above equation, and by suitable algebraic rearrangement, new imputed values of T_B, S_B, (Q_B+R_B), HS_B and HE_B can be computed individually, each on the assumption that the other values are held constant. The marginal changes in these parameters to obtain this change in imbalance are the required sensitivities. The imbalance, Trouble Index, and sensitivities are computed and tabulated by FORTRAN subprogram PERTURB, which is given in Appendix A.

Treatment of Confidence Information

The aquifer parameters are of two types in the way that confidences are usually assigned. The first is the type where the confidence is expressed as some proportion of the value initially assigned. The

parameters S, T, Q, and R fall into this class. Examination of the typical statements of confidence intervals given earlier shows that the multiplier used to find the lower confidence bound is the reciprocal of that used to find the upper confidence bound. Since each bound represents the same confidence level, a non-linear relationship exists between the parameter values and confidence values. The most manageable curve that represents this type of relationship is the exponential. The form is:

Adjusted Parameter = Original Parameter x exp(f x maxp), where maxp is the maximum perturbation exponent derived from the assigned confidence level, and f is the fraction of confidence interval to be consumed in the perturbation (positive for increases in the parameter, and negative for decreases). For example, if the confidence bounds for Q_i are $0.50Q_i$ and $2.00Q_i$, maxp would be assigned a value of 0.70, since exp(-1.0 x .70) = 0.50 and exp(1.0 x .70) = 2.0. Then, if the algorithm establishes that Q_i should be perturbed by 5% of its allowable confidence in the negative direction, f would be assigned a value of -0.05, and the new value of Q_i would be Q_i x exp(-0.05 x 0.70) or $0.965Q_i$.

The second way of expressing confidence bounds is by a translation of the quantity by amounts not proportioned to the value of the parameter. The confidence bounds for the HS and HE parameters are of this type. These bounds can simply be entered directly into the program, and the adjustment equation would be as follows:

Adjusted Parameter = Original Parameter + f x Confidence Bound,

where f has the same meaning as in the first type.

The Weightings

With sensitivity, confidence and trouble data in suitable form, the next step is to combine them in a suitable way so that basic perturbation coefficients may be assigned to each parameter at each node. This is done in two weighting steps by SUBROUTINE PERTURB which is reproduced in Appendix A. An example of the calculations for an interior node of the model is given in Table I on the following page.

The first weighting step involves the combination of sensitivity and confidence data. These are combined so that each sensitivity, originally expressed as a magnitude, is represented as a fraction of the allowable perturbation of that parameter. The sign of the quantity is retained as an indicator of the direction of the change. These new measures, called sensitivity ratios, are denoted A_{Bj}, where the subscripts represent the parameter j of node B and j takes on values one through six, one value for each parameter at a node.

These sensitivity ratios have inverse relationships to the amounts of perturbation required. That is, those sensitivity ratios having the larger magnitudes are those to which the model is relatively insensitive, and vice versa. To invert and standardize these, SUBROUTINE PERTURB finds new values (X_{Bj}) such that $A_{Bj} \times X_{Bj}$ is constant, and such that the

Table I. Values Assigned to a Node by Subroutine PERTURB

(THIS EXAMPLE IS CONTINUED IN TABLE IV ON PAGE 58)

PARAMETER (NOTE 1)	CONSTRAINTS (LOWER AND UPPER BOUNDS)	SENSITIVITY RATIO	CONFIDENCE WEIGHTED AND STANDARDIZED	WITH TROUBLE INDEX WEIGHTING APPLIED (NOTE 2)
Q(10,9)	X 3/5 , X 5/3	0.011	0.059	-0.107
R(10,9)	X 5/11, X 11/5	-0.021	-0.031	0.057
S(10,9)	X 2/3 , X 3/2	-0.046	-0.014	0.025
T(10,9)	X 1/2 , X 2.0	-0.043	-0.015	0.027
HS(10,9)	±40 FEET	0.00134	0.486	-0.899
HE(10,9)	±30 FEET	0.00165	0.395	-0.723
	SUM OF ABSOLUTE VALUES		1.000	1.83

NOTES:

1. NODE(10,9) HAS IMBALANCE OF -0.38 FEET/YEAR AND TROUBLE INDEX OF -1.83 FEET.

2. THESE VALUES BECOME THE BASIC WEIGHTS FOR USE IN THE ADJUSTMENT AND TEST PROGRAMS OF THE FOLLOWING CHAPTER. SEE TABLE IV ON PAGE 58.

summation over j of the absolute values of the X_{Bj} at node B is equal to 1.0. The values for A_{Bj} and X_{Bj} at node B = (10,9) are shown in the third and fourth columns, respectively, of Table I.

The second weighting step is the trouble weighting step. Each X_{Bj} is multiplied by the Trouble Index. Sample values of this computation are given in the fifth column of Table I.

Adjustment and Test

The next step is to use the data prepared in subprogram PERTURB in conjunction with the iterative routines and error analysis subprograms to apply perturbations collectively in small amounts and to determine the amount of improvement obtained. This is done by selecting a factor--say 0.01--as the initial collective perturbation coefficient. All parameters are then adjusted by the amount that this implies for the individual parameters as determined by the data from subprogram PERTURB. Limits are applied to insure that no parameter is adjusted beyond the range originally permitted in the confidence assignment. The aquifer response is then computed and the final water levels compared with the water level data. Error data are tabulated, and the program is permitted to try larger collective adjustment factors until marginal improvement is small.

Experiments conducted thus far establish some confidence that repeated applications of this procedure may reduce the error to any

desired level, given parameter bounds that are physically capable of producing coherent or consistent results.

The TRINSUCOSE Procedure

The TRINSUCOSE procedure or algorithm, in summary, consists of the following steps:

1. Obtain condifence data in addition to the customary model parameter data. Put it in appropriate form.

2. Using SUBROUTINE STATIC, determine the overall water balance. If the data have a reasonable level of coherence at this point, proceed to the next step. Otherwise, return to step 1 and attempt to resolve the outstanding problems with the hydrologist/geologist team.

3. Using SUBROUTINE RATING, make the data coherence tests on an individual decoupled node basis. Examine the outputs for cases of extremely inconsistent data. For these cases obtain a review by the investigating team to identify mistakes in data preparation and to "educate their discretion" somewhat on the likelihood of the model coming into the desired level of calibration with the present data and bounds. If adjustments are obtained, repeat step 3 for the nodes for which adjusted information was obtained. When all identifiable mistakes have been removed and some level of static data coherence established, proceed to step 4.

4. Using SUBROUTINE PERTURB, compute and assemble the Trouble Index, Subjective Confidence and Sensitivity information into the properly weighted collective perturbation parameters.

5. Using appropriate iterative and error analysis programs, perform the collective adjustments at the minimum level necessary to obtain the desired improvement in model performance.

6. Review the results and return to any of the above steps that may be indicated if results are not satisfactory. Otherwise, assign the values obtained by the above procedures to the calibrated model.

Step 5 was conducted in the present case with PROGRAM TEST 115 which is reproduced in Appendix A along with two of the subprograms which it calls: ALDIR and ERROR. Most of the results of Chapter 4 were produced using these programs.

CHAPTER 4

RESULTS

The procedures developed in Chapter 3 were applied to the truncated (166 node) model of the Tucson aquifer. All six of the parameters at each node were perturbed in accordance with the weights assigned by SUBROUTINE PERTURB. Six different collective perturbation factors were used: 0.0, 0.02, 0.04, 0.06, 0.08 and 0.10. The first one, of course, is no adjustment at all and represents the original model data. The results in terms of error performance are tabulated in Table II on the following page. These results were each scaled to 100% for the original data and plotted in Figure 3 on the second following page.

Observe at the 0.06 level of collective perturbation that mean error is reduced to 30% of its original value. This shows that the overall water volume balance responds well to these adjustments. The other three measures all reflect model performance in a dynamic sense. At the 0.06 level all are reduced to less than 70% of their original values. Beyond the 0.06 level the marginal improvement is small. This indicates that the improvement obtainable in the first round of TRINSUCOSE computations has been reached. The amount of subjective constraint consumed thus far in the adjustment procedure can be measured in several ways.

