WATER
WATER
WATER
WATER
WATER
WATER
WATER
WATER
WATER
WATER

PROJECT COMPLETION
REPORT NO. 425X

Integrated System Identification and Optimization For Conjunctive Use Of Ground and Surface Water Phase I

October 1973

by
Yacov Y. Haimes
Associate Professor
Systems Engineering Department
Case Institute of Technology
Case Western Reserve University
Cleveland, Ohio

United States Department
of the Interior

CONTRACT NO.
B-045-OHIO

State of Ohio
Water Resources Center
Ohio State University

INTEGRATED SYSTEM IDENTIFICATION AND OPTIMIZATION

FOR

CONJUNCTIVE USE OF GROUND AND SURFACE WATER

PHASE I

by

Yacov Y. Haimes
Associate Professor
Systems Engineering Department
Case Institute of Technology
Case Western Reserve University
Cleveland, Ohio

WATER RESOURCES CENTER
THE OHIO STATE UNIVERSITY
COLUMBUS, OHIO 43210

October, 1973

This research was supported in part by the
Office of Water Resources Research
U.S. Department of the Interior
under Project B-045-OHIO.

PROJECT STAFF

Yacov Y. Haimes, Ph.D.
Principal Investigator

Leon S. Lasdon, Ph.D.
Co-Investigator

Prasanta Das, M.S. - *Research Assistant*
Yosef Driezin, M.S. - *Research Assistant*
Hernan Lopez, M.S. - *Research Assistant*

Richard L. Perrine, Ph.D.
Project Consultant
University of California
Los Angeles

David K. Todd, Ph.D.
Project Consultant
University of California
Berkeley

ACKNOWLEDGEMENTS

We would like to express our gratitude to the following organizations and individuals who have been invaluable in our research efforts in this project. First to the Office of Water Resources Research, U. S. Department of the Interior for their support of this study. To the Ohio State University, Water Resources Center, and in particular to Dr. K. S. Shumate, former Director of the Center, Dr. Robert C. Stiefel, Director, and to Mr. James F. Teeple, Assistant to the Director for their assistance and advice. To the Miami Conservancy District, Dayton, Ohio and specifically to Messrs. Robert Schroer and Paul Plumer who provided data and significant advice and support. To the staff of the U. S. Geological Survey, Water Resources Division in Washington, D. C. and particularly to Drs. Nicholas C. Matalas and Thomas Maddock, III who provided a documented aquifer simulation model. The collaboration and continuous exchange of information with the staff of the Water Resources Division, U.S.G.S. have significantly enhanced the progress of this Project. Finally, we would like to thank the faculty members at Case Western Reserve University who served on the thesis committees of the graduate students involved in this project. In particular to Professors M. A. Cardenas and Gilmer L. Blankenship.

EXECUTIVE SUMMARY

Highlights of Phase I Project Efforts

The following four points indicate the direction and accomplishments of the first year of project activity.

1. In a continuing effort to improve the knowledge of the aquifer system characteristics, a computer program which simulates an aquifer using an alternating direction implicit-iterative procedure has been obtained from the U. S. Geological Survey. This aquifer simulation model was utilized as a component in the system identification effort.

2. The search for a real data base for the final implementation and validation of the models has been very successful. We have established a working relationship with the Miami Conservancy District and their staff have expressed its interest in our research and promised to collaborate and assist in supplying groundwater and other hydrological data from the District. This has aided in identifying the model parameters and in validating and testing the model once the identification was completed. Furthermore, water quality models currently under development by the District engineers should be useful when the validated models are used to derive improved water management policies.

3. Our efforts have focussed on the development of improved solution methodologies for the identification and optimization problems. A system identication methodology, applicable to calibration of confined and unconfined (under certain constraints) aquifer models described by parabolic partial differential equations, has been developed. Work has been completed on the formulation and coding of a digital computer

software package capable of estimating the values of transmissivity, as a function of location within the aquifer. The estimation is based on measured and observed data on the waterhead and the various flows from and into the aquifer. The identification methodology has been successfully applied to the parameter identification of an aquifer model simulating the behavior of a real aquifer system. The Fairfield-New Baltimore aquifer in southern Ohio was chosen as the problem site in collaboration with the Miami Conservancy District. After several meetings with personnel from the Miami Conservancy District and after the initial stage of data collection was completed, a two dimensional partial differential equation model was postulated for the region. Calibration of the model was pursued using the new system identification methodology developed in this research. The results are satisfactory and establish that analytical means, as opposed to simulation by trial-and-error techniques, are feasible and yield excellent results. The identification phase has been completed and will be integrated with the overall ground and surface management model in Phase II.

4. The product of Phase I is a complete, validated aquifer model which we believe is most valuable to engineers and managers concerned with groundwater systems. It can be used both for simulation and/or optimization. In particular it is of direct value to engineers and analysts wishing to know the response of the aquifer system to various demands placed upon it. The model will be fully utilized, of course, for the conjunctive management of ground and surface water developed in Phase II of this project. The model is in the form of a FORTRAN V computer program.

Application of Research Results

This research was conducted in close cooperation with the Miami Conservancy District, Dayton, Ohio. We expect that our models and results will be utilized by the above agency as well as others. In addition, we are closely communicating with the U. S. Geological Survey, Water Resources Division, in Washington, D. C. Our results are being evaluated by this agency and we again hope that they will be utilized successfully.

Work Remaining, and Progress Contemplated During Phase II

Phase II will continue the development of the management model by utilizing the identification and groundwater models developed in Phase I. This development for the overall optimal conjunctive management of ground and surface water includes the following:

1. Objectives and constraints arising when conjunctive use of ground and surface water is considered will be formulated. In particular, emphasis will be focussed on the joint formulation of the system modeling (identification) and optimization of ground and surface water resources for their optimal management.
2. Alternative courses of action will be identified and formulated mathematically as decision variables within the model.
3. A forecasting subsystem will be developed to provide estimates of water demands of users for both long and short term models.

In points 1-3 above, the experience of the Miami Conservancy District engineers will be most valuable in assuring that the

management model developed is realistic and useful. Usage data from that District will serve as input to the forecasting subsystem, and cost data will aid in establishing objective functions.

Surface water data may have to be augmented from sources other than the Miami Conservancy District for this project. A decision on this matter will be made later in Phase II when the inventory and evaluation of the data available from the District is completed.

4. A number of optimization problems will be formulated, whose solutions should yield improved policies for water management. These fall into two categories:

   Short term planning model. This has an objective to show how user demands over a 3-5 year period may be met at minimal cost.

   Long term planning model. This model includes as alternatives various capital construction projects for increasing water supply. It provides information on which project should be constructed and when and how the water from these and from existing projects should be distributed and used. The time horizon is 30-40 years.

5. Data on costs, forecasts of demands, and various assumptions on natural flows into the water resource system will be gathered or generated. These latter assumptions can be used to answer questions as to what are the best water management policies under drought conditions, as well as under normal inflow conditions. The effects of various assumptions on industrial and residential growth will also be investigated.

6. Solution methods for the joint system modeling and optimization will be developed, coded, and tested with real data.

7. Results will be analyzed, and if indicated the model will be adjusted and rerun. Here again, close cooperation with the

Miami Conservancy District will ensure that the final product is realistic and of value to other areas with real decision problems.

8. A report documenting all the research phases and findings will be made available at the end of Phase II.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES

Chapter 1

PROBLEM DEFINITION AND DATA COLLECTION

1.1 Problem Definition

More than half of the water supply of the United States is extracted directly from groundwater systems. Groundwater also acts as an enormous regulating reservoir providing for the natural base flows of streams. Increasingly man-induced recharge is being used to augment natural replenishment of aquifers. In many locations groundwater reserves are utilized for interim economic development. Aquifers in general are capable of storing most of the cumulative excess water runoff or of being recharged with water from other sources for use in water shortage periods. This often eliminates the need for constructing expensive large dams and reservoirs.

Groundwater is clearly one of the major elements of our water resources. Its wise management must be considered a necessity. Key to the optimal management of this resource for any of its multiple beneficial uses is an ability to predict with reasonable accuracy the response of the system to decisions affecting recharge and withdrawals. This in turn is dependent upon the values of the important parameters such as permeability, specific yield, transmissivity, etc., as these are distributed over the aquifer formations in space. Once the physical properties and characteristics of an aquifer are known, it is possible to apply appropriate physical laws and to predict the response of the aquifer system to demands placed on it.

Without such knowledge prediction is impossible.

Groundwater should not be considered as an isolated resource, however. Extensive research has been devoted to the problem of the conjunctive use of ground and surface water. The overall goal of this study is thus to develop methodologies for the optimal planning and management of the conjunctive use of groundwater and surface water facilities by considering the problem of estimation of physical hydrologic response model parameters and the optimization of conjunctive water system operation and management together in an integrated fashion. In addition, multiple system objectives are to be considered simultaneously.

Obtaining the required aquifer system parameter values directly in a sufficiently dense space-time network, by an extensive data observation system would be prohibitively expensive. For this reason, most of the parameter values used are deduced from the behavior of the system rather than by direct observation. Because of analytical limitations the values presently being obtained represent essentially an estimate of the parameters near the well being observed, i.e., a point observation. The aquifer as a system, however, responds to these values as distributed over the system, not just in the vicinity of the well. What is needed is an analytical procedure which will permit a determination of those parameters which will reflect, through an optimum weighting process, the effect of the distribution of their values as a function of time and space in a way which provides the most relevant information for the management optimization problems. This is the approach of the aquifer model investigation.

One of the major objectives of this research is to develop a generalized analytical procedure whereby the past history of an aquifer or its response

to a future series of planned drafts can be utilized. In particular:

(i) To evaluate the basic aquifer parameters as a function of position.

(ii) To optimize this evaluation so that the results have the maximum relevance to the determination of the optimum management policy.

The numerical values of the parameters assigned to various points in the aquifer not only represent an average of the distributed effect of that parameter but are also properly weighted with regard to the effect of management decisions regarding extractions and withdrawals.

Phase II of the research plan involves the joint consideration of the system modeling and optimization of ground and surface water resources for their optimal management. The identification of aquifer system parameters, (e.g. transmissivity and storage functions) as part of system modeling, has traditionally been considered and treated separately from the optimization procedure. However, since the unknown aquifer parameters are used in determining the optimal decision variables as shown in earlier work, the integrated system identification and optimization will be considered for the optimal management of ground and surface water. The objectives of this study are to utilize advanced aquifer system models in a total planning model which considers the conjunctive effects of ground and surface hydrology, multiple planning objectives and both long and short-range real water resources management situations.

The objectives of the research are attained by the following steps:

(i) Detailed formulation of the joint system identification and optimization problem.

(ii) Improvement of the aquifer system models previously developed so as to fit the integrated formulation.

(iii) The development of improved solution methodologies for the identification and optimization problems. (Decomposition and multilevel techniques will be considered.)

(iv) The implementation of the above formulation and solution to a real ground and surface water system.

The first two steps outlined in the objectives involve the application of more realistic mathematical models of aquifers. The forms of the equations in such models are well known (such as the diffusion equation). As stated above however, the model parameters such as transmissivity and storage are distributed (i.e. vary with spatial location) and are usually unknown. In this research mathematical methods have been applied for system identification to determine these parameters. The last two steps use the aquifer parameters obtained above in an optimization scheme whose results can lead to a policy for better management.

Although such optimization has been proposed previously in the literature, most such work has not taken into account the distributed and dynamic character of aquifers. Phase II of this study will involve the joint consideration of the system modeling and optimization of ground and surface water resources for their optimal management. By this joint consideration, shortcomings of previous developments will be overcome.

It should be emphasized that a real ground and surface water resource system has been selected for study and the above methodology applied to this system.

1.2 Description of Real Aquifer System: Miami Conservency District

The area modeled for the validation of the identification algorithm is the Fairfield-New Baltimore area of the Miami Conservency District, which consists of 32 square miles of the Great Miami River Valley southwest of Hamilton, Ohio. The area modeled possesses a sand and gravel aquifer that is bounded by the bedrock walls of the Great Miami River Valley. The bedrock walls form the boundary of the aquifer, with the exception of the west and north, where the boundaries are arbitrary. For the west boundary the dry fork of the White Water River, located about two miles west of New Baltimore was selected as the boundary. For the northern boundary a line through Fairfield near the southern city limit of Hamilton was chosen.

Geologically, the aquifer under study consists of glacial outwash sands and gravels of the Pleistocene Age. From the hydrogeological point of veiw, the aquifer area can be conveniently divided into three parts; these are described as follows.

In the central part of the area the aquifer material consists of stratified sand and gravel situated 150-200 feet below ground surface. Widely scattered lenses of clay and silt are also present but are not of sufficient areal extent to cause any perceptible confining effects. In the southwest corner area the sand and gravel is only about 80 feet thick,

Along the eastern edge of the area some three square miles consists of a sand and gravel aquifer, which is about 100 to 150 feet thick and is overlain by about 100 feet of clay and silt.

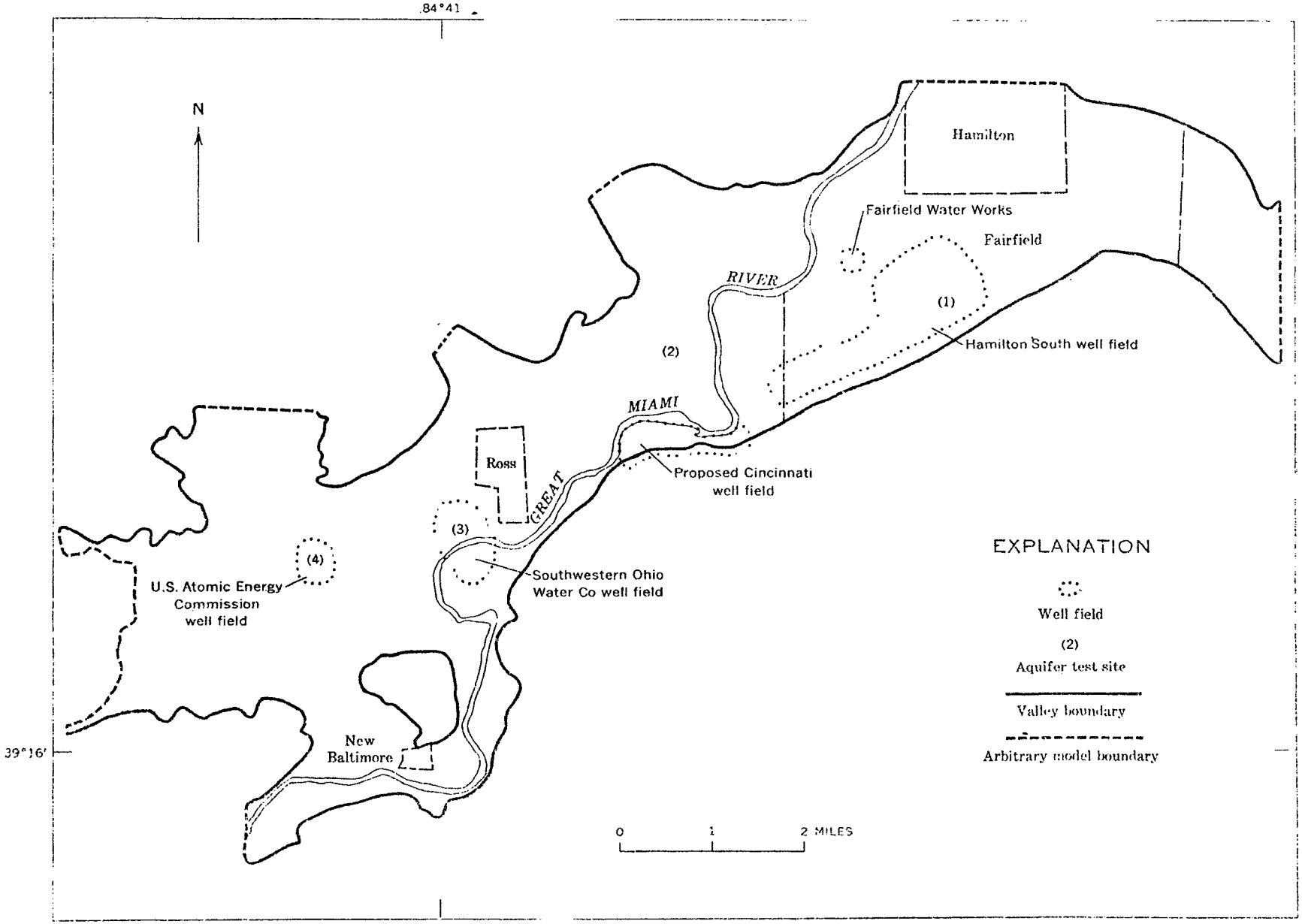
In the western-most portion of the Fairfield-New Baltimore area, which covers about eight square miles, the aquifer is about 200 feet thick and is capped with a complex layer of till, silt and clay.

Groundwater occurs under unconfined conditions throughout most of the area. However, the mathematical condition that the drawdown be small as compared to the saturated thickness of the aquifer is satisfied. This condition permits use of the identification technique developed in the present work.

The hydrologic and geologic characteristics of the Fairfield-New Baltimore aquifer have been extensively studied and a report [Spieker, 1968] provides an excellent source of information for the area.

1.2.1 Estimation of the Input-Output Water Balance

Concerning the hydrologic boundaries (i.e., boundary conditions), the aquifer is bounded by the vertical bedrock wall of the buried Miami Valley. The permeability of this rock is low, yet it can contribute a significant amount of water to the system due to the very large contact area; therefore, a leakage boundary is introduced in the model. A second source of water is provided by the Great Miami River which traverses the aquifer as shown (Figure 1.1). The river strongly interfaces with the aquifer and is one of the most important components of the ground and surface water system.



DESCRIPTION OF THE LOWER GREAT MIAMI RIVER VALLEY, OHIO

FIGURE 1.1

The input-output water balance of the aquifer is made up of the following components:

(i) Recharging of Induced Stream Infiltration

This is a difficult system's input to estimate. It is a highly variable quantity whose interaction with the aquifer depends on many factors such as width and depth of the river, velocity of the streamflow, permeability of the streambed. The most critical of all these factors is the stream infiltration rate under conditions of low streamflow. Two estimates of this factor have been made for the problem area and based on them a range of 240,000 to 500,000 gpd per acre has been determined as the expected range of variation for the maximum infiltration rate all year round [Spieker, 1968]. Such a range indicates that the river is a large source of water for the aquifer; consequently, in the aquifer model the river has been modeled as a constant head boundary.

(ii) Recharge from Boundaries

The perimeter of the aquifer modeled is 220,000 feet, of which 180,000 feet is along the bedrock valley walls. The permeability has been estimated to be on the order of 1.5 gpd per sq. ft. These figures, when multiplied by the total area, yield 6.8 mgd coming from the bedrock formations into the aquifer. This last figure is used in this study.

(iii) Pumping

Pumping is concentrated in three well fields, namely, the Hamilton south well field (Fairfield), the Southwestern Ohio Water Co. well field, and

the U. S. Atomic Energy Commission well field. Pumping started in 1943 with eleven wells in Fairfield. These wells were operated from 1943 to 1945. Then, from 1945 to 1952 there was no significant pumping in the area. In 1952 Southwestern Co. installed a new well, S-1 (Figure 1.1). This well was pumped from 1952 to 1955 at an average rate of 10 mgd. In 1955 a second well was installed, S-2 (Figure 1.1). The combined pumpage of S-1 and S-2 from 1955 to 1962 averaged 13.8 mgd. In 1956 the city of Hamilton installed a new well field (F - 8, F - 10, F - 11) which was pumped from 1956 through 1962 at an average of 7.5 mgd. The U. S. Atomic Energy Commission well field, has been pumped at an average of 1 mgd since 1952.

(iv) Initial Conditions

Records of water level in the area were not kept until pumping had started; therefore, it is difficult to determine the initial conditions of the system. Spieker [Spieker, 1968] estimated those conditions based on existing hydrographs of the area, present water levels measurements, model's results, and river stages. In the present work, initial conditions for groundwater levels in the area were averaged to 550 ft.

For the Fairfield-New Baltimore area only four reliable pumping tests have been performed to determine the aquifer transmissivity. Locations of test points are shown as T_1, T_2, T_3, T_4, (Figure 1.1). The average storage coefficient has been estimated to be 0.145. This is the only available value.

The construction and validation of an aquifer model for the Fairfield-New Baltimore area is an important step in this project since no prediction of the real system behavior can be made without such a component.

Chapter 2

MODELING AND IDENTIFICATION OF AQUIFER SYSTEMS

2.1 Introduction

There are numerous approaches to constructing models which will reliably predict the response of an actual groundwater aquifer to a time history of multiple well pumpages and to estimate natural water inflows and outflows into and from the system. This chapter is intended to provide an introduction to various types of such models. It is important to realize that the predictive ability of a given model is the key feature of value in water resource planning and management. The fact that the hydrologic parameters in the model may not directly correspond to the "micro" physical characteristics of the aquifer soil should not be a disturbing occurence in such models.

In the following sections and in more detail in the following chapters, the models devised and validated as part of the OWRR Project Phase I will be explained and discussed.

2.2 Literature Survey

The ability to forecast water levels and quality distribution in aquifers is essential for planning groundwater management programs. Mathematical models describing groundwater flows are well established. The mathematical models are based on the balances of water volumes which result

in partial differential equations or difference equations. The aquifer system parameters such as transmissivity and storage coefficient are an integral part of the equations and must be known in order to predict the response of the aquifer to various demands placed on it. Since comprehensive physical measurements of aquifer parameters are seldom feasible, the identification of parameter by suitable mathematical formulations is highly desirable.

The problem of identification of the aquifer parameters is equivalent to the problem of parameter identification of partial differential equations (P.D.E.).

Many identification techniques for systems described by ordinary differential equations (O.D.E.) have been developed and are available. The same can not be said of the P.D.E. counterpart. The problem occurs since the theory of P.D.E. is more complex and difficult to apply. Most P.D.E. of interest in engineering have no analytical solutions, and the existing numerical techniques to solve them are not completely satisfactory.

In the area of identification of P.D.E., most techniques focus on the identification of constant-parameter, one-dimensional equations. Few of the techniques consider the identification of variable-parameter, one-dimensional equations, and only very few on the varying-parameter, multi-dimensional equation case.

The survey of the state-of-the-art in identification of P.D.E. has been divided in two sections.

(i) a survey of general techniques of parameter identification of P.D.E.,

(ii) a survey of literature with particular application to groundwater parameter identification.

(i) The best known technique for parameter identification of P.D.E. is the transformation of the P.D.E. into a set of O.D.E. followed by application of quasilinearization, [Bellman & Kalaba, 1969], [Schenke & Haimes, 1973].