Table II. Errors vs. Collective Perturbation Factor

COLLECTIVE PERTURBATION FACTOR	MEAN ERROR (FT.)	MEAN ABSOLUTE DEVIATION (FT.)	STANDARD DEVIATION (FT.)	PEAK ERROR RANGE (FT.)
0.00	6.11	20.11	28.77	159.15
0.02	4.45	17.60	24.55	126.56
0.04	3.02	15.36	20.96	100.75
0.06	1.80	13.44	18.09	86.36
0.08	1.19	12.54	16.75	78.89
0.10	1.07	12.26	16.26	78.52

Table III. Constraint Analysis at Perturbation Level 0.10

PARAMETER	NO. OF NODES HAVING THAT PARAMETER	NUMBER AT A CONSTRAINT	PERCENTAGE AT A CONSTRAINT
Q	66	0	0.0
R	67	7	9.6
S	166	6	3.6
T	166	0	0.0
HS	166	9	5.4
HE	166	13	7.9
TOTAL	797	35	4.4

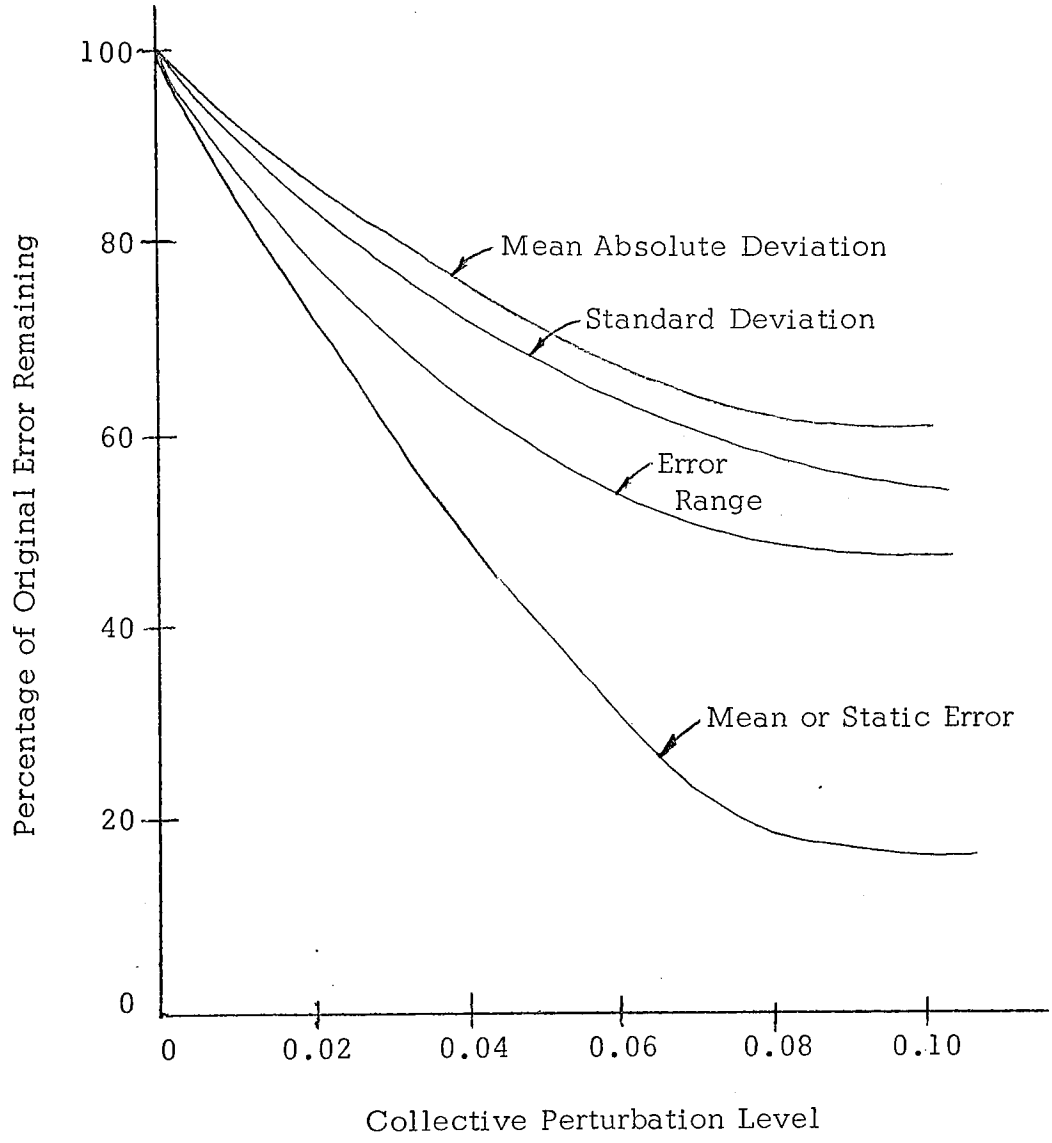

Fig. 3. Error Reduction at Various Perturbation Levels

Table III shows that very few of the parameters have reached their subjective constraints, even at the 0.10 collective perturbation level. Clearly a substantial amount of room for adjustment is available for second, third, etc. applications of the algorithm to reduce errors further.

Table IV shows the specific effects of adjustments at an interior node of the model. This is an extension of the sample computations for Node (10,9) contained in Table I on page 50. For this particular node the pattern of error reduction was as follows:

Collective Perturbation Level	Model Prediction Error (feet)
0.00	-11.3
0.02	-10.0
0.04	- 8.7
0.06	- 7.5
0.08	- 6.2
0.10	- 4.7

The node may be considered a representative node in many ways, but no node in the model can be considered typical.

The results reported here and other considerations that arose during the study form the basis of the concluding remarks of the next chapter.

Table IV. Example of Adjustment at One Node

(AN EXTENSION OF TABLE I ON PAGE 50)

PARAMETER	BASIC WEIGHT (NOTE 1)	TIMES 0.10 (NOTE 2)	ADJUSTMENT OPERATOR (NOTE 3)	CONSTRAINT REACHED (NOTE 4)	ORIGINAL VALUE	NEW VALUE
Q(1,9)	-0.107	-0.0107	X 0.9894	NO	22200	21965
R(1,9)	0.057	0.0057	X 1.0057	NO	-7340	-7382
S(1,9)	0.025	0.0025	X 1.0025	NO	0.1500	0.1504
T(1,9)	0.027	0.0027	X 1.0027	NO	63.80	63.97
HS(1,9)	-0.890	-0.0890	- 3.6	NO	2686.7	2683.1
HE(1,9)	-0.723	-0.0723	- 2.2	NO	2620.7	2618.5

NOTES:

1. TRANSCRIBED FROM THE LAST COLUMN OF TABLE I ON PAGE 50.
2. THIS EXAMPLE IS FOR PERTURBATION LEVEL 0.10.
3. ADJUSTMENT OPERATORS ARE FACTORS OR INCREMENTS DEPENDING ON PARAMETER TYPE.
4. IF CONSTRAINT IS NOT REACHED APPLY THE ADJUSTMENT INDICATED. IF A CONSTRAINT IS REACHED MOVE ONLY TO THE BOUNDARY.

CHAPTER 5

CONCLUSIONS, LIMITATIONS AND FUTURE PROSPECTS

The conclusions reached in the course of this investigation may be summarized briefly as follows:

1. Prior to mathematical manipulation of the aquifer model, parameter quality information can be generated which will measure the coherence of data, predict which nodes will be troublesome in the manipulation, and help to identify mistakes in preparation of data for the model.

2. Subjective information can be quantified in a useful way and can provide narrower constraints than the corresponding physical constraints.

3. Subjective information can be formulated into an objective function so that the adjustment of parameters proceeds in a way that the desired level of calibration is obtained with a near minimum of application of confidence-weighted adjustments to the originally assigned parameters.

4. Error propagation and migration studies of developing models can provide important insight into future model behavior and help determine whether the decoupling scheme proposed here would be applicable.

5. Nodes which cannot be brought into the desired range of calibration with the above procedures can be clearly identified by a final Trouble Index.

6. Subjective over-reaction to the iteration results is avoided by the procedures described, since for the most part the subjective information is extracted from the investigators in advance, and the opportunities for subjective interaction in the automated or semi-automated calibrations process are limited.

There are some precautions and limitations that should be observed in application of the above procedure:

1. Should the model fail to come into calibration within the established parameter bounds, it may simply mean that insufficient data are available or that the confidence bounds are too severe. But, it may also mean that the postulated model structure is incorrect. That is, the the boundaries may be incorrectly stated, assumptions of isotropic flow may be incorrect, leakiness or artesian behavior may not be accounted for, the parameters are not stable with time or level, or that the selection of nodes does not represent adequately the regions of homogeneity and heterogeneity of the aquifer. The TRINSUCOSE procedures cannot specify which of these problems may be present. In these cases the geologist/hydrologist team would have to review the results and attempt to find assignable causes to the failure of the model to come into calibration.

2. While data quality analyses are presented, in the case of the Tucson aquifer the mistakes in assembling the data were not clearly

identified. Since so many of the nodes had data of great incoherence, those data with misplaced decimal points, etc., were almost completely submerged in a sea of other troubles.

3. When the data permit optimization over several consecutive time periods, the data for the water levels at the end of one calibration period become the data for the beginning of the next. This means that the freedom to perturb water level data must be constrained at intermediate time steps. What may appear as the best water level map for behavior at the end of one time period might introduce additional errors at the subsequent time step.

4. In aquifers of high transmissivity, or in any case where the model calibration period is long with respect to the ability of errors at one node to introduce errors in distant nodes, these procedures may not be valid. Tests similar to those described at the close of Chapter 2 may help identify these conditions.

During these studies several ideas for extension and further development of these ideas arose. Some of these are as follows:

1. The subprogram PERTURB could be modified to furnish its information directly to the adjustment and test programs. In this way the intermediate data decks of weighted perturbation information could be avoided.

2. The trial multipliers for parameter perturbation could be selected by the adjustment procedures rather than by the programmer. Then a proper one could be selected by the computer for most effective adjustment.

3. The entire procedure could be made to re-evaluate itself as it proceeded. That is, after application of a selected collective adjustment factor, the results could automatically be evaluated and repeated with new Trouble Indices and sensitivities until the desired degree of error reduction was obtained, or until too many of the troubled nodes had used up all of the available perturbation.

4. The method should be applied to other regions of the Tucson aquifer, and the resulting adjustments should be compared with those obtained by the trial and error methods.