Falkenburg [Falkenburg, 1971] identifies variable parameter one-dimensional equations by transforming the P.D.E. into an appropriate integral equation. Using a functional approach, he generates an approximate solution for the distributed system using the integral equation. This approximate solution is then used to identify the equation parameters on a least-square basis. Extensions of this methodology to handle two dimensional P.D.E. have not been devised to date and therefore cannot be applied for the purposes of this project.

Karplus & Kawamoto [Karplus & Kawamoto, 1966] apply sensitivity analysis to identify constant parameters in multidimensional P.D.E. Seinfeld [Seinfeld & Chen, 1971] follows the same approach. The identification problem is posed as a minimization problem. The solution of the P.D.E. is required to match the measured response of the physical system. The parameters are identified on a least-squares basis using a steepest descent method. The main drawback of this approach is the slow convergence rate of the steepest descent method.

Phillipson [Phillipson, 1971] solves the problem of identifying initial and boundary conditions for systems described by linear parabolic and second order hyperbolic P.D.E. He casts the problem within a variational framework

and characterizes extremals of quadratic functionals constrained by P.D.E. by applying known results from the theory of optimal control of distributed parameter systems developed by Lions [Lions, 1971].

(ii) Trial and error procedures and system simulation have been commonly used for identifying transmissivity and storage coefficients in groundwater system modeling. An RC-Network electric analog to simulate the behavior of the aquifer system was employed by Spieker [1968] and Bear and Schwartz [1966] among others. The parameters were determined by trial and error so that, after each run a better agreement was obtained between water levels observed in the aquifer at the end of the calibration period and the corresponding values determined from the model. The work of Vemuri and Karplus [Vemuri & Karplus, 1969] provided an improved method for identifying aquifer parameters. They used a hybrid computer to obtain optimal values of aquifer transmissivity. Their model, however, is unable to determine the storage coefficient except by trial and error.

A unique analytic approach for the identification of aquifer parameters through application of decomposition and multilevel optimization techniques was developed by Haimes et al [Haimes, 1967], [Haimes, Perrine and Wismer, 1968], [Wismer, Perrine and Haimes, 1970]. These investigations involved integral solutions to the flow problem together with decomposition of the aquifer system into a set of independent subsystems. Each of the subsystems was described by a one dimensional partial differential equation with constant parameter. This approach is attractive because of its relative simplicity; however, it may suffer from limited applicability.

Yeh and Tauxe [Yeh and Tauxe, 1971], applied the technique of quasi-linearization to identify the parameters of a homogeneous and isotropic confined

aquifer system. A further extension of this model to a finite leaky aquifer system was studied by Marino and Yeh [Marino & Yeh, 1973]. The applicability of quasilinearization to groundwater modeling and identification is very limited due to problems of high dimensionality and small region convergence.

Kleinecke [Kleinecke, 1971], applied a linear programming to determine groundwater model parameters. A large set of linear equations results from writing the water-balance equations for each cell of a multi-cell aquifer model for which data on water levels are available. Unfortunately, this approach has been found to be very sensitive to errors in measurement [Kleinecke, 1971].

2.3 Aquifer Models

2.3.1 Introduction

The objective of this section is to introduce the aquifer model to be used in this work. The model presented is the well known nonlinear partial differential equation which describes unsteady groundwater flow. The equation is based on Darcy's law and the law of conservation of mass.

Since no analytical solution exists for the continuous model, the P.D.E. is discretized and put into form amenable to numerical solution. The alternate implicit direction method, [Peaceman and Rachford, 1955], is applied to solve it. Finally, a discussion on the model calibration or aquifer model identification problem is presented.

2.3.2 Aquifers

An aquifer is a geologic formation or stratum that can store and transmit significant quantities of water. Under appropriate hydraulic gradients this water can be extracted and used for various purposes. Aquifers can also be regarded as underground storage reservoirs which are replenished by natural and artificial inputs. From a systems viewpoint, an aquifer can be viewed as a black box with an input and an output. The hydrologic input is composed of subsurface inflows and percolation from streambeds, irrigation, artificial recharge and precipitation. The hydrologic output is composed of subsurface outflows, effluent streams, pumpage, and losses such as evapotranspiration and leakage. Aquifers are classified as either confined (artesian) or unconfined (phreatic) aquifers. A confined aquifer is bounded above and below by impervious formations. An unconfined aquifer has a water table as an upper boundary and an impervious formation as a lower boundary.

The physical properties of aquifers are described by such parameters as: hydraulic conductivity (K), which indicates the ability of the aquifer material to conduct water; storativity (S), which measures the storage capacity of the aquifer, etc. For detailed discussion on aquifers the reader is referred to [Bear, 1972], and [Walton, 1970].

Confined and unconfined aquifer systems can be classified as follows:

Confined Aquifers:

Class No.

I $$T\left[\frac{\partial^2\phi}{\partial x^2} + \frac{\partial^2\phi}{\partial y^2}\right] = S\frac{\partial\phi}{\partial t} + Q(x,y,t)$$

II $$\frac{\partial}{\partial x}\left[T_x\frac{\partial\phi}{\partial x}\right] + \frac{\partial}{\partial y}\left[T_y\frac{\partial\phi}{\partial y}\right] = S\frac{\partial\phi}{\partial t} + Q(x,y,t)$$

III $$K\left[\frac{\partial}{\partial x}\left(b\frac{\partial\phi}{\partial x}\right) + \frac{\partial}{\partial y}\left(b\frac{\partial\phi}{\partial y}\right)\right] = S\frac{\partial\phi}{\partial t} + Q(x,y,t)$$

IV $$\frac{\partial}{\partial x}\left[(k\cdot b)\frac{\partial\phi}{\partial x}\right] + \frac{\partial}{\partial y}\left[(k\cdot b)\frac{\partial\phi}{\partial y}\right] = S\frac{\partial\phi}{\partial y} + Q(x,y,t)$$

V $$\frac{\partial}{\partial x}\left[(k\cdot b)\frac{\partial\phi}{\partial x}\right] + \frac{\partial}{\partial y}\left[(k\cdot b)\frac{\partial\phi}{\partial y}\right] = S(x,y)\frac{\partial\phi}{\partial t} + Q(x,y,t)$$

Unconfined Aquifers:

I $$T\left[\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2}\right] = S\frac{\partial h}{\partial t} + Q(x,y,t)$$

II $$\frac{k}{2}\left[\frac{\partial^2(h^2)}{\partial x^2} + \frac{\partial^2(h^2)}{\partial y^2}\right] = S\frac{\partial h}{\partial t} + Q(x,y,t)$$

III $$\frac{\partial}{\partial x}\left(T_x\frac{\partial h}{\partial x}\right) + \frac{\partial}{\partial y}\left(T_y\frac{\partial h}{\partial y}\right) = S\frac{\partial h}{\partial t} + Q(x,y,t)$$

IV $$\frac{\partial}{\partial x}\left[(k\cdot h)\frac{\partial h}{\partial x}\right] + \frac{\partial}{\partial y}\left[(k\cdot h)\frac{\partial h}{\partial y}\right] = S\frac{\partial h}{\partial t} + Q(x,y,t)$$

V $$\frac{\partial}{\partial x}\left[(k\cdot h)\frac{\partial h}{\partial x}\right] + \frac{\partial}{\partial y}\left[(k\cdot h)\frac{\partial h}{\partial y}\right] = S(x,y)\frac{\partial h}{\partial t} + Q(x,y,t)$$

VI $$K\{\frac{\partial}{\partial x}[(h-d)\frac{\partial h}{\partial x}] + \frac{\partial}{\partial y}[(h-d)\frac{\partial h}{\partial y}]\} = S(x,y)\frac{\partial h}{\partial t} + Q(x,y,t)$$

VII $$\frac{\partial}{\partial x}[K(h-d)\frac{\partial h}{\partial x}] + \frac{\partial}{\partial y}[K(h-d)\frac{\partial h}{\partial y}] = S(x,y)\frac{\partial h}{\partial t} + Q(x,y,t)$$

where

$T(x,y,t)$ - Aquifer's transmissivity, over space and time.

$K(x,y)$ - Hydraulic conductivity, over space.

$S(x,y)$ - Storativity of an aquifer.

$b(x,y)$ - Aquifer thickness.

$P(x,y,t)$ - Piezometric head.

$Q(x,y,t)$ - Vertical outflow (inflow if negative).

$d(x,y)$ - Elevation of aquifer bottom.

$h(x,y,t)$ - Elevation of water table.

Each of the above classes of aquifer systems is characterized by its associated parameters. The following tables summarize the variations in the system parameters.

Confined Aquifers:

Class No.	b(x,y)	K (x,y,z)	T(x,y,z)	S(x,y,z)
I	CONST.	CONST.	CONST.	CONST.
II	CONST.	d(x,y)	T(x,y)	CONST.
III	b(x,y)	CONST.		CONST.
IV	b(x,y)	K(x,y)		CONST.
V	b(x,y)	K(x,y)		S(X,Y)

UNCONFINED AQUIFERS:

Class No.	d(x,y)	K(x,y,z)	T(x,y,t)	S(x,y,z)
I	CONST.	CONST.	CONST.	CONST.
II	CONST.	CONST.		CONST.
III	CONST.	K(x,y)	T(x,y)	CONST.
IV	CONST.	K(x,y)		CONST.
V	CONST.	K(x,y)		S(x,y)
VI	d(x,y)	CONST.		S(x,y)
VII	d(x,y)	K(x,y)		S(x,y)

2.3.3 A Continuous Aquifer Model

Aquifer models are vital tools in the planning of groundwater management programs. They help the decision maker to forecast aquifer water levels and quality distributions resulting from alternative policies without having to run expensive and impractical real tests. The following are some important points that must be taken into account when selecting a model, [Schenke and Haimes, 1973].

1. Model's goals and assumptions
2. Geologic and hydrologic properties of the real system.
3. Information available to calibrate the model.
4. Amount and accuracy of input data.
5. The system's output data.

Point 4 refers to the fact that a very sophisticated model is of little value if the accuracy of the input data is poor, whereas point 5 stresses the importance of having the system's output data to check for accuracy of the model response.

There are several mathematical models in the literature, all of which are based on Darcy's law and the law of conservation of mass. Single cell models assume that average conditions describe the entire aquifer. Partial Differential Equation models apply the laws of physics which describe the flow of fluids through porous media, to derive dynamic equations relating aquifer water levels or pressures to forcing functions and aquifer parameters. Finite difference models are either discrete approximations of P.D.E. models or sets of algebraic equations resulting from physical considerations similar to the ones leading to P.D.E. models. Both P.D.E. and discrete models are used in this work. The continuous model will be used in the theoretical developments whereas the discrete model is used in numerical computations. The P.D.E. model is stated below.

The flow in the aquifer is assumed to be essentially horizontal, since the thickness of the aquifer is small compared with horizontal dimensions. If the aquifer properties are nonhomogeneous, isotropic and confined, then the application of Darcy's law and the law of conservation of mass leads to the following relation:

$$\frac{\partial}{\partial x}\left[T(x,y)\,\frac{\partial h}{\partial x}\right] + \frac{\partial}{\partial y}\left[T(x,y)\,\frac{\partial h}{\partial y}\right] = S\,\frac{\partial h}{\partial t} + Q \qquad (2.1)$$

where

T = $T(x,y)$ -- Non-homogeneous transmissivity (ft/sec)

h = $h(x,y,t)$ -- Hydraulic head (ft)

S = $S(x,y)$ -- Coefficient of storage (dimensionless)

Q = $Q(x,y,t)$ -- Pumping rate per unit area (ft/sec)

Equation (2.1) relates changes of water head to pumping in a confined aquifer, (eventually Q could include recharge terms and leakage terms).