5. The method should be applied to the original data of other aquifers, preferably ones that are now considered to be well calibrated. Here, however, it would be difficult to obtain uncontaminated subjective information of the type needed.

6. Other criteria should be considered for incorporation into objective functions for parameter adjustment.

APPENDIX A

FORTRAN PROGRAMS

In the course of the investigation many FORTRAN programs and subprograms were used. PROGRAM TEST 115 produced most of the results reported in Chapter 4 and is reproduced here.

Also reproduced below are the following:

1. SUBROUTINE STATIC. Subprogram STATIC computes a complete volume balance on the aquifer and prints the results in tabular form as many times as specified by a calling argument. The printed outputs show the percentage error in volume balance as well as the net inflow/outflow in acre-feet summed from both the pumping/recharge information and the change in storage imputed by assigned storage coefficients and historical changes in water levels. The average storage coefficient is also computed.

2. SUBROUTINE RATING. Subprogram RATING is designed to permit a detailed look at the data assigned to each node and that of its immediate neighbors, to permit an initial evaluation of its validity. Computations are based on the assumption that the height of the water table at any node changes linearly with time from the initial value to the end value. Based on this assumption, the imputed values of storage coefficient, transmissivity, net recharge, volume imbalance, and end height

of the water table at the central node of each local system surrounding each node are computed. The latter figure is called the Trouble Index and is used to compute an alphabetic Node Rating at each node. In addition to the tabulated information the Trouble Index and Node Ratings are placed in arrays for call by SCLMAP if desired.

3. SUBROUTINE PERTURB. The basic structure of Subprogram PERTURB is similar to that of Subprogram RATING in the computation of volume imbalance and Trouble Index, but it uses a simpler output format. It goes beyond the computations of Subprogram RATING in that it computes the sensitivities and assigns them weights in accordance with the requirements of the TRINSUCOSE algorithm. It places its results in arrays so that they may be made available for use in the final adjustment programs.

4. SUBROUTINE ALDIR. ALDIR is the subprogram that predicts future aquifer water levels based on the present state and input stresses to the aquifer. It uses the alternating direction implicit method approximate solution of the model equations which are in simultaneous linear form.

5. SUBROUTINE ERROR. The small Subprogram ERROR was constructed to compute and print out certain identifying information and a simple analysis of the array at the bottom of any array map. Information includes mean value, mean absolute value, standard deviation, root mean square value, maximum and minimum values.

Two other subprograms were used extensively in the experiments reported here. These are the housekeeping programs SETUP and SCLMAP. They are not reproduced here but are described briefly as follows:

1. SUBROUTINE SETUP. This subroutine is a general purpose one for working with arrays of aquifer parameters. It clears arrays, fills them with constant values, reads in arrays, array headings, and scale factors in variable FORMAT from data cards, scales arrays by any desired factor, and/or prepares new data cards from array information. The modes of operation are selected by the calling arguments or use of the ENTRY point PUNCHEM. Use of the subroutine avoids the multiple placement of array manipulation statements throughout the main program.

2. SUBROUTINE SCLMAP. Subprogram SCLMAP is called when an array is to be printed in matrix form. The array, number of maps, scale factor and heading are selected by the calling arguments. The scaling for printing does not change the values in the array being printed. For alphabetic maps the ENTRY point MAPALFA is used.

Program Test 115

```
      PROGRAM TEST 115 (INPUT, OUTPUT, TAPE 5 = INPUT)

C RESTRUCTURED APRIL 19, 1971.
C ROBERT E. LOVELL - SYSTEMS ENGINEERING 410.
C ITERATION TESTS UNDER PARAMETER PERTURBATION.

      COMMON KIS,KIE, KS(51),KE(51), KJS,KJE, ISM1,IEP1,FMTY
     1 , SCLFTR, FHEAD(8), FMT1(6), WIDTH, AREA, TIME, ITER
     2 , TRATIO, AN
      DIMENSION HS(25,14), HE(25,14), T(25,14), S(25,14)
     1 , R(25,14), HT(25,14), H(25,14), PES(25,14)
     2 , PEQ(25,14), PER(25,14), SP(25,14), TP(25,14)
     3 , PERCT S(25,14), PERCT T(25,14), PERCT Q(25,14)
     4 , CS(25,14), CT(25,14), CQ(25,14), CR(25,14),FMT8T(8)
     5 , FMT8Q(8), FMT8R(8), Q(25,14), PET(25,14), FMT8(8)
     6 , PHS(25,14), PHE(25,14), PEHS(25,14), PEHE(25,14)
     7 , QRP(25,14), PERCT R(25,14), FMT8S(8)
      ITER = 0  $  TRATIO = 0.0
      TIME = 10.0  $  WIDTH = 5280.0
      AREA = WIDTH * WIDTH / 43560.0
C READ IN OUTPUT HEADING AND FIRST SUB-HEADING.
      READ(5,1) FHEAD  $  READ(5,1) FMT8
    1 FORMAT (8A10)
C READ IN THE AQUIFER LOCATION DECK.
      READ(5,2)KIS,KIE,(KS(I),I=1,25),(KE(I),I=1,25),KJS,KJE
      ISM1 = KIS - 1  $  IEP1 = KIE + 1
    2 FORMAT (2014)
   87 FORMAT (1X, A10, F10.2, 5X, 8A10)
      PRINT FMT8

      CALL SETUP (PERCT S, 25, 14, 0.0, 5, 1.0)
      CALL SETUP (PERCT T, 25, 14, 0.0, 5, 1.0)
      CALL SETUP (PERCT Q, 25, 14, 0.0, 5, 1.0)
      CALL SETUP (PERCT R, 25, 14, 0.0, 5, 1.0)
      CALL SETUP ( SP, 25, 14, 0.15, 5, 1.0)
      CALL SETUP ( TP, 25, 14, 0.0 , 5, 1.0)
      CALL SETUP ( HT, 25, 14, 0.0 , 5, 1.0)
      CALL SETUP (QRP, 25, 14, 0.0 , 5, 1.0)
      CALL SETUP (  H, 25, 14, 0.0 , 5, 1.0)
C READ IN THE HJ47 DECK.
      CALL SETUP  (HS, 25, 14, 0.0, 1, 1.0)
      PRINT 87, FMTY, SCLFTR, FMT1
C READ IN THE HH66 DECK.
      CALL SETUP  (HE, 25, 14, 0.0, 1, 1.0)
      PRINT 87, FMTY, SCLFTR, FMT1
```

```
C READ IN THE T (TRANSMISSIVITIES) DECK.
      CALL SETUP  ( T, 25, 14, 0.0, 1, 1.0)
      PRINT 87, FMTY, SCLFTR, FMT1
C READ IN THE S (STORAGE COEFFICIENTS) DECK.
      CALL SETUP  (  S, 25, 14, 0.15, 1, 1.0)
      PRINT 87, FMTY, SCLFTR, FMT1
C READ IN THE Q19 (19-YEAR PUMPING) DECK.
      CALL SETUP  ( Q, 25, 14, 0.0, 1, 1.0)
      PRINT 87, FMTY, SCLFTR, FMT1
C READ IN THE R19 (19-YEAR RECHARGE) DECK.
      CALL SETUP  ( R, 25, 14, 0.0, 1, 1.0)
      PRINT 87, FMTY, SCLFTR, FMT1
C FOLLOWING DECKS PREPARED BY SUBROUTINE PERTURB.
C   TRANSMISSIVITY PERTURBATIONS (PET-TI).
      CALL SETUP (PET, 25, 14, 0.0 , 1, 1.0)
      PRINT 87, FMTY, SCLFTR, FMT1
C    STORAGE COEFFICIENT PERTURBATIONS (PES-TI).
      CALL SETUP (PES, 25, 14, 0.0 , 1, 1.0)
      PRINT 87, FMTY, SCLFTR, FMT1
C   PUMPING PERTURBATIONS (PEQ-TI).
      CALL SETUP (PEQ, 25, 14, 0.0 , 1, 1.0)
      PRINT 87, FMTY, SCLFTR, FMT1
C   RECHARGE PERTURBATIONS (PER-TI).
      CALL SETUP (PER, 25, 14, 0.0 , 1, 1.0)
      PRINT 87, FMTY, SCLFTR, FMT1
C   STARTING WATER LEVEL PERTURBATIONS (PEHS-TI).
      CALL SETUP(PEHS, 25, 14, 0.0 , 1, 1.0)
      PRINT 87, FMTY, SCLFTR, FMT1
C   ENDING WATER LEVEL PERTURBATIONS (PEHE-TI).
      CALL SETUP(PEHE, 25, 14, 0.0 , 1, 1.0)
      PRINT 87, FMTY, SCLFTR, FMT1
C FOLLOWING DECKS CONTAIN PERTURBATION BOUNDS.
C   TRANSMISSIVITY BOUNDS (PT).
      CALL SETUP  (CT, 25, 14, 0.0, 1, 1.0)
      PRINT 87, FMTY, SCLFTR, FMT1
C   STORAGE COEFFICIENT BOUNDS (PS).
      CALL SETUP  (CS, 25, 14, 0.0, 1, 1.0)
      PRINT 87, FMTY, SCLFTR, FMT1
C   PUMPING BOUNDS (PQ).
      CALL SETUP  (CQ, 25, 14, 0.0, 1, 1.0)
      PRINT 87, FMTY, SCLFTR, FMT1
C   RECHARGE BOUNDS (PR).
      CALL SETUP  (CR, 25, 14, 0.0, 1, 1.0)
      PRINT 87, FMTY, SCLFTR, FMT1
C   STARTING WATER LEVEL BOUNDS (PHS).
      CALL SETUP (PHS, 25, 14, 0.0, 1, 1.0)
      PRINT 87, FMTY, SCLFTR, FMT1
C   FINAL WATER LEVEL BOUNDS(PHF).
      CALL SETUP (PHE, 25, 14, 0.0, 1, 1.0)
      PRINT 87, FMTY, SCLFTR, FMT1
C PERTURBATION TESTS.
```

```
C  READ IN ERROR MAP HEADING.
      READ(5,1) FMT8
      DO 75 K = 2,5
      FAC = FLOAT(1-K) * 0.02
      DO 11 I=KIS,KIE $ JS=KS(I) $ JE=KE(I) $ DO 11 J=JS,JE
      TEMPA = PHS(I,J)  $  TEMP = PEHS(I,J) * FAC * TEMPA
      H(I,J) = HS(I,J) + SIGN(AMIN1(ABS(TEMP), TEMPA), TEMP)
      TEMP = PET(I,J) * FAC  $ TEMPA = CT(I,J)
      TP(I,J)=T(I,J)*EXP(SIGN(AMIN1(ABS(TEMP),TEMPA),TEMP))
      TEMP = PES(I,J) * FAC  $ TEMPA = CS(I,J)
      SP(I,J)=S(I,J)*EXP(SIGN(AMIN1(ABS(TEMP),TEMPA),TEMP))
      TEMP = PEQ(I,J) * FAC  $ TEMPA = CQ(I,J)
      Q TEMP =Q(I,J)*EXP(SIGN(AMIN1(ABS(TEMP),TEMPA),TEMP))
      TEMP = PER(I,J) * FAC  $ TEMPA = CR(I,J)
      R TEMP =R(I,J)*EXP(SIGN(AMIN1(ABS(TEMP),TEMPA),TEMP))
   11 QRP(I,J) = Q TEMP + R TEMP
      CALL ALDIR (6, 1.0, 25, 14, QRP, SP, TP, H, HT)
      DO 16 I=KIS,KIE $ JS=KS(I) $ JE=KE(I) $ DO 16 J=JS,JE
      TEMPA = PHE(I,J)  $  TEMP = PEHE(I,J) * FAC * TEMPA
   16 HT(I,J) = HT(I,J) - HE(I,J) - SIGN(AMIN1(ABS(TEMP),
     1 TEMPA), TEMP)
      CALL SCLMAP (1, HT, 25, 14, 10.0, FMT8, 8)
      CALL ERROR (HT,25,14)
   75 CONTINUE
      STOP
      END
```