The geometry of the aquifer's boundary, the intial head, and the boundary conditions complete the mathematical statement of the aquifer model.

Equation (2.1) also describes phreatic aquifers when the drawdown is small compared to the saturated thickness. Thus, by analogous development and with Dupuit's approximations [Jacob, 1950] result in:

$$\frac{\partial}{\partial x}\left[Kh\frac{\partial h}{\partial x}\right] + \frac{\partial}{\partial y}\left[Kh\frac{\partial h}{\partial y}\right] = S\frac{\partial h}{\partial t} + Q \qquad (2.2)$$

where $K = K(x,y)$ is the hydraulic conductivity of the aquifer (ft/sec) averaged over the aquifer depth. Equation (2.2) relates changes of water head to pumping in an unconfined aquifer.

Equation (2.1) is well suited for the aquifer underlying the Fairfield-New Baltimore area, because the unconfined aquifers in the region satisfy the condition that drawdown is small compared to the saturated thickness.

Since no analytical solution to equation (2.1) exists, a finite difference model is used for the numerical computations.

2.3.4 A Discrete Aquifer Model

There are several ways of discretizing equation (2.1) using finite difference methods, [Ritchmeyer, 1967]. The main method is to replace the domain of the independent variables (x,y,t) by a finite set of points $\{x_i,y_j,t_n\}$. The discretization process is accomplished by superimposing a grid, whose shape is usually rectangular in order to obtain simple equations. The water head is required to satisfy difference equations obtained by replacing partial derivatives with difference approximations. There are two types of finite-difference schemes: explicit and implicit.

In an explicit scheme, the water head at grid point (x_i,y_j) at time t_{n+1} is calculated explicitly in terms of known values of the head (at surrounding grid points) at time t_n. In an implicit scheme the water head at grid point (x_i,y_j) at time t_{n+1} is a function of known and unknown values of the head at time t_n and t_{n+1}, respectively.

Explicit schemes are easy to solve, but they impose severe requirements on the grid sizes and time steps to control stability problems. A typical time step constraint is given by, [Bear, et al, 1972]:

$$\Delta t \leq \frac{S}{2T} \frac{(\Delta x \cdot \Delta y)^2}{(\Delta x^2 + \Delta y^2)} \tag{2.3}$$

where Δt is a time step, $\Delta x \cdot \Delta y$ is the area of a grid cell, S is storage and T is transmissivity. Condition (2.3) imposes a small time step, expecially in confined aquifers with a small storage coefficient and a large transmissivity. For problems extending over large values of time, this could result in an excessive amount of computation.

The implicit method overcomes the above difficulties at the expense of a somewhat more complicated calculational procedure. It generates a large set of simultaneous equations which are solved with iterative techniques that may also be time-consuming. However, the implicit schemes are unconditionally stable, and the size of the time step is limited only by considerations of convergence rate and accuracy.

The Alternate Implicit-Direction (A.I.D.) method, [Peaceman and Rachford, 1955], avoids these disadvantages and yet still manages to use a system of equations with tridiagonal coefficient matrices instead of a single five-triagonal matrix resulting from the implicit method. Matrices can be inverted directly using Thomas' algorithm [Peaceman & Rachford, 1965]. Essentially the A.I.D. method employs two difference equations which are used in turn over successive time-steps, each of duration equal to half of the step size. The first equation is implicit only in the x - direction

and the second is implicit only in the y - direction. In this study, the alternate implicit direction method is used in the finite difference approximation of aquifer model. Application of the A.I.D. method to equation (2.1) can easily be developed.

For simplicity T and S are assumed constants and Q zero. Then, equation (2.1) becomes

$$\frac{T}{S}\left[\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2}\right] = \frac{\partial h}{\partial t} \qquad (2.4)$$

The time interval Δt is divided into halves. In the first time sub-interval, $\frac{\partial^2 h}{\partial y^2}$ is replaced by a second difference approximation evaluated in terms of the unknown values of the water head h, whereas $\frac{\partial h^2}{\partial y^2}$ is replaced by a second difference approximation in terms of known values of h. The following equation is obtained:

$$\frac{T}{S}(\Delta x)^2[h_{ij-1jk} + \frac{\Delta k}{2} - 2h_{ijk} + \frac{\Delta k}{2} + h_{i+1jk} + \frac{\Delta k}{2} +$$

$$\frac{T}{S}(\Delta x)^2[h_{ij-1k} - 2h_{ijk} + h_{ij+1k}] = \frac{2}{\Delta k}[h_{ijk+\frac{\Delta k}{2}} - h_{ijk}] \qquad (2.5)$$

where h_{ijk} denotes the water head evaluated at the grid point at time k. Δk denotes time interval. Equation (2.5) is implicit in the x direction.

For the second time sub-interval, $\frac{\partial^2 h}{\partial y^2}$ is replaced by a second difference approximation in terms of unknown values of h. The rest of the substitutions are similar to those leading to (2.5). The following

equation is obtained:

$$\frac{T}{S}(\Delta X)^2\left[h_{i-1jk+\frac{\Delta k}{2}} - 2h_{ijk+\frac{\Delta k}{2}} + h_{i+1jk+\frac{\Delta k}{2}}\right] +$$

$$\frac{T}{S}(\Delta y)^2\left[h_{ijk-1k+\Delta k} - 2h_{ijk+\Delta k} + h_{ijk+1k+\Delta k}\right] =$$

$$\frac{2}{\Delta k}\left[h_{ijk+\Delta k} - h_{ijk+\frac{\Delta k}{2}}\right] \qquad (2.6)$$

Equation (2.6) is implicit in the y direction.

Each grid point generates two equations similar to (2.5) and (2.6). A set of row grid points generates a set of coupled equations implicit in the x direction which can be arranged as follows, [Peaceman and Rachford, 1955]:

$$\text{(Boundary Points)} \quad B_1h_1 + C_1h_2 = D_1$$

$$\text{(Interior Points)} \quad A_ih_{i-1} + B_ih_i - C_ih_{i+1} = D_i \quad 2 \leq i \leq n-1 \qquad (2.7)$$

$$\text{Boundary Points)} \quad A_nh_{n-1} + B_nh_n = D_n$$

where A, B, C, D are constants and the subscripts denote row points. A similar equation can be obtained for a set of column grid points.

For a given time step $k + \Delta k$, a solution $h_{ijk+\Delta k}$ is obtained as follows:

1. Solve equation (2.7) for every row (x direction in the grid). The result is an intermediate solution $h_{ijk+\frac{\Delta k}{2}}$.

2. Using the partial result in 1, above, solve an equation similar to (2.7) for every column (y direction). The result is $h_{ijk+\Delta k}$.

The sizes of the time steps and spatial increments in equation (2.5) are determined by considering trade-offs between convergence and computational efficiency. Experience indicates the size for each case.

To complete the discretization of equation (2.1) the boundary conditions should be properly handled. Some techniques are considered below:

1. For a grid size on the boundary, a grid refinement is usually desirable to follow closely the geometry of the aquifer boundary; however, this destroys the computational simplicity of equation (2.7). A tradeoff between accuracy and computational efficiency leads to steps 2 through 4 below.
2. Impervious boundaries, for example, can be handled by setting T=0 at those grid points which lie outside the boundary.
3. Constant head boundary nodes are handled by assigning an initial constant waterhead.

Finally, the assumption that T and Q are constant is dropped and the complete discrete version of equation (2.1) is obtained by following similar steps to those leading to the discretization of (2.4). The equation, implicit in the x direction, is

$$T_{\underline{ij}}h_{i-1jk} + \frac{\Delta K}{2} - [T_{\underline{ij}} + T_{\bar{ij}} + P - I_{ij}^{(n)}]h_{ijk} + \frac{\Delta k}{2}$$

$$+ T_{\bar{ij}}h_{i+1jk} + \frac{\Delta k}{2} = \tag{2.8}$$

$$-T_{\underline{ij}}h_{ij-1k} + [T_{\underline{ij}} + T_{i\bar{j}} - P - I_{ij}^{(n)}]h_{ijk}$$

$$- T_{i\bar{j}}h_{ij+1k} + Q_{ij}$$

The equation, implicit in the y direction, is

$$T_{\underline{ij}}h_{i-1jk} + \Delta k - [T_{\underline{ij}} + T_{i\bar{j}} + P - I_{ij}^{(n)}]h_{ijk} + \Delta k$$

$$+ T_{i\bar{j}}h_{ij+1k} + \Delta k = \tag{2.9}$$

$$-T_{\underline{ij}}h_{i-1jk} + \frac{\Delta k}{2} + [T_{\underline{ij}} + T_{\bar{ij}} - P - I_{ij}^{(n)}]h_{ijk} + \frac{\Delta k}{2}$$

$$- T_{\bar{ij}}h_{i+1jk} + \frac{\Delta k}{2} + Q_{ij}$$

where,

$I_{ij}^{(n)}$ denotes the iteration parameter at the $i^{\underline{th}}$ & $j^{\underline{th}}$ node at the $n^{\underline{th}}$ iteration and is defined as

$$I_{ij}^{(n)} = I_{ij}^{(n-1)} \quad \exp\left[\frac{\ln \frac{2xm}{\pi}}{\lambda - 1}\right] \tag{2.10}$$

and,

λ = the number of iteration parameter desired,

x_n = the larger of the total number of rows or total number of columns of the grid.

$$I^{(1)} = \frac{\Pi^2}{2xm} [T_{i - 1/2\ j} + T_{i + 1/2\ j} + T_{ij + 1/2}]$$

Computational details on how to use equations (2.8) and (2.9) to forecast water levels are given in Appendix A. A digital program implementing these equations is presented there.

where $\underline{i} = i - 1/2$, $\bar{i} = i + 1/2$ and similar expressions for $\underline{j}$, $\bar{j}$ apply,

also $P = \frac{\Delta X^2 S}{\Delta t}$,

Assumptions made to simplify the formulations above were:

$$\Delta X = \Delta Y, \qquad S = \text{constant}$$

To increase accuracy, transmissivity is expressed in half-node expansions:

$$T_{\underline{i}j} = \frac{T_{i-1j} + T_{ij}}{2}$$

$$T_{i\underline{j}} = \frac{T_{ij-1} + T_{ij}}{2}$$

Similar expressions apply for $T_{\bar{i}j}$ and $T_{i\bar{j}}$.

Equations (2.8), (2.9) are similar to (2.5), (2.6), except for the
that known values have been arranged on the right-hand side of the
:r equations.

The iteration parameter $I_{ij}^{(n)}$ was introduced in eqns. (2.8) and (2.9)
iprove the accuracy of the A.I.D. method by cycling the solutions
n a time step until the difference $(h_{ijk+1}^{(n+1)} - h_{ijk+1}^{(n)})$ for all i's
's becomes less than a convergence factor, [Bredehoeft & Pinder, 1970]
Rubin, 1968].

› The Aquifer Identification Problem

To use the models described by (2.1) and (2.7) in order to forecast
:er water levels, the following information should be obtained:

1. The transmissivity function T
2. The storage function S
3. The forcing function Q
4. The aquifer initial conditions (i.c.)
5. The aquifer boundary conditions (b.c.)

rmination of the above five types of input data or parameters comprises
aquifer system identification problem. Identifying each of these
s of information represents a difficult problem. For example, identifying
quires the determination of pumpage and recharge patterns, rain infiltra-
, river, lakes, percolation, leakages, losses, in order to make a
r-balance of the total water input into the aquifer. Conditions (i.c.,b.c.)
similar level of involvement is the determination of the aquifer initial
boundary. In the literature, this latter problem is called the state

identification problem. A solution approach to this problem is presented by Phillipson [Phillipson, 1971]. Transmissivity and storage are highly variable discrete distributed parameters. This is largely due to the great variety of geological materials and structures forming an aquifer. Such irregularities pose serious problems to the identification of aquifer model transmissivity and storage or, as it is known in the literature, the aquifer identification problem.

Generally speaking, the above five types of information are related to each other and can be considered as a single problem composed of many subproblems. The identification model addresses itself to a single sub-problem, i.e., to identify the aquifer model transmissivity function. It is assumed that storage, i.c., b.c., and Q are already known. The problem is stated as follows:

Given the aquifer initial and boundary conditions, given the storage coefficient, given well pumpage records and an associated historical record of water levels in the aquifer, estimate the value of T (aquifer model transmissivity) on the basis of these records, using curve fitting criteria. Some of the factors which complicate the solution to this problem are:

1. The historical records do not cover an entire basin, and, as a consequence, an overall distributed parameter function is being estimated from non-uniformly distributed data.
2. Initial and boundary conditions are difficult to determine.
3. Great difficulties are involved in the accurate estimation of the input function Q. Most of the aquifer water sources are random variables.

4. Considerable noise is associated with the data measurements.

Of necessity, certain simplifications such as the assumptions of noise less data measurements, that boundary conditions are given, and that the aquifer water balance has been estimated, have to be made if insight is to be gained into the problem. The developments in following chapters should be judged bearing these simplifications in mind.

2.4 Aquifer System Identification

The identification problem in groundwater hydrology involves determining the distribution of parameters which characterize an aquifer from observations of pumping and recharge rates, flows at boundaries, and water levels. Most aquifers are heterogeneous. Thus, a realistic representation must include distributed parameters characterizing the *normal water system*.

The basic flow equations which characterize the system model are well established. Darcy's Law describes the slow flow of an incompressible fluid through the porous medium. Usually the gradients involved are small, and when continuity is imposed on a confined aquifer or reservoir system, linearization can be employed. The result, is that the transient pressure response of the system can be described by the diffusion equation (2.1). Only two space dimensions are considered because with typical systems vertical flow seldom is important. Waterhead is denoted by h, and Q represents a source strength (production rate per unit area). The coefficients in the equation characterize the porous medium. Transmissivity,

$T(x,y)$, is a measure of the ease with which fluid moves through the system. The storage function, $S(x,y)$, is a measure of system capacity. Both are spatially distributed parameters.

At this point appropriate boundary conditions, including production rates, and values for T and S must be specified to be able to predict future system response. A difficulty, of course, is that detailed knowledge of the variation of $T(x,y)$ and $S(x,y)$ is not available. On the hand, waterhead and time data, h and t, respectively, can be obtained by observing the system at specified locations. Thereby, an inverse problem in the aquifer description is created: given some function $F(h_{observed} - h(T,S)_{calculated})$, how must T and S be chosen so that F is minimized? A solution to the inverse problem enables one to accurately predict system response to future modes of operation.

Thus, it is assumed that a useful description of the system is given by specifying a number of transmissivity and storage values: T_i and S_i which will minimize an appropriate criterion function.

Two basically different approaches may be utilized to achieve useful representations for the spatial distribution of properties in the system. One approach is to subdivide the basin into a finite number of areas of specified geometry, each of which is assumed to be homogeneous with respect to transmissivity and storage. The simplest such case is the analysis of a lumped system, for which the entire aquifer is considered to be homogeneous with respect to these characteristics. The second approach defines aquifer properties through a functional relationship which provides the spatial variation.

There are many ways in which each of these two basic approaches can be implemented. Using the first approach, areas can be defined as rectangular spaces (the typical grid for a discretization), other polygons, wedge-shaped areas converging at the center of a well field, annular rings centered on individual wells, or less regular geometries defined by geological boundaries, for aquifer subsystems. The complexity of the identification problem is determined in this instance largely by the number of such distinct spatial regions considered and hence the number of parameters to be identified. A second factor to be considered, which may be far from trivial, is whether or not the aquifer simulation problem implied by the geometrical configuration is tractable.

In the second approach, the infinity of mathematical functions which might approximately represent transmissivity and storage can be reduced to a polynomial representation. In this instance the complexity of the identification problem largely rests upon the degree of the polynomial selected, and thus the number of coefficients which must ultimately be determined.

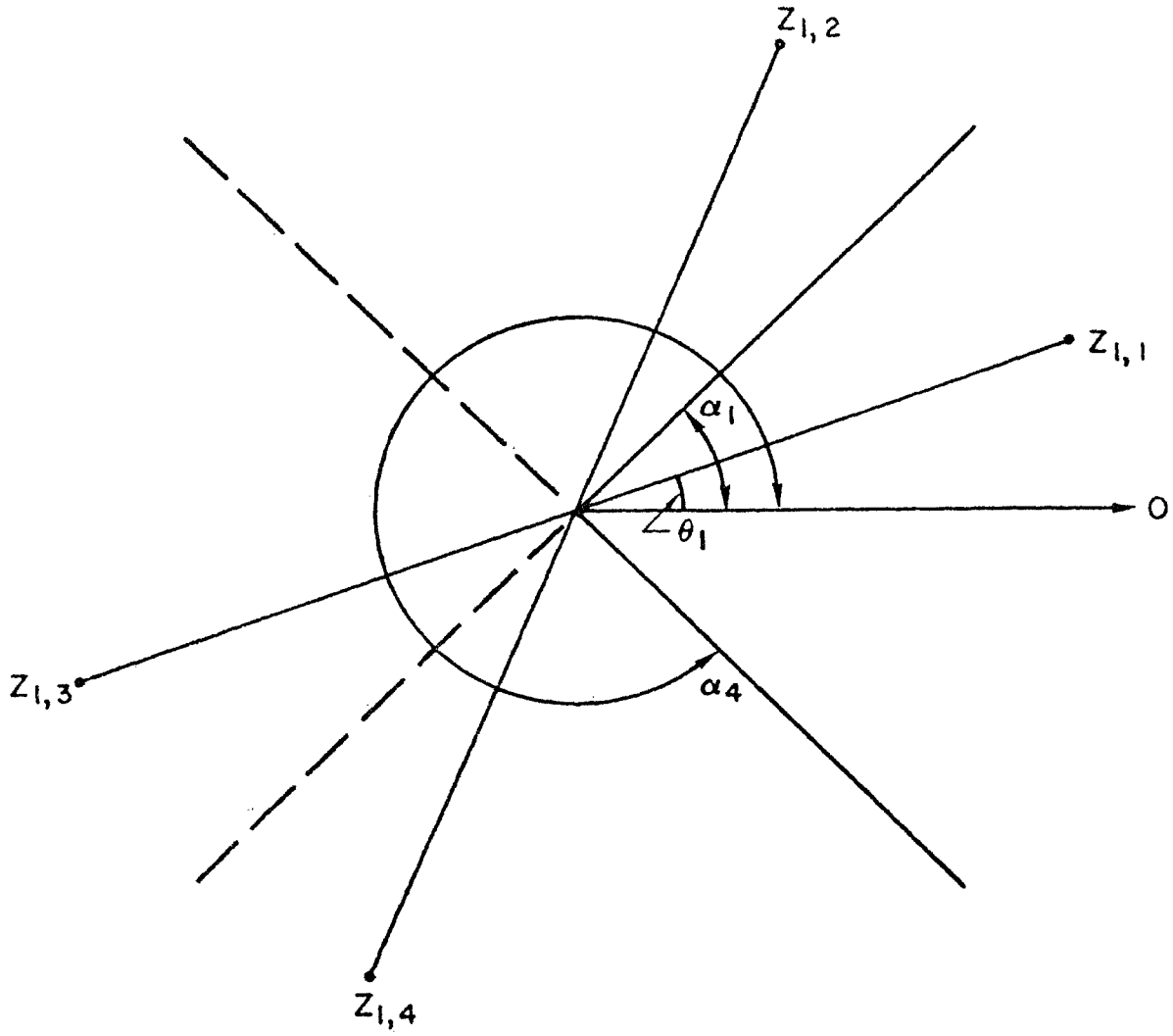
In the course of this work several approaches to the identification problem have been carefully reviewed and considered. Some are the outgrowth of earlier work by the present authors and some are primarily the result of efforts by other investigators. In the following sections of this report, both the philosophy and physical reasoning underlying behind these alternatives and the selection from them is given. The new model which was developed during this study is presented in Chapter 3.

2.4.1 Model One

Consider the characteristics of a typical large aquifer. Quite often fluids are produced through a cluster of wells located near its center. Starting from any established initial condition, over reasonably large periods of time no effect of system boundaries is likely to be felt. The system is large in area compared to its vertical dimensions, and hence can be represented as an infinite two-dimensional system containing a cluster of wells in a region of primary interest, [Haimes, 1967] and [Haimes et al, 1968]. In the complex domain, coordinates for N producing wells (see Figure 2.1) are as follows:

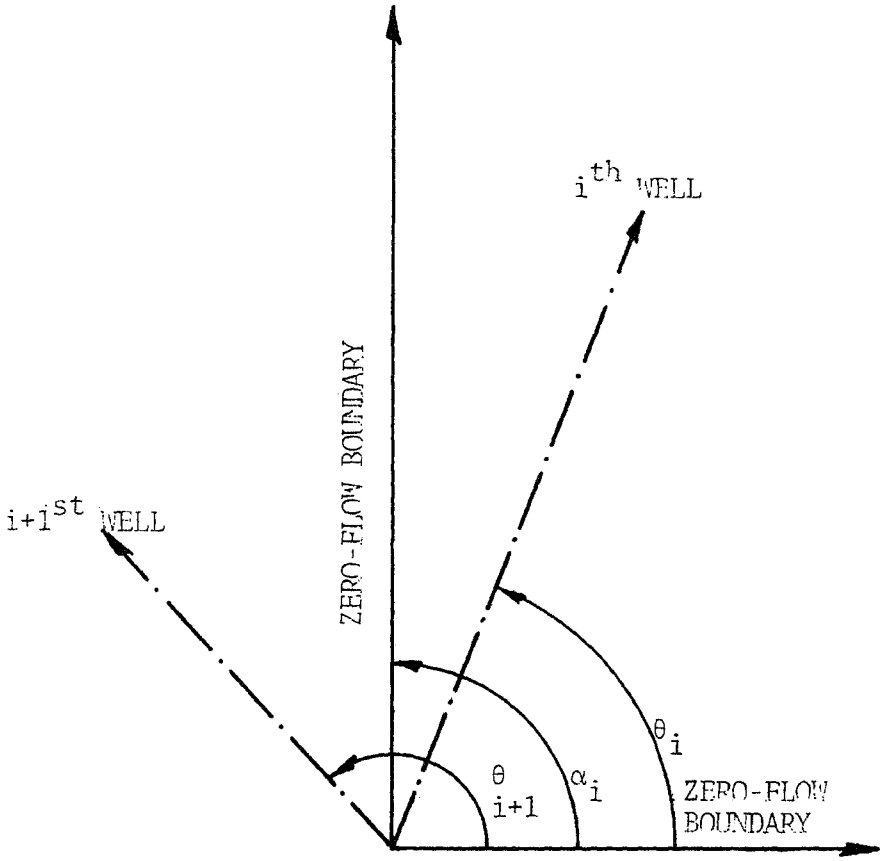
$$Z_i = r_i e^{j\theta_i}, \quad i = 1,2,\ldots,N \quad j = \sqrt{-1} \tag{2.11}$$

Water stored in the region midway between two producing wells could flow to either. Actually, at any time, there exists a line lying between any pair of producing wells across which no flow occurs. On opposite sides of this line flow will occur toward different wells of the pair. Thus, when one considers a well cluster in a large aquifer, at the edge of the well cluster each such line extends out toward the infinitely distant boundary. Mathematically, the effect of each such "no flow" line is exactly the same as if the line constituted an impermeable physical boundary (See Figure 2.2). Also, because of convergence effects, system behavior as reflected by head measurements in the producing wells, is determined largely by the hydraulic properties of the aquifer near the production well. Thus the sensitivity of the response to system properties at large distance



REPRESENTATION OF THE AQUIFER
SYSTEM GEOMETRY

FIGURE 2.1



MODEL ONE GEOMETRY

FIGURE 2.2

from the well cluster is small.