Subroutine STATIC

```
      SUBROUTINE STATIC(S,DH,QR,IR,IC,NUMBER,FMTQ,IFMTQ)
C RESTRUCTURED APRIL 19, 1971.
C COMPARISON OF NET QR WITH CHANGE IN STORAGE.
C THIS SUBROUTINE RECOMPUTES AN (THE NUMBER OF NODES).
      COMMON KIS,KIE, KS(51),KE(51), KJS,KJE, ISM1,IEPL,EMTY
     1 , SCLFTR, FHEAD(8), FMT1(6), WIDTH, AREA, TIME, ITER
     2 , TRATIO, AN
      DIMENSION S(IR,IC), DH(IR,IC), QR(IR,IC), FMTQ(IFMTQ)
C     COMPUTE AVERAGE STORAGE COEFFICIENT (ASC).
      ASC = 0.0  $  NODES = 0
      DO 8 I =KIS,KIE
      JS = KS(I)  $  JE = KE(I)
      DO 8 J = JS, JE
      NODES = NODES + 1
    8 ASC = ASC + S(I,J)
      AN = NODES  $  ASC = ASC/AN
C ENTRY CAN BE MADE HERE AFTER THE FIRST CALL TO AVOID
C     RECOMPUTATION TIME.
      ENTRY STATIC 1
C COMPUTATIONS ON THE QR ARRAY.
      NPQR = NNQR = 0  $  SUMPOS = SUMNEG = ZERO = 0.0
      DO 845 I =KIS,KIE
      JS = KS(I)  $  JE = KE(I)
      DO 845 J = JS, JE
      TEMP = QR(I,J)
      IF(TEMP) 843, 845, 844
  843 SUMNEG = SUMNEG - TEMP
      NNQR = NNQR + 1
      GO TO 845
  844 SUMPOS = SUMPOS + TEMP
      NPQR = NPQR + 1
  845 CONTINUE
C COMPUTATIONS ON THE DELTA H ARRAY.
      SUMDHP = SUMDHN = SP = SN = 0.0  $  NP = NN = 0
      DO 854 I=KIS,KIE $ JS=KS(I) $ JE=KE(I) $ DO854 J=JS,JE
      TEMP = DH(I,J)  $  IF(TEMP) 852,854,853
  852 NN = NN + 1  $  SN = SN - TEMP*S(I,J)
      SUMDHN = SUMDHN - TEMP  $  GO TO 854
  853 NP = NP + 1
      SP = SP + TEMP*S(I,J)
      SUMDHP = SUMDHP + TEMP
  854 CONTINUE
C CLOSING COMPUTATIONS.
      OUTFLOW = SUMPOS - SUMNEG
      VOLINC = SN * 640.0  $  VOLDEC = SP * 640.0
```

```
      DELVNEG = VOLDEC - VOLINC
      NZ = NODES - NP - NN   $  NZOR = NODES - NPOR - NNOR
      ANPOR = NPOR $  ANNOR = NNOR  $  ANP = NP  $  ANN = NN
      IF(NPOR .EQ. 0) ANPOR = 1.0E99
      IF(NNOR .EQ. 0) ANNOR = 1.0E99
      IF(NP   .EQ. 0) ANP   = 1.0E99
      IF(NN   .EQ. 0) ANN   = 1.0E99
      ASPOR = SUMPOS / ANPOR
      ASNOR = SUMNEG / ANNOR
      AVD   = VOLDEC / ANP
      AVI   = VOLINC / ANN
      ASP = SUMDHHP / ANP
      ASN = SUMDHHN / ANN
      PMN = SUMDHHP - SUMDHHN
      APMN = PMN / AN
      DVN = DELVNEG / AN
      AOUT = OUTFLOW / AN
      ESC = DELVNEG / (PMN * AREA)
      AMFGH2O = (OUTFLOW / DELVNEG - 1.0) * 100.0
C PRINT OUT RESULTS.
      DO 737 KQ = 1, NUMBER
      PRINT FHEAD $ PRINT FMTQ $ PRINT 716, WIDTH,AREA,TIME
      PRINT 711,NPOR,SUMPOS,ASPOR,NP,SUMDHHP,ASP,VOLDEC,AVD
      PRINT 714
      PRINT 711,NNOR,SUMNEG,ASNOR,NN,SUMDHHN,ASN,VOLINC,AVI
      PRINT 715
      PRINT 711,NZOR,ZERO,ZERO,NZ,ZERO,ZERO,ZERO,ZERO
      PRINT 715
      PRINT711,NODES,OUTFLOW,AOUT,NODES,PMN,APMN,DELVNEG,DVN
  737 PRINT 712, ASC, ESC, AMFGH2O
      RETURN
  711 FORMAT(/* STATIC ANALYSIS OF OVERALL MODEL.*20X,F6.1,
     1 * FEET BETWEEN NODES.*/54X, F6.1, * ACRES PER NODE.*
     2 /54X,F6.1,* YEARS.*////20X*--------------OR DATA-----*
     3 *----------      ------------------------DELTA  H  D*
     4 *ATA-----------------------------*/35X*TOTAL*8X*AVERAGE*
     5 18X *TOTAL*10X*AVERAGE*8X*TOTAL* 10X *AVERAGE* / 20X
     6 *NODES* 9X *VOLUME* 9X *VOLUME      NODES      DELTA *
     7 *HEAD       DELTA HEAD      DELTA VOL.      DELTA VOL.*
     8 // * POSITIVE INPUT*)
  711 FORMAT()H+, 14X, I11, 2F15.0, I10, 2F15.1, 2F15.0)
  712 FORMAT (/// *   AVERAGE STORAGE COEFFICIENT =*, F7.3,/
     1 * EFFECTIVE STORAGE COEFFICIENT =*, F7.3, //   * TOT*
     2 *AL NR EXCEEDS STORAGE CHANGE BY* F7.1* PERCENT.*///
     3 * ALL VOLUMES ARE IN ACRE-FEET.*)
  714 FORMAT (* NEGATIVE INPUT*)
  715 FORMAT (*     ZERO INPUT*)
  716 FORMAT (*0TOTALS/OUTFLOW*)
      END
```