Viewed together, these properties of the system suggest that a very simple model might prove useful for computing the hydraulic characteristics of the aquifer in the vicinity of each producing well. A logical decision is to represent the no-flow line between wells as if it were an impermeable boundary which for the simplest applicable geometry is a straight line. (While a more realistic model would bound each segment by an arbitrary curve, the resulting model would be intractable.) But two such boundary lines, on opposite sides of a single well, will intersect to form a wedge. And so, for our cluster of wells in an infinite aquifer, we arrive at the following analytically tractable yet reasonable model. The aquifer is divided into N wedge-shaped homogeneous regions, each enclosing a single producing well (see Figure 2.1 for four producing wells). (A trivial extension could include several wells within a single region.) The regions are separated by N straight, impermeable boundaries radiating from the origin at azimuths by N straight, impermeable boundaries radiating from the orgin at azimuths α_i. Hence $\theta_i < \alpha_i < \theta_{i+1}$, $i = 1,2,\ldots,N$, where $\theta_{N+1} \equiv \theta_1$. No *a priori* knowledge of the location of effective flow boundaries between wells is possible. Thus, the response of the system will be permitted to select the optimal location of each such no-flow line between wells. Optimal values of transmissivity and storage coefficient also will be selected within each region.

Each wedge of the aquifer lying between no-flow boundaries, can be treated as separate sub-system. The location of the boundaries, however,

affects more than one well, and thus links the several subsystems into a total aquifer system. A mathematical solution for the response of the system results when the continuous line source solution is applied. Boundary conditions for a single wedge-shaped subsystem are satisfied approximately by employing a finite number of image wells, where $Z_{i,k}$, $k = 1,2,\ldots,m_i$, restricted to even integers, is the number of images corresponding to the ith production well. Note that $k = 1$ refers to the producing well. The system of producing and image wells extends throughout the entire plane of the aquifer, also the following relation must be met:

$$\alpha_i - \alpha_{i-1} + 2\pi \cdot \delta_{i,1} = \frac{2\pi}{m_i} \quad i = 1,2,\ldots,N \tag{2.12}$$

where

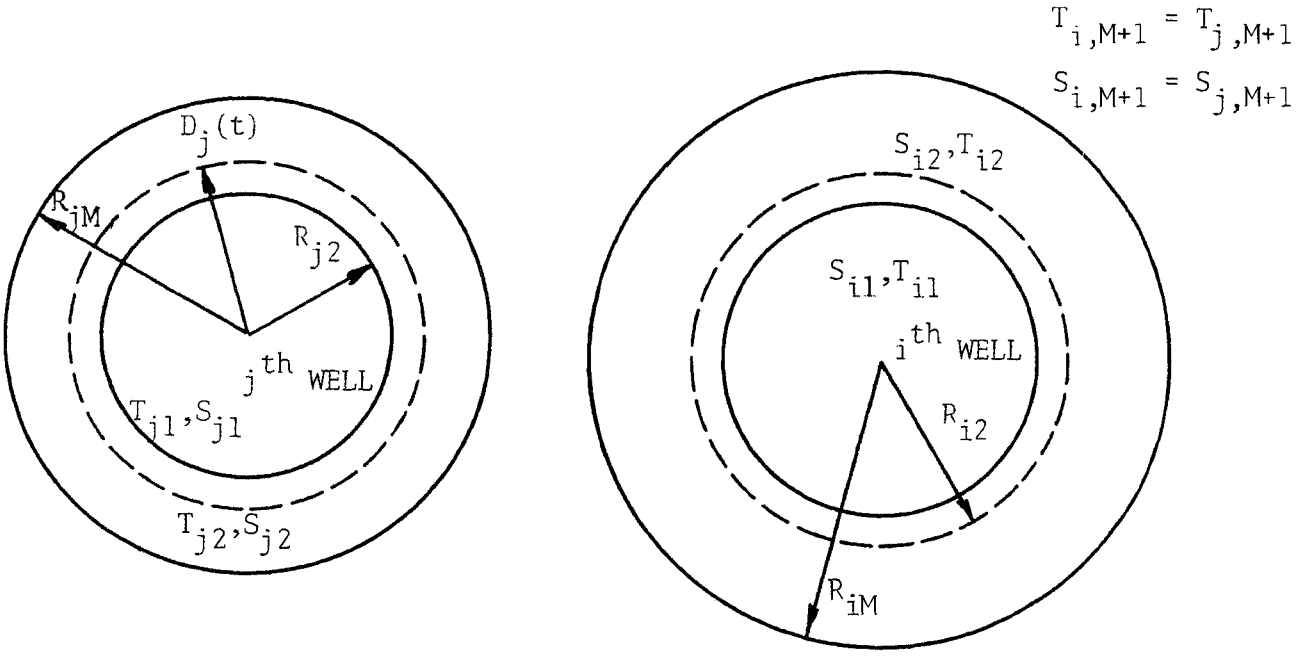
$\delta_{i,1}$ is the kronecker delta and $\alpha_o \equiv \alpha_N$.

The parameters which represent transmissivity and storage generally will change abruptly on crossing a "boundary." This is no real limitation to the identification scheme, however, once the logical basis for the choice of the model is understood and accepted. Parameters determined in this way represent effective values, established on the basis of system response, within a defined spatial region. Near-well regions have a disproportionate influence on the results obtained, and so the results might be expected to vary from segment to segment. A strength of the model is that detailed geometry is actually determined by system behavior. Representative model geometry is illustrated in the figures.

2.4.2 Model Two

The second basic system model which has been considered takes particular advantage of the fact that waterhead response is more strongly influenced by near-well properties than by those further away [Wismer, Perrine & Haimes, 1970]. In addition, early response is controlled solely by near-source properties, and as time goes on properties further out are reflected with gradually diminishing importance. Accordingly, a useful conceptual model starts by specifying annular regions concentric with each well, such that constant effective parameter values can be used within each region. An external region, outside the last defined ring around any well, must extend to the system boundaries (or to infinity). This external region is common to all wells. Thus, at some distance from any one well a set of uniform system properties, common to all subsystems, is assumed. This model geometry is illustrated in Figure 2.3.

This aquifer model has several unique characteristics. Concentric rings around any one well, with defined effective properties, also form part of the external region surrounding the outer ring of any one other well. Thus, to the extend that each ring contributes a measure to the overall effective average properties of the system, the parameter values obtained are not single-valued. In other words, in a way which one might expect to be much like the observed response of the actual aquifer, any one well "sees" the detailed variation in properties immediately surrounding a second well only as part of the distant, system-average properties. The subsystems represented by individual wells again are



MODEL TWO GEOMETRY

FIGURE 2.3

related in a simple way. In this instance they are linked by the common external region, and its common set of properties.

Another important characteristic becomes apparent when fluid is removed at a constant rate. At any time there is a radial distance from each well beyond which measurable reduction in waterhead has not yet occured, and over most of the region out to this point, waterheads fall almost uniformly with time, after an initial transient period. This characteristic leads to important simplifications in actual solution of the differential equations required to simulate the aquifer using this model.

2.4.3 Model Three

The first two modeling approaches described above both represent subdivisions of the aquifer into spatially defined regions within which properties such as transmissivity and storage are uniform. The present model represents the other approach, in which aquifer properties are defined in space by means of a functional relationship. Because a very substantial part of this year's progress has been made using such a model, and it will be thoroughly described in subsequent portions of the report, no details are presented.

2.4.4 Commentary on Model Selection

In the above discussion the need for several simulation models was considered. It might be well asked why not proceed simply to determine

which model is "best," and then exploit that model to increase our knowledge of the groundwater system.

There are a number of reasons for the approach which has been taken. First of all, to establish which modeling approach is "best" would require establishing restrictive criteria on which the determination would be based. Should the criterion for model selection be based on the simulative computational efficiency? The details which can be presented? The simplicity of the identification procedure? The flexibility of the model to changes and adaptation? Each may lead to a different conclusion and all such considerations are of importance.

There are over-riding factors which would dictate consideration of at least several class-representative models. Almost any model likely to yield useful results will tend to be complex, and require substantial computational requirements when ultimately used as part of an optimum management scheme. For this reason every effort should be made to keep the model as simple as possible. Conversely, any model for which results can be obtained with reasonable limits on computer time will necessarily be an oversimplified representation of the real system. Thus, we always tend to seek more detailed views of the subject aquifer which utilizes more efficient computational techniques.

In this situation the "best" result may achieved by joint use of two basically different models, each computationally efficient through the identification process and resting on sound physical grounds. As an analogy, consider the problem of defining the topography of a mountain. Two conventional photographs, one from the front and one from the side,

together provide much more information than either by itself, and yet the two together are much more readily obtained than any three-dimensional representation we can construct.

In a completely parallel way, a very substantial increase in information obtained may accrue from the simultaneous use of several simple aquifer models. One would expect that these benefits would also improve subsequent optimization of aquifer management.

As an example, consider an aquifer in which transmissivity shows a distinct trend in a specified direction. At the same time, three distinct lobes of altered transmissivity immediately surround the well field. In addition, the directional trend in transmissivity shown by the lobes is the opposite of the superimposed, groundwater basin-wide trend. Choosing to model the system using Model One, the existence of and direction of the lobes altered transmissivity are correctly identified, but the over-riding trend is missed entirely. Use of Model Three, on the other hand, readily picks up the trend but misses critical detail near the well field. At the same time, any one function flexible enough to encompass all this variation would lead to an impossibly complex system identification problem.

The obvious conclusion is that there is much to be gained from a thoughtful investigation of several models within a study such as this. In the past year, primary emphasis has necessarily been placed on only one approach. Including the work planned under Phase II, the total project will include consideration of all alternatives discussed within this section.

Chapter 3

GROUNDWATER IDENTIFICATION MODEL - MODEL 3

3.1 Introduction

The groundwater identification model and solution algorithm developed in this research project were first tested on a simplified finite dimensional example problem referred to as problem B. Implementation and testing of the algorithm on a real aquifer were then performed successfully using data from the Miami Conservency District in Dayton, Ohio. The solution of the model identification problem is referred to in this report as the identification algorithm (I.A.). The development of this algorithm constitutes the main objective of this chapter.

The structure of Chapter 3 includes a formal, brief statement of the identification problem presented in Section 3.2. In Section 3.3 a simplified problem (Problem B) is formulated for the groundwater transmissivity identification problem. The reduction in dimensionality (from an infinite dimensional space to an Euclidean space) is achieved by assuming a known transmissivity function where specific parameter must be identified.

The problem formulation is carried over within a static optimization framework. This results in the minimization of a nonlinear function subject to a nonlinear constraint.