Subroutine RATING

```
      @SUBROUTINE RATING(MODE,IR,IC,HS,HE,T,S,QR, SI,TI,QRI
     1 ,F IF)
C RESTRUCTURED APRIL 19, 1971.
C      MODE = 0 MEANS TO PREPARE ARRAYS ONLY.
C      MODE = 1 MEANS TO PRINT TABLES AND PREPARE ARRAYS.
C      TI = TROUBLE INDEX MAP, SI = THE RATING MAP,
C      QRI = THE IMBALANCE MAP, F(IF) = OUTPUT HEADING.

C PART I - DECLARATIONS AND INITIALIZATION.
      INTEGER SI, TIMAX
      @COMMON KIS,KIE, KS(5),KE(5)), KJS,KJE, ISM1,IEP1,FNTY
     1 , SCLFTR, FHEAD(8), FNT1(6), WIDTH, AREA, TIME, ITER
     2 , TRATIO, AN
      @DIMENSION HS(IR,IC), HE(IR,IC), S(IR,IC), T(IR,IC)
     1 , SI(IR,IC), TI(IR,IC), QRI(IR,IC), QR(IR,IC)
     2 , TA(4), HSA(4), HEA(4), TAB(4), DHS(4), DHE(4)
     3 , DT(4), FI(4), F(IF), DHA(4), DHAI(4)
      DO 3 IB = 1,IR  $  DO 3 JB = 1,IC  $  SI(IB,JB) = 57B
    3 TI(IB,JB) = QRI(IB,JB) = 0.0
      CA = 365.25 / (7.4805 * 43.56 * AREA)
      CC = 1.0/TIME  $  CB = CC/AREA
      NODE = ICOUNT = IPAGE = 0

C PART II - OPERATIONS.
C COMPUTATIONS OF IMPUTED VALUES AND SENSITIVITY.
      DO 617 IB = KIS,KIE  $  JS = KS(IB)  $  JE = KE(IB)
      INN = IB - 1  $  ISS = IB + 1  $  IEE = IWW = IB
      DO 617 JB = JS,JE
      JWW = JB - 1  $  JEE = JB + 1  $  JNN = JSS = JB
      NODE = NODE + 1
C SELECT REQUIRED VALUES FROM STORAGE.
      TB = T(IB,JB)  $  HSB = HS(IB,JB)  $  HEB = HE(IB,JB)
      TA(1)=T(INN,JNN)$HSA(1)=HS(INN,JNN)$HEA(1)=HE(INN,JNN)
      TA(2)=T(IEE,JEE)$HSA(2)=HS(IEE,JEE)$HEA(2)=HE(IEE,JEE)
      TA(3)=T(ISS,JSS)$HSA(3)=HS(ISS,JSS)$HEA(3)=HE(ISS,JSS)
      TA(4)=T(IWW,JWW)$HSA(4)=HS(IWW,JWW)$HEA(4)=HE(IWW,JWW)
      SB = S(IB,JB)  $  QRB = QR(IB,JB)  $  DHB = HEB - HSB
      TEMP1  = TEMP2 = TEMP10 = 0.0
C MAKE COMPUTATIONS AROUND ADJACENT NODES.
      DO 417 IG = 1,4
      DHS(IG) = HSA(IG) - HSB
      DHE(IG) = HEA(IG) - HEB
      DHA(IG) = (DHS(IG) + DHE(IG)) / 2.0
      TAB(IG) = TA(IG) + TB $ IF(TAB(IG).EQ.0.0)TAB(IG)=1.0
      DT(IG) = DHA(IG) * TA(IG)
```

```
      FI(IG) = DT(IG) / TAB(IG)
      TAB(IG) = 2.0 * TA(IG) * TB / TAB(IG)
      TEMP 10 = TEMP10 + FI(IG)
      TEMP1 = TEMP1 + (HEA(IG) + DRS(IG)) * TAB(IG) / 2.0
  417 TEMP2 = TEMP2 + TAB(IG) / 2.0
C BEGIN FINAL COMPUTATIONS.
      TEMP1 = CA * TEMP1                    $ TEMP2 = -CA * TEMP2
      TEMP3 = CC * HSB                      $ TEMP4 = -CB * QRB
      TEMP5 = HEB * TEMP2                   $ TEMP6 = SB * TEMP3
      TEMP7 = -CC * SB * HEB
C COMPUTE IMBALANCE AND STORE IN QRI.
      QRI(IB,JB) = TEMP1 + TEMP5 + TEMP4 + TEMP7 + TEMP6
C COMPUTE TROUBLE INDEX AND STORE IN TI AND NODE RATING
C     AND STORE IN SI.
      TEMP 8 = SB*CC-TEMP 2 $ IF(TEMP 8.EQ.0.0)TEMP 8=1.0E6
      HEBI = (TEMP1 + TEMP4 + TEMP6) / TEMP8
      TI(IB,JB) = TINA = TINY = HEBI - HEB
      TINXX = 7000B
      IF(TINX .EQ. ABS(TINX)) GO TO 116
      TINX = ABS(TINX)
      TINXX = 7100B
  116 IF(TINX .GT. 3.0) GO TO 117
      SI(IB,JB) = TINAX + 01B
      GO TO 121
  117 IF(TINX .GT. 6.0) GO TO 118
      SI(IB,JB) = TINAX + 02B
      GO TO 121
  118 IF(TINX .GT. 10.0) GO TO 119
      SI(IB,JB) = TINXX + 03B
      GO TO 121
  119 IF(TINX .GT. 25.0) GO TO 120
      SI(IB,JB) = TINAX + 04B
      GO TO 121
  120 SI(IB,JB) = TINAX + 05B
      IF(MODE .EQ. 0) GO TO 617
C COMPUTATION OF IMPUTED T.
  121 TEMP11 =  (TEMP4 + TEMP7 + TEMP6) / (-2.0 * CA)
      IF(TEMP 10 .EQ. 0.0) TEMP 10 = 1.0E-6
      TBI = TEMP 11/ TEMP 10
      DO 214 IQ = 1,15
      TEMP 10 = 0.0    $  IZ = IQ
      DO 213 JQ = 1,4
      TBIN = TBI + TA(JQ)  $  IF(TBIN .EQ. 0.0) TBIN = 1.0E-6
      FI(JQ) = DT(JQ) / TBIN
  213 TEMP 10 = TEMP 10 + FI(JQ)
      IF(TEMP 10 .EQ. 0.0) TEMP 10 = 1.0E-6
      TBIN = TEMP 11 / TEMP 10
      IF(TBI     .EQ. 0.0) TBI     = 1.0E-6
      IF(ABS((TBIN-TBI)/TBI)         .GE. 0.01) 214, 215
  214 TBI = TBIN
  215 TBI = TBIN
```

```
      DO 216 IQ = 1,4
      TBIN = TBI + TA(IQ)
      IF(TBIN .EQ. 0.0) TBIN = 1.0E-6
  216 FI(IQ) = 2.0 * TBI * TA(IQ) / TBIN
C COMPUTE OTHER IMPUTED VALUES.
      TEMP 8 = CC * DHB $ IF(TEMP 8 .EQ. 0.0)TEMP 8 = 1.0E-5
      SBI = (TEMP1 + TEMP4 + TEMP5) / TEMP8
      QRBI = (QRI(IB,JB) - TEMP 4) / CB
      DHBI = HEBI - HSB
      DO 372 IG = 1,4
  372 DHAI(IG) = (DHS(IG) + HEA(IG) - HEBI) / 2.0

C PART III - PRINT UP TABULAR RESULTS.
      IF (ICOUNT .GT. 0) GO TO 713
      PRINT FHEAD $ PRINT F
      IPAGE = IPAGE + 1 $ PRINT 901, IPAGE
  901 FORMAT (1H+, 124X, *PAGE*, I4, *.*)
  713 ICOUNT = ICOUNT + 1 $ IF(ICOUNT .GE. 6) ICOUNT = 0
  714 PRINT 902,NODE,IB,JB,TINY,SI(IB,JB),IZ,QRT(IB,JB),HSB
     1 , HEB
  902 FORMAT (*0NODE*, I4, *, (*, I2, *,*, I3, *)  TROUBLE *
     1 *INDEX*, F8.2, * FEET.  NODE RATING*, F4, 15X,
     2  *IMPUTED  IMPUTED ASSIGNED*, 20X,I11, / 25X, *IMBAL*
     3 *NCE*F8.2* FEET.*13X*NODE     FLOW     FLOW    AVERAGE *
     4 *AVERAGE DELTA H DELTA H*/61X,3(7X*T*),*     DELTA H *
     5 *DELTA H    START     END   H START   H END*/26X,*BASE*
     6 30X,*BASE*56X,2F8.2,/25X,*DELTA H      T         S      *
     7 * OR          N*)
      IG = 0
  715 IG = IG + 1
      IF(TA(IG) .EQ. 0.0) GO TO 617
      PRINT 903, TA(IG), TAB(IG), FI(IG), DHAI(IG), DHA(IG)
     1 , DHS(IG), DHE(IG), HSA(IG), HEA(IG)
  903 FORMAT (1H+, 63X, 9F8.2)
      GO TO 816
  617 PRINT 904, TA(IG), TAB(IG), HSA(IG), HEA(IG)
  904 FORMAT (1H+, 63X, 2F8.2, 16X,*BOUNDARY*, 16X, 2F8.2)
  816 GO TO (716, 717, 718, 617), IG
  716 PRINT 906, DHB, TB, SB, QRB
  906 FORMAT (16X, *ASSIGNED*, F8.2, F8.2, F8.2,F8.0,7X,*E*)
      GO TO 715
  717 PRINT 907, DHBI, TBI, SBI, QRBI
  907 FORMAT (17X, *IMPUTED*, 3F8.2, F8.0, 7X, *S*)
      GO TO 715
  718 PRINT 908
  908 FORMAT (63X, *J*)
      GO TO 715
  617 CONTINUE
      RETURN
      END
```

Subroutine PERTURB

```
      SUBROUTINE PERTURB (IR,IC,HS,HE,T,S,Q,R,QRI,TI,CT,CS
     1 ,CC,CR,CHS,CHE,F,IF)
C RESTRUCTURED APRIL 19, 1971.
C     THE FIRST 6 VARIABLE NAMES HAVE THE STANDARD MEANINGS.
C     QRI = IMBALANCE IN FEET.  TI = TROUBLE INDEX IN FEET.
C     THE CX ARRAYS BRING IN THE LN MAX PERTURBATION
C        INFORMATION AND RETURN THE WEIGHTED PERTURBATION
C        EXPONENTS FOR CORRECTION OF 1 UNIT OF ERROR.
C     THE CHX ARRAYS BRING IN THE MAXIMUM PERMITTED
C        CORRECTION TO WATER LEVEL AND RETURN THE WEIGHTED
C        CORRECTION INCREMENT.
C     F(IF) IS THE HEADING FOR THE OUTPUT LIST.

C PART I - DECLARATIONS AND INITIALIZATION.
      COMMON KIS,KIE, KS(51),KE(51), KJS,KJE, ISM1,IEP1,FMIY
     1 , SCLFTR, FHEAD(8), FMT1(6), WIDTH, AREA, TIME, ITER
     2 , TRATIO, AN
      DIMENSION HS(IR,IC), HE(IR,IC), T(IR,IC), S(IR,IC)
     1 , R(IR,IC), QRI(IR,IC), TI(IR,IC), CT(IR,IC)
     2 , CC(IR,IC), CR(IR,IC), TA(4), HSA(4), HEA(4), F(IF)
     3 , CHS(IR,IC), CHE(IR,IC), Q(IR,IC), CS(IR,IC)
      P MAX HS = P MAX HE = -1.0E+99  $  P MIN HS = P MIN HE
     1 = 1.0E+99
      CA = 365.25 / (7.4805 * 43.56 * AREA)
      CC = 1.0/TIME  $  CR = CC/AREA  $  NODE = 0
      DEL QRI = -0.01 / CR
      BALMAX=TIMAX=PERMAXT=PERMAXS=PERMAXO=PERMAXR =-1.0E+99
      BALMIN=TIMIN=PERMINT=PERMINS=PERMINO=PERMINR = 1.0E+99
      PRINT FHEAD  $  PRINT F
      PRINT 42

C PART II - OPERATIONS.
      DO 617 IB = KIS,KIE  $  JS = KS(IB)  $  JE = KE(IB)
      INN = IB - 1  $  JSS = IB + 1  $  IEE = IWW = IB
      DO 617 JB = JS,JE  $  NODE = NODE + 1
      JWW = JB - 1  $  JEE = JB + 1  $  JNN = JSS = JB
C SELECT REQUIRED VALUES FROM STORAGE.
      TB = T(IB,JB)  $  HSB = HS(IB,JB)  $  HEB = HE(IB,JB)
      QB = Q(IB,JB)  $  RB = R(IB,JB)  $  QRB = QB + RB
      SB = S(IB,JB)
      TA(1)=T(INN,JNN)$HSA(1)=HS(INN,JNN)$HEA(1)=HE(INN,JNN)
      TA(2)=T(IEE,JEE)$HSA(2)=HS(IEE,JEE)$HEA(2)=HE(IEE,JEE)
      TA(3)=T(ISS,JSS)$HSA(3)=HS(ISS,JSS)$HEA(3)=HE(ISS,JSS)
      TA(4)=T(IWW,JWW)$HSA(4)=HS(IWW,JWW)$HEA(4)=HE(IWW,JWW)
      TEM HEA = TEM HSA = TTA = 0.0
```

```
C MAKE COMPUTATIONS AROUND ADJACENT NODES.
      DO 417 IG = 1,4  $  TT = TA(IG)/(TA(IG) + TB)
      TTA = TTA + TT
      TEM HEA = TEM HEA + HEA(IG)*TT
  417 TEM HSA = TEM HSA + HSA(IG)*TT
C BEGIN FINAL COMPUTATIONS.
      TTA = TTA * CA  $  TTT = TTA * TB
      TEM HEA = TEM HEA * CA  $  SUM HEA = TEM HEA * TB
      TEM HSA = TEM HSA * CA  $  SUM HSA = TEM HSA * TB
      TEM HEB = HEB * TTA  $  SUM HEB = TEM HEB * TB
      TEM HSB = HSB * TTA  $  SUM HSB = TEM HSB * TB
C LET GINFLOW = UNDERGROUND INFLOW.
C     SSTORE = STORAGE AT START (RELATIVE).
C     SINFLOW = SURFACE INFLOW.
C     ESTORE = STORAGE AT END (RELATIVE).
      GINFLOW = SUM HEA + SUM HSA - SUM HEB - SUM HSB
      SSTORE = CC * SB * HSB  $ SINFLOW = -CR * QRB
      ESTORE = CC * SB * HEB
C COMPUTE AND STORE IMBALANCE AND TROUBLE INDEX.
      QRI(IB,JB) = BALANCE = GINFLOW + SINFLOW+SSTORE-ESTORE
      TI(IB,JB) = TINDEX = BALANCE / (TTA*TB + CC*SB)
      BALMAX = AMAX1(BALMAX,BALANCE)
      TIMAX = AMAX1(TIMAX,TINDEX)
      BALMIN = AMIN1(BALMIN,BALANCE)
      TIMIN = AMIN1(TIMIN,TINDEX)
C COMPUTATION OF DELTAS FOR A 0.01 FOOT/YEAR INCREASE
C     IN IMBALANCE.
      TEMP = HSB - HEB
      IF(ABS(TEMP) .LT. 0.001) TEMP = SIGN(0.001,TEMP)
      DEL SB = 0.01 * TIME / TEMP
      TEMP = TEM HEA + TEM HSA - TEM HEB - TEM HSB
      IF(ABS(TEMP) .LT. 1.0E-6) TEMP = SIGN(1.0E-6,TEMP)
      DEL TB = 0.01 / TEMP
      TEMP = CC*SB - TTT
      IF(ABS(TEMP) .LT. 1.0E-6) TEMP = SIGN(1.0E-6,TEMP)
      DEL HSB =  0.01 / TEMP
      DEL HEB = -0.01 / (CC * SB + TTT)
C COMPUTE FRACTIONAL CHANGE IN PARAMETER AND RATIO WITH
C     LOG PERTURBATION MAX.
      RS=DEL SB/AMAX1(SB,0.001)+1.0$IF(RB.EQ.0.0)RB=-0.01
      RT=DEL TB/AMAX1(TB,0.001)+1.0$IF(QB.EQ.0.0)QB= 0.01
      RR = DEL QRB / RB + 1.0
      RQ = DEL QRB / QB + 1.0
      IF(RS .LE. 0.01)RS = 0.01 $ CSB = CS(IB,JB)
      AS = ALOG(RS) / CSB
      IF(RT .LE. 0.01)RT = 0.01 $ CTB = CT(IB,JB)
      AT = ALOG(RT) / CTB
      IF(RR .LE. 0.01)RR = 0.01 $ CRB = CR(IB,JB)
      AR = ALOG(RR) / CRB
      IF(RQ .LE. 0.01)RQ = 0.01 $ CQB = CQ(IB,JB)
      AQ = ALOG(RQ) / CQB
```

```
      CHSB = CHS(IB,JB)   $  AHS = DEL HSB / CHSB
      CHEB = CHE(IB,JB)   $  AHE = DEL HEB / CHEB
     CTEMP A = 100.0/(1.0/ABS(AS)+1.0/ABS(AT)+1.0/ABS(AR)
     1 +1.0/ABS(AQ) + 1.0 / ABS(AHS) + 1.0 / ABS(AHE))
      TEMP = TINDEX  * 0.01
      XS = TEMP A / AS   $  CS(IB,JB) = AS = XS * TEMP
      XT = TEMP A / AT   $  CT(IB,JB) = AT = XT * TEMP
      XR = TEMP A / AR   $  CR(IB,JB) = AR = XR * TEMP
      XQ = TEMP A / AQ   $  CQ(IB,JB) = AQ = XQ * TEMP
      XHS= TEMP A / AHS $ CHS(IB,JB) = AHS= XHS* TEMP
      XHE= TEMP A / AHE $ CHE(IB,JB) = AHE= XHE* TEMP
      PERMAX T = AMAX1(PERMAX T, AT)
      PERMIN T = AMIN1(PERMIN T, AT)
      PERMAX S = AMAX1(PERMAX S, AS)
      PERMIN S = AMIN1(PERMIN S, AS)
      PERMAX Q = AMAX1(PERMAX Q, AQ)
      PERMIN Q = AMIN1(PERMIN Q, AQ)
      PERMAX R = AMAX1(PERMAX R, AR)
      PERMIN R = AMIN1(PERMIN R, AR)
      P MAX HE = AMAX1(P MAX HE, AHE)
      P MIN HE = AMIN1(P MIN HE, AHE)
      P MAX HS = AMAX1(P MAX HS, AHS)
      P MIN HS = AMIN1(P MIN HS, AHS)
      PRINT 48,NODE,IM,JB,BALANCE,TINDEX,AT,AS,AQ,AR,AHS,AHE
  617 CONTINUE
     CPRINT 44, BALMAX,TIMAX,PERMAX T, PERMAX S, PERMAX Q
     1 , PERMAX R, P MAX HS, P MAX HE
     2 ,          BALMIN,TIMIN,PERMIN T, PERMIN S, PERMIN Q
     3 , PERMIN R, P MIN HS, P MIN HE
      RETURN
  420FORMAT(/28X*TROUBLE*11X*BASIC PERTURBATION EXPONENTS*
     1 13X*BASIC INCREMENTS*/*  NO.  ROW  COL  IMBALANCE    *
     2 * INDEX*7X*T*11X*S*11X*Q*11X*R*10X*HS*10X*HE*/)
  440FORMAT(//8X*MAXIMUM*2F10.2,6F12.5/8X*MINIMUM*
     1 2F10.2,      6F12.5)
  48 FORMAT (1X, I4, 2I5, 2F10.2,          6F12.5)
      END
```