In Section 3.4 the static optimization problem is approximately solved by using the approach developed in Appendix A. First a quadratic problem which approximates the original one is studied. Then, in Section 3.5, the

idea of approximating the original problem by a successive set of quadratic approximations is implemented by using Marquardt's Algorithm [Marquardt, 1963]

Section 3.6 summarizes the main ideas of Chapter 3 and Appendix A by discussing the similarities and differences between the simplified and real (complicated) identification problems.

3.2 Statement of the Model Identification Problem

The model of concern is described by equation (2.1), repeated for convenience below:

$$\frac{\partial}{\partial x}\left(T\frac{\partial h}{\partial x}\right) + \frac{\partial}{\partial y}\left(T\frac{\partial h}{\partial y}\right) = S\frac{\partial h}{\partial t} + Q \qquad (3.1)$$

$$h(x,y,o) = h_o \qquad (3.2)$$

$$\frac{\partial h}{\partial n}\Big|_{r_1} = \phi, \text{ and/or } h(x,y,t)\Big|_{r_2} = h_1 \qquad (3.3)$$

$$\text{where } Q \in R$$

where (3.2) is the initial condition of the system, (3.3) the boundary conditions, which may be constant flow and/or constant head, r_1 and r_2 denote the boundary geometry, R is the domain of (3.1)-(3.3) defined as $\Omega x[0,T]$, where Ω is the aquifer's area, and t represents time. It is assumed that all parameters and functions in the model are known except for T(x,y), the model's transmissivity function. Initial and boundary conditions are also assumed known.

The model described by (3.1)-(3.3) is not completely determined because the function T(x,y) is unknown; therefore, the question arises as how to determine T(x,y). The identification of the function T(x,y) for a specific aquifer system is, of course, known as a parameter identification, system identification, parameter estimation, or model calibration. The goal is to find that T(x,y) for which both the response of the real system and the model under the same input is as "close" as possible (usually closeness is measured by a least-squares norm). The response of the model will never exactly match that of the real system because of the many assumptions required in developing the model. After-all an aquifer model is only a simplification of a complex reality.

In groundwater studies the analyst is usually supplied with an arbitrary input-output historical record and his solution approach has to be flexible enough to adjust to the information available. Subsequently, in designing a calibration method he must seek a technique applicable to any arbitrary input information, deterministic inputs, and noise corrupted outputs. It is advantageous if the method would not put any restrictions either on the number of test wells required or on their placements.

These attributes were considered in developing the calibration technique discussed in this chapter; however, it is felt that more computational experience will be necessary before one is able to determine the minimum levels of information required and possible accuracy restrictions imposed on the number and placement of test wells.

A least-squares norm of the output error, namely the error between observed and calculated waterheads, is selected as a loss function. This

norm is mathematically tractable and possesses some desirable properties, such as assigning equal weights to both positive and negative deviations and penalizing large errors with large weights.

Now the parameter identification problem (P.I.P.) can be formulated. Given (i) a class of systems described by equation (3.1)-(3.3), input-output information denoted here by $\hat{Q}(x,y,t)$ (mainly pumping and recharge) and $\hat{h}(x,y,t)$ (real aquifer output), respectively, and (iii) an output error criterion defined by

$$J_1(T) = \int_0^t \int_\Omega [h(x,y,t;T) - \hat{h}(x,y,t)]^2 dt d\Omega \qquad (3.4)$$

where $h(x,y,t;T)$ is the model (waterhead) output for a given o function $T(x,y)$ and $\hat{h}(x,y,t)$ is the aquifer observed output (waterhead).

The parameter identification problem is to minimize (3.4) with respect to $T(x,y)$. To complete the formulation of the problem it is necessary to define the functional space where $T(x,y)$ is to reside. Initially, T will be limited to the space of positive real functions and possibly (because of (3.4)) to having its second powers integrable in Ω. Assuming the system is identifiable, i.e., the characteristics of $T(x,y)$ can be retrieved from input-output observations of the real system, the question is: How can the observation data be used to obtain an "accurate" estimate of $T(x,y)$?

Equations (3.1) - (3.4) represent a variational problem, that of searching for a function $T(x,y)$ which minimizes the non-linear functional (3.2) under constraints (3.1) - (3.3). The solution of such a problem depends strongly on the solution properties of the constraint equations. However, the theory of non-linear P.D.E. is not well developed yet.

Consequently, the derivation of existence theorems as well as necessary and sufficient conditions for $T(x,y)$ to be extremum of (3.2) is still an open question. Existence theorems for a similar problem to the one defined by (3.1) - (3.4) have been studied by Lions [Lions, 1971] and Cesari [Cesari, 1972].

Due to the above limitations, an approximate solution to the identification problem (3.1) - (3.4) is developed. Linearization of (3.1) - (3.3) transforms (3.2) into a quadratic functional in $T(x,y)$. The minimization of quadratic functionals subject to linear constraints is a well understood problem with many available solution procedures. After linearizing (3.1) - (3.3) around nominal values of $h = h_o$ and $T = T_o$, an auxiliary objective functional quadratic in the perturbation functions, δ and δT is formulated. The minimization of this auxiliary objective functional is performed subject to a linearized P.D.E. related to (3.1) - (3.3). Letting $\delta h = u$ the control variable of a dynamic optimization problem is solved by least-squares techniques in Hilbert space. A detailed development of the procedure is presented by Lopez [Lopez, 1973] and mathematical justification is provided by Lions, [Lions, 1971].

3.3 Parameter Identification as a Finite Dimensional Problem (Problem B)

Reduction in dimensionality of the identification problem demands knowledge of the structure of the transmissivity function within a set of parameters. The problem of selecting such a structure is discussed below.

Assuming that the model equation (2.1) accurately represents the characteristics of the real system, then a perfect simulation would be attained if the exact value of the aquifer transmissivity were substituted in the model. However, (2.1) is just an approximation of the real system, and therefore the model's T(x,y) is not expected to coincide with the real transmissivity. In fact, a discrepancy between them may be required in order to compensate for approximations introduced in other parts of the model. From a physical viewpoint, the transmissivity indicates the capacity of the aquifer to transmit water through its entire thickness. It is defined as

$$T = K D \tag{3.5}$$

where

K = coefficient of permeability (ft/sec.)

D = saturated thickness (ft)

K is a function of the geologic characteristics of the aquifer and therefore in non-homogeneous aquifers K is a function of x,y and z. In confined aquifers, if D is approximately constant and K is usually averaged on the aquifer depth, yielding a T which is a function of x and y only

$$T(x,y) = K(x,y)D \tag{3.6}$$

In general, T(x,y) is a non-continuous function whose structure depends on the aquifer geology. Since the geologic characteristics change from

aquifer to aquifer, it is not possible to establish a general expression for T(x,y) based on physical grounds. For specific cases, determination of the structure of T(x,y) demands complete knowledge of the geology of the aquifer. It is the recognition of the difficulties involved in determining T from physical measurements which forces the hydrologist to adopt indirect methods.

In summary, it does not seem feasible to determine a general representation for T(x,y) based on physical grounds. Therefore, it is decided to represent T(x,y) with some general mathematical structure capable of approximating the real structure as close as possible.

A polynomial representation of T(x,y) is proposed. The degree of the polynomial as well as the coefficients are selected using a least-squares criterion. Specifically, the identification problem consists of determining both the degree and the coefficients of the polinomial which represents T(x,y). Clearly, the higher the degree of the polynomial the more complex the identification problem becomes.

Let it be assumed that T(x,y) belongs to the space of positive polynomials in x and y. For simplicity in the developments, assume T(x,y) is given by

$$T(x,y) = b_1 x + b_2 y + b_3 \tag{3.7}$$

where b_i, i = 1,2,3, are unknown coefficients to be determined. The general case is handled in the same manner as explained below. The identification problem described by relations (3.1) - (3.4) can be reformulated as follows.

First the aquifer equation (2.1) is rewritten in compact form as

$$h(x,y,t;\ T) = F(x_1,\ x_2,\ x_3,\ b_1,\ b_2,\ b_3) \tag{3.8}$$

where $x_1 = x$, $x_2 = y$, $x_3 = t$, b_i's = transmissivity function coefficients. Notice that h in equation (3.8) is now a function of the parameters b_1, b_2, and b_3. Rather than identifying a function T (an infinite dimensional problem) the identification problem now involves the identification of a vector of parameter $\underline{b} = (b_1,\ b_2,\ b_3)$ (a finite dimensional problem).

Using (3.8) the objective functional (3.4) becomes

$$J_1(\underline{b}) = \int_0^t \int_\Omega [F(\underline{x},\ t,\ \underline{b}) - \hat{F}(\underline{x})]^2 dt\ dr \tag{3.9}$$

where

$$\underline{x} = (x_1,\ x_2,\ x_3)$$

$$\hat{F}(\underline{x}) = \hat{h}(x,\ y,\ t)$$

Since in a practical case it is not possible to obtain (observe) $\hat{F}(\underline{x})$ for every point in the aquifer and for every time, a more realistic expression for equation (3.9) should be derived, namely:

$$\phi_1(\underline{b}) = \sum_{i=1}^{n} (F(\underline{x}_i,\ \underline{b}) - \hat{F}(\underline{x}_i))^2 \tag{3.10}$$

where $\underline{x}_i$ represents a discretization of the domain, R, of equation (3.8),

the subscript i denotes the $i^{\underline{th}}$ observation, and n represents the number of output observation points collected. Define the following problem as problem B: Find an estimate of the parameter vector $\underline{b}$ such that the error function (3.10) is minimized subject to the constraints (3.8).

3.4 Problem B: A Quadratic Approximation

An approximate solution to problem B can be obtained by approximating the non-linear objective function (3.10) with a quadratic function. This is achieved by linearizing the constraint (3.8) about an initial guess of the vector of parameters being identified $\underline{b}$, $\underline{b}^\circ$.