Subroutine ALDIR

```
      SUBROUTINE ALDIR (NUMITER,RATIO,IR,IC,QR,S,T,H,HT)
C RESTRUCTURED APRIL 19, 1971.
C FIRM REQUIREMENT - S ARRAY MUST BE NON-ZERO EVERYWHERE.
C FIXED DIMENSIONS NEED CHANGE ONLY IF IR OR IC EXCEED 51.
C NUMITER = NO. ITERATIONS DESIRED, RATIO = RATIO BETWEEN
C CONSECUTIVE TIME STEPS, IR AND IC ARE ROW AND COLUMN
C DIMENSIONS, QR IN ACRE-FEET OVER THE STATED TIME PERIOD,
C S DIMENSIONLESS, T IN KGALS/FT./DAY, H = STARTING AND
C ENDING WATER LEVELS IN FEET, AND HT = TEMPORARY AND ENDING
C WATER LEVELS IN FEET.
     0COMMON KIS,KIE, KS(51),KE(51), KJS,KJE, ISM1,IEP1,FMTY
     1 , SCLFTR, FHEAD(8), FMT1(6), WIDTH, AREA, TIME, ITER
     2 , TRATIO, AN
     0DIMENSION QR(IR,IC), S(IR,IC), T(IR,IC), H(IR,IC)
     1 ,HT(IR,IC),AC(51),BC(51),CC(51),DC(51),W(51),G(51)
      TFUNCT(ADJ T) = CONST*ADJ T*BASE T/(ADJ T + BASE T)
      TIME S = 365.25 * TIME  $  AREA B = AREA * 43560.0
      CONST = 2.0 * 1000.0 / 7.4805 $ QRFAC = 43560.0/TIME S
      ITER = NUMITER  $  TRATIO = RATIO
      KISP1 = KIS + 1  $  KJSP1 = KJS + 1
      IF(RATIO .NE. 1.0) GO TO 16
      DELT = TIME S / FLOAT(NUMITER)  $  GO TO 17
   16 DELT = TIME S/RATIO*(RATIO - 1.0)/(RATIO**NUMITER-1.0)
   17 DO 75 KNUM = 1,NUMITER
      DELT = DELT*RATIO  $  FACTOR = 2.0 * AREA B / DELT
      DO 110 J = KJS,KJE  $  DO 115 I = KIS,KIE
      BASE T = T(I,J)  $  GAM = FACTOR * S(I,J)
      IF(BASE T .NE. 0.0) GO TO 32
      AC(I) = CC(I) = 0.0 $ BC(I) = DC(I) = -GAM $ GO TO 115
   32 AC(I) = TFUNCT(T(I-1,J))  $  EC =TFUNCT(T(I,J-1))
      CC(I) = TFUNCT(T(I+1,J))  $  FC =TFUNCT(T(I,J+1))
      BC(I)=-CC(I)-AC(I)-GAM
     0DC(I)=-EC*H(I,J-1)+(EC+FC-GAM)*H(I,J)-FC*H(I,J+1)
     1 + QR(I,J)*QRFAC
  115 CONTINUE
      W(KIS) = BC(KIS)  $  G(KIS) = DC(KIS) / W(KIS)
      DO 120 K=KISP1,KIE $ W(K)=BC(K)-AC(K)*CC(K-1)/W(K-1)
  120 G(K)=(DC(K)-AC(K)*G(K-1))/W(K)  $  HT(KIE,J) = G(KIE)
      DO 110 K = KISP1,KIE  $  KBW = KIE - K + KIS
  110 HT(KBW,J)=G(KBW)-CC(KBW)*HT(KBW+1,J)/W(KBW)
      DO 95 I = KIS,KIE  $  DO 95 J = KJS,KJE
   95 H(I,J)=HT(I,J)
      DO 80 I = KIS,KIE  $  DO 85 J = KJS,KJE
      BASE T = T(I,J)  $  GAM = FACTOR * S(I,J)
```

```
      IF(BASE T .NE. 0.0) GO TO 42
      AC(J) = CC(J) = 0.0 $ BC(J) = DC(J) = -GAM $ GO TO 85
   42 AC(J) = TFUNCT(T(I,J-1))   $  EC =TFUNCT(T(I+1,J))
      CC(J) = TFUNCT(T(I,J+1))   $  FC =TFUNCT(T(I-1,J))
      BC(J)=-CC(J)-AC(J)-GAM
     CDC(J)=-H(I+1,J)*EC+H(I,J)*(EC+FC-GAM)-H(I-1,J)*FC
     1 + QR(I,J)*ORFAC
   85 CONTINUE
      W(KJS) =BC(KJS)   $  G(KJS) = DC(KJS) / W(KJS)
      DO 90 K = KJSP1,KJE $ W(K)= BC(K)-AC(K)*CC(K-1)/W(K-1)
   90 G(K)=(DC(K)-AC(K)*G(K-1))/w(K)   $  HT(I,KJE) = G(KJE)
      DO 80 K = KJSP1,KJE   $  KBW = KJE - K + KJS
   80 HT(I,KBW)=G(KBW)-CC(KBW)*HT(I,KBW+1)/W(KBW)
      DO 75 I = KIS,KIE   $  DO 75 J = KJS,KJE
   75 H(I,J)=HT(I,J)   $  RETURN
      END
```

Subroutine ERROR

```
      SUBROUTINE ERROR (A, IR, IC)
C TO COMPUTE STATISTICS AND PRINT THEM ON BOTTOM OF MAPS.
C RESTRUCTURED APRIL 19, 1971.
     0COMMON KIS,KIE, KS(51),KE(51), KJS,KJE, ISM1,IEP1,FMTY
     1 , SCLFTR, FHEAD(8), FMT1(6), WIDTH, AREA, TIME, ITER
     2 , TRATIO, AN
      DIMENSION A(IR,IC)
      KONTROL = 0
    3 ZMAX = -1.0E+99  $  NODES = 0
      ZMIN =  1.0E+99  $  E = EE = AN = AE = 0.0
      DO  1 I=KIS,KIE $ JS=KS(1) $ JE=KE(I) $ DO  1 J=JS,JE
      NODES = NODES + 1  $  AIJ = A(I,J)
      ZMAX = AMAX1(AIJ,ZMAX)  $  E = E + AIJ
      ZMIN = AMIN1(AIJ,ZMIN)  $  EE = EE + AIJ*AIJ
    1 AE = AE + ABS(AIJ)
      AN = NODES  $  RMS = SQRT(EE/AN)  $  AE = AE/AN
      EE = SQRT((AN*EE - E*E) / (AN*(AN - 1.0))) $ E = E/AN
      IF (KONTROL .EQ. 1) RETURN
C ENTRY POINT FOR PRINTING OF RESULTS.
      ENTRY ERROR 1
     0PRINT 2, AREA, ITER, ZMAX, AE, ZMAX, AE, TIME, TRATIO
     1 , E, RMS, E, RMS, NODES, ZMIN, EE, ZMIN, EE
      RETURN
C ENTRY POINT FOR COMPUTATION OF VALUES ONLY.
      ENTRY ERROR 0
      KONTROL = 1  $  GO TO 3
    20FORMAT(1X,F5.0* ACRES PER NODE.    *I2* ITERATIONS.*4X
     1 *MAXIMUM VALUE =*F7.2* FEET.  MEAN ABSOLUTE VALUE =*
     2 F7.2* FEET.*2E12.2/1X,F5.0* YEARS.*10X,F5.2* TIME RA*
     3 *TIO.*7X*MEAN VALUE =*F7.2* FEET.*12X*RMS VALUE =*
     4 F7.2* FEET.*2E12.2/1X,I5* NODES.*31X*MINIMUM VALUE =*
     5F7.2* FEET.   STANDARD DEVIATION =*F7.2* FEET.*2E12.2)
      END
```

APPENDIX B

RESULTS OF ERROR PROPAGATION STUDIES

Appendix B provides additional details of the nine error propagation tests conducted on the uncalibrated model of the Tucson basin aquifer. It amplifies the information furnished in Chapter 2 on pages 26 and 27.

Three tests were conducted to observe how errors in the pumping/recharge parameters of the model would move through the aquifer in different durations of time. The behavior of the pumping Q and recharge R parameters are indistinguishable in the model (Eq. 1 on page 11). They are called QR in the iterative program, SUBROUTINE ALDIR.

First, QR at the interior node (13,9) was increased by 500 acre-feet for the first year only. Because of the sign convention adopted in this model, this represented additional withdrawal of that amount from the node equivalent to a local error in water level of -5.2 feet. Over an extended period, this perturbation would represent a mean error over the 166-node model of -0.03 feet. At the end of two years the peak error was down to 10% of its original value and was still located at the node where the error was originally introduced. By the tenth (and last) year

of the test, the peak error was located at node (13,10), one mile east of the node where the error was originally introduced and was then less than 2% of the original value.

The second QR error experiment was conducted in the same way except that the error was introduced at a boundary node, (13,5). This happened to be a region of lower transmissivity. In this case the peak error remained at node (13,5) throughout the ten years of the test. By the end of the tenth year the error was reduced to 10% of its original value.

The last QR error experiment was similar to the second except, in this case, the 500 acre-foot error at node (13,5) was introduced in each of the 10 years of the run. In this case the peak error remained at node (13,5) throughout the test but was 30% of the value of the total error introduced over the 10 years.

Three tests were conducted to observe how an error in the starting water level (HS) would behave. For the first test, an error of -5.2 feet was introduced at interior node (12,8), and the behavior was observed at one-year intervals for 10 years. This error corresponded to a volume error of 500 acre-feet and, therefore, was equal in sign and magnitude to the mean error associated with the first and second QR error tests described above. At the end of the first year the peak error was 12% of the original value and at the end of the tenth year was less than 2% of the

original value. These results were similar to those observed in the first QR test. However, in the HS case, the peak error exhibited a propensity to drift to nearby nodes. At the end of the first year the peak was found at the adjacent node to the east, (12,9). By the second year it had returned to its original location. At the end of the sixth year it had moved to the node immediately to the west (12,7), and by the tenth year was located two nodes to the west. This opened the question of how it might behave over an extended time period.

The second HS test was an extension of the first, with tabulated output every 5 years for 45 years. By the fifteenth year the node with peak error had migrated to the west boundary 3 miles away, (12,5), and was less than 1.5% of the original size. The peak error remained at this node until the end of the 45-year period but, by that time, was almost uniformly distributed over the model.

A third HS test was conducted to observe whether the same sort of behavior would be observed with the sign of the water level error changed to positive. A 30-year test was conducted with the expectation that changing the sign would cause error migration in the opposite direction. This, however, was not the case. Except for the reversals of sign the results throughout the 30-year comparison period were identical with those of the second HS test. The direction was exactly the same.

A storage coefficient (S) test was conducted by doubling the storage coefficient at interior node (11,7) and observing the behavior at one-year intervals for a period of 10 years. In this case all measures of error increased with time. However, this can be attributed to the falling water level at that node. The drop in water level over the 10-year period was approximately 30 feet. Each drop in water level introduced an additional increment of water into the overall water balance and, therefore, increased the mean error. Also, as this surplus water spread to other nodes, the measures of dynamic error also increased. The important result from this test was that the node in which the error was introduced retained the peak error throughout the test.

A transmissivity (T) error test was conducted by doubling the transmissivity at interior node (12,8) and observing behavior at one-year intervals for 10 years. The mean error remained 0.00 at all times since the volume balance had not been disturbed. The peak positive error appeared at the adjacent downstream node (11,8), and the peak negative error at the adjacent upstream node (13,8), and remained there throughout the test. The build-up of water to the north and loss of water to the south is accounted for by the increased transmissivity and the fact that the primary direction of flow through node (12,8) is to the north. The peak errors were +1.33 feet and -2.40 feet at the end of the first year. By the end of the tenth year they had increased to +3.11 feet and -2.96

feet. Standard deviation increased from 0.30 feet to 0.61 feet over the same period. In an attempt to determine error trends over the aquifer, a series of least-squared error planes was fitted to the error data -- one for each time step. An analysis of the results showed that the greatest slope of the error plane increased approximately linearly with time at the rate of 0.003 feet per year.

Clearly, model errors which are caused by transmissivity errors are much more likely to increase with time than are model errors associated with errors in the other parameters. A transmissivity error at one node only can easily be identified by the appearance of an increasing positive error on one side and an increasing negative error on the other. In the general problem of model calibration, in which the transmissivity at each node is subject to uncertainty, the pattern of transmissivity errors would not likely be identifiable by inspection. However, the fact that the error, at all times, remains centered about the node in which the error occurs provides additional support for the decoupling and local adjustment concepts described in Chapter 3.

In some types of pumping tests of aquifer values for the parameter T/S are obtained. Other tests or assumptions are required to separate these into two parts. Therefore, an additional error test was conducted in which the value of T/S was assumed correct, but both T and S were doubled at node (12,8) so that the effect of incorrect separation of

the T and S parameters could be identified. The results were almost an algebraic sum of the individual S and T error tests. All measures of error increased as the water level fell over a period of time. Here again the peak errors remained clustered about the node in which the original error was introduced.

In summary, these tests indicate that, in the case of the Tucson basin aquifer, the logical place to begin correction of errors observed in the calibration process is at the nodes where the errors are observed. For the node size, transmissivities and the time span of the calibration period used here, there is no substantial probability that prediction errors at one node are caused primarily by the parameter or data errors at some distant node.

REFERENCES CITED

Allison, Stephen V. Cost, Precision, and Value Relationships of Data Collection and Design Activities in Water Development Planning, Water Resources Center Contribution No. 120, Hydraulic Laboratory, U. of California, Berkeley, Technical Report 6-27, May 1967.

Bouwer, Herman. Planning and Interpreting Soil Permeability Measurements, Journal of the Irrigation and Drainage Div., ASCE, 95 (IR3), 391-402, 1969.

Bruce, G. H., D. W. Peaceman, H. H. Rachford, Jr., and J. P. Rice. Calculations of Unsteady-state Gas Flow through Porous Media, Petroleum Transactions, Am. Inst. Minging Engrs., 198, 79-92, 1953.

Coats, K. H., J. R. Dempsey, and J. H. Henderson. A Practical Method for Determining Reservoir Description from Field Performance Data, Society of Petroleum Engineers of the AIME, Paper presented at the 43rd Annual Fall Meeting, Houston, Sept. 29-Oct. 2, 1968.

Darcy, Henri. Les Fontaines Publiques de la Ville de Dijon, Paris, 1856.

Emsellem, Y., and G. de Marsily. Présentation d'une méthode nouvelle de resolution du "problème inverse," École Nationale Supérieure des Mines de Paris, Centre d'Informatique Géologique, May 1969.

Essenwanger, Oskar M. Computerized Quality Control as Applied to Upper Atmosphere Data, Proceedings of the Fifteenth Conference on the Design of Experiments in Army Research, Development and Testing, ARO-D Report 70-2, 1970.

Gates, Joseph S. Development of the Digital-Computer Model of the Groundwater Reservoir in the Tucson Basin, and its Future Use in Error and Worth-of-Data Studies, Progress Report to the Water Resources Research Center, University of Arizona, September 1970.

Gates, Joseph S. Calibration of the Digital-Computer Model of the Tucson Basin, and Development of the Small-Scale Model (1-mile Nodal Spacing) of the Basin, Appendix III of a contract proposal to the City of Tucson, Dept. of Water and Sewers by the Hydrology and Water Resources Office, University of Arizona, January 1971.

Haimes, Y. Y., R. L. Perrine, and D. A. Wismer. Identification of Aquifer Parameters by Decomposition and Multilevel Optimization, Israel Journal of Technology, 6(5), 322-329, 1968.

Hubbert, M. King. The Theory of Ground-Water Motion and Related Papers, Hafner, 1969.

McMillan, William D. Theoretical Analysis of Groundwater Basin Operations, Water Resources Center Contribution No. 114, Hydraulic Laboratory, U. of California, Berkeley, Technical Report 6-25, November 1966.

Moench, Allen. Digital Model of the Tucson Basin -- Initial Development, Hydrology and Water Resources Office, University of Arizona, Intraoffice memorandum 14a, August 1, 1969.

Pinder, George F. and J. D. Bredehoeft. Application of the Digital Computer for Aquifer Evaluation, Water Resources Research, 4(5) 1069-1093, October 1968.

Pliska, Stanley R. A Mathematical Model of a Ground Water Basin, Paper presented at the ORSA/TIMS Meeting, San Francisco, May 1-3, 1968.

Thiessen, Alfred H. Precipitation Averages for Large Areas, Monthly Weather Review, 39, 1082-1084, July 1911.

Tyson, H. N., Jr., and E. M. Weber. Groundwater Management for the Nation's Future - Computer Simulation of Groundwater Basins, Proc. Am. Soc. Civil Engr. 90, HY4, 59-78, 1964.

Vemuri, Venkateswararao and Walter J. Karplus. Identification of Nonlinear Parameters of Ground Water Basins by Hybrid Computation, Water Resources Research, 5(1) 172-185, February 1969.

Warren, J. E., and H. S. Price. Flow in Heterogenous Porous Media, Petroleum Transactions, Am. Inst. Mining Engineers, 222, 155-169, 1965.