Linearization of (3.8) yields

$$h_\ell(\underline{x}_i, \underline{b}^\circ + \underline{S}_t) = F(\underline{x}_i, \underline{b}^\circ) + \nabla F^T(\underline{x}_i, \underline{b}^\circ)\underline{S}_t$$

$$= F(\underline{x}_i, \underline{b}^\circ) + \sum_{j=1}^{K} \left[\frac{\partial F}{\partial b_j}(\underline{x}_i, \underline{b})\Big|_{\underline{b} = \underline{b}^\circ}\right] S_{tj} \quad (3.11)$$

where

R_D = Discretization of R

$\underline{S}_t$ = vector of perturbations

h_ℓ = linearized aquifer model output

K = number of parameters being sought

Accordingly, the objective function (4.10) becomes

$$\phi_1(\underline{b}_o + \underline{S}_t) \simeq \phi_2(\underline{S}_t) = \sum_{i=1}^{n} [h_\ell(\underline{x}_i, \underline{b}^\circ + \underline{S}_t) - \hat{h}(\underline{x}_i)]^2 \qquad (3.12)$$

The function ϕ_2 in (3.12) is a quadratic function of the perturbation $\underline{S}_t$ (since h_ℓ is a linear function of $\underline{S}_t$). Solution of this quadratic problem is now considered. The equations which characterize the optimum $\underline{S}_t^*$ are derived by substituting (3.11) in (3.12) as follows:

$$\phi_2(\underline{S}_t) = \sum_{i=1}^{n} [F(\underline{x}_i, \underline{b}_o) - \hat{F}(\underline{x}_i) + (\nabla F^T(\underline{x}_i, \underline{b}^\circ)|_{\underline{b} = \underline{b}^\circ} \cdot \underline{S}_t)^2] \qquad (3.13)$$

Differentiating (3.13) with respect to $\underline{S}_t$ and setting the result equal to zero the following linear equation which characterize $\underline{S}_t^*$ is obtained

$$A\,\underline{S}_t = \underline{g} \qquad (3.14)$$

where

$$A = P^T P$$

$$P = \begin{bmatrix} \frac{\partial F_1}{\partial b_1} & \frac{\partial F_1}{\partial b_2} & \frac{\partial F_1}{\partial b_3} \\ \frac{\partial F_2}{\partial b_1} & \frac{\partial F_2}{\partial b_2} & \frac{\partial F_2}{\partial b_3} \\ \vdots & \vdots & \vdots \\ \frac{\partial F_n}{\partial b_1} & \frac{\partial F_n}{\partial b_2} & \frac{\partial F_n}{\partial b_3} \end{bmatrix}$$

(All derivatives are evaluated at $\underline{b}^\circ$)

n = number of observation points

$F_i = F(\underline{x}_i, \underline{b})$

$\underline{g} = (g_1, g_2, g_3)$ = gradient vector

$$g_j = \sum_{i=1}^{n} [\hat{F}_i - F_i] \frac{\partial F_i}{\partial b_j} \qquad j = 1, 2, 3$$

$$\hat{F}_i \triangleq \hat{h}(\underline{x}_i)$$

At this point, some of the computational requirements of (3.12) are considered. Since (3.14) is based on (3.11), the latter is studied first. On the right hand side of (3.11) the first term, $F(\underline{x}_i, \underline{b}^\circ)$, requires the solution of the aquifer equation (2.1) with $T(x,y)$ being replaced by $b_1^\circ x + b_2^\circ y + b_3^\circ$. Since (2.1) does not have an analytical solution, an approximate numerical solution is found by using the discrete model (2.5) - (2.9) introduced in Chapter 2. An aquifer simulator [Maddock, 1972] helped in the implementation of the solution. This program was modified and made compatible to be run on the UNIVAC 1108 computer facilities at Case Western Reserve University. Details are given in Appendix A.

The second term of (3.11), $\sum_{i=1}^{n} \frac{\partial F_i}{\partial b_j} \cdot \underline{S}_{tj}$ requires the calculation of the model sensitivity function, $\frac{\partial F_i}{\partial b_j}$, as follows: Perform the differentiations in (3.1) - (3.3) to get

$$T \nabla^2 h + T_x h_x + T_y h_y = S h_t + Q \qquad (3.15)$$

where

$$\nabla^2 h = \frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2}$$

$$T_x = \frac{\partial T}{\partial x}$$

$$T_y = \frac{\partial T}{\partial y}$$

$$h_x = \frac{\partial h}{\partial x}$$

$$h_y = \frac{\partial h}{\partial y}$$

$$h_t = \frac{\partial h}{\partial t}$$

Differentiating (3.15) with respect to (w.r.t.)b_j yields:

$$T \nabla^2 h_j + T_x h_{xj} + T_y h_{yj} = S h_{tj} - \hat{Q} \tag{3.16}$$

where

$$h_j = \frac{\partial h}{\partial b_j}$$

$$h_{xj} = \frac{\partial}{\partial x} \left(\frac{\partial h}{\partial b_j}\right)$$

$$h_{yj} = \frac{\partial}{\partial y} \left(\frac{\partial h}{\partial b_j}\right)$$

$$\hat{Q} = T_j \nabla^2 h + T_{xj} h_x + T_{yj} h_y \tag{3.17}$$

where $$T_j = \frac{\partial T}{\partial b_j}$$

$$T_{xj} = \frac{\partial}{\partial b_j} \left(\frac{\partial T}{\partial x}\right)$$

$$T_{yj} = \frac{\partial}{\partial b_j} \left(\frac{\partial T}{\partial y}\right)$$

Initial and boundary conditions for (3.16) are obtained by differentiation of I.C. and B.C. of (3.1) - (3.3) w.r.t. b_j. Since those conditions do not depend on the transmissivity function, the set of I.C. and B.C. for (3.16) are homogeneous. Notice that (3.16), the sensitivity equation of the aquifer model, is also a parabolic differential equation of the same type as (3.1) - (3.3). This fact greatly aids the development of the identification digital package, since no additional coding has to be written to calculate the sensitivity equations of the model. As a consequence, a discrete version of equation (3.16), similar to the discrete version of of eqns. (2.8) - (2.9) obtained for the aquifer equation, can be derived. Here, only the discretization of $\hat{Q}$ will be explained since discretization of any other term in equation (3.16) follows the same steps as in section 2.3.4. For the sake of simplicity only the $\hat{Q}$ corresponding to the sensitivity equation for j=1, namely b_1 will be discretized. Equation (3.17) becomes

$$\hat{Q} = T_1 \nabla^2 h + T_{x1} h_x + T_{y1} h_y \qquad (3.18)$$

note that

$$T = b_1 x + b_2 y + b_3$$

thus

$$T_1 = x, \; T_2 = y, \; T_3 = 1,$$

accordingly

$$\hat{Q} = \nabla^2 h + h_x \qquad (3.18)$$

Replacing h by central differences and h_x by forward differences, equation (3.8) becomes

$$\hat{Q} = j \cdot (D_{ijk}) + (h_{ijk+1k} - h_{ijk}) \qquad (3.19)$$

where

$$D_{ijk} = (h_{i-1jk} - h_{ijk}) + (h_{i+1jk} - h_{ijk}) +$$

$$(h_{ij-1k} - h_{ijk}) + (h_{ij+1k} - h_{ijk})$$

Going back to equation (3.14) some of the computational effort required to solve the equation which characterizes $\underline{S}_t^*$ now becomes clear. To construct the matrix A in (3.14), first, the aquifer equation (2.1) has to be solved, then the sensitivity equation (3.17) is solved for b_1, b_2, and b_3 (i.e., three equations), and finally a matrix multiplication of the sensitivity functions P is performed, namely P^TP. Similar manipulations are required to define $\underline{g}$. $\underline{S}_t^*$ is obtained by pre-multiplication of $\underline{g}$ by the inverse of A, A^{-1}.

3.5 Development of the Identification Algorithm

Once the quadratic problem is solved the next step is to decide on some way of improving the initial guess T_o in order to get better quadratic approximations to the identification problem described by equations (3.8) - (3.10).

Any scheme to correct the initial guess must consider the characteristics of the "surface" $\phi_1(\underline{b})$. If F_i is linear in $\underline{b}$, then, $\phi_1(\underline{b})$ represents a surface whose contours are ellipsoids. For the nonlinear case the surface becomes distorted according to the severity of the nonlinearity. This "ill-conditioned" characteristic of the $\phi_1(\underline{b})$ surface accounts for almost

all the trouble found when, attempts to improve $\underline{b}^{\circ}$ using traditional methods are made. For example, the well known Gauss-Newton method, [Marquardt, 1963[proceeds to improve $\underline{b}^{\circ}$ by iterating the solution method of equation (3.14). In practice, a value of $\underline{b}^{\circ}$ far from the optimum of $\phi_1(\underline{b})$ causes the Gauss-Newton method to diverge. On the other hand, a steepest descent technique would alleviate the problem caused by starting points far from the optimum, but would not handle the entire optimization procedure, due to its natural slow convergence compounded with the poor conditioning of the $\phi_1(\underline{b})$ surface.

Marquardt [Marquardt, 1963], considered these problems and derived an algorithm which combines efficiently the Gauss-Newton method and the steepest descent, reinforcing each other whenever local conditions of the $\phi_1(b)$ surface requires it.

Marquardt's algorithm fits the requirements of the present problem well (improvement of $\underline{b}^{0}$) and it is utilized as part of the Identification Algorithm.

By "improvement of $\underline{b}^{\circ}$" via the application of Marquardt's algorithm (M.A.) we mean any $\underline{b}$ which yields a further decrease in the objective function (the error criterion).

M.A. is based on the concept that the best direction, $\underline{S}$, for finding a local reduction in the value of $\phi_1(b)$, lies between (Gauss-Newton method direction) and $\underline{s}_g$ (steepest descent method direction). Marquardt proposes to find that direction by solving the following equation:

$$(A + \lambda I)\, S = \underline{g} \tag{3.20}$$

Where A and $\underline{g}$ were defined in (3.14) and λ is Marquardt parameter. Notice that if λ is very small, then equation (3.20) is reduced to (3.14), while if λ is very large, $\underline{S} = \lambda^{-1}\underline{g}$, which is the steepest descent direction; thus, M.A. is indeed a hybrid between the steepest descent and Gauss-Newton methods. The most critical step in M.A. is the selection of the parameter λ. The goal is to adjust λ which yields to a reduction in $\phi_1(\underline{b})$. The mathematical basis of the algorithm are stated in three theorems which can be found in [Marquardt, 1963]. Those theorems assure the existence of a sufficiently large λ which satisfies the condition,

$$\phi_1(\underline{b}^{(r+1)}) < \phi_1(\underline{b}^{(r)}) \qquad (3.21)$$

where the superscript (r) denotes the iterations number. Marquardt's strategy for choosing λ seeks to use small values whenever the local conditions of the $\phi_1(\underline{b})$ surface are such that equation (3.11) represents a suitable linear approximation of the aquifer model (3.1) - (3.3). In general, he uses large values of λ only when necessary to satisfy condition (3.21); otherwise, equation (3.20) would become a steepest descent technique with all the drawbacks previously cited.

Following this line of thinking, Fletcher, [Fletcher, 1972], proposed an improved strategy to select λ by comparing the actual reduction in the sum of squares, $\phi_1(\underline{b}^{(n)}) - \phi_1(\underline{b}^{(n)} + \underline{S})$, and the reduction, $\phi_2(\underline{b}^{(n)}) - \phi_2(\underline{b}^{(n)} + \underline{S})$, that would have taken place if the quadratic approximation function (3.12) were exactly correct, where,

$$\phi_2(\underline{b}^{(n)}) - \phi_2(\underline{b}^{(n)} + \underline{S}) = -[2\,\underline{r}^T(b^{(n)})\,P(b^{(n)})\,\underline{S} + \underline{S}^T A\,\underline{S}]$$

where $\underline{r}$ is the vector of residuals:

$$\underline{r}(\underline{b}) = (r_1(\underline{b}),\ldots,r_n(\underline{b}))$$

and $r_i(\underline{b}) = F(\underline{x}_i,\underline{b}) - \hat{F}(\underline{x}_i)$

when the ratio,

$$R = \frac{\Phi_1(\underline{b}^{(n)}) - \phi_1(\underline{b}^{(n)} + \underline{S})}{\phi_2(\underline{b}^{(n)}) - \phi_2(\underline{b}^{(n)} + \underline{S})}$$

is close to 1, the quadratic approximation (3.12), is good and λ should be reduced to allow large steps. If R is negative or near zero then (3.12) is bad and λ should be increased. As a consequence, Fletcher suggests the following strategy, [Fletcher, 1972]:

(i) If $R \in [\rho,0]$, then make $\lambda^{(n)} = \lambda^{(n-1)}$. Otherwise go to (ii)

(ii) If $R < \rho$ make $\lambda^{(n)} = \hat{R} \cdot \lambda^{(n-1)}$

where

$$\hat{R} = 2 + \frac{\phi_1(\underline{x}_i,\ \underline{b}^{(n-1)}) - \phi_1(\underline{x}_i,\ \underline{b}^{(n+1)} + \underline{S}^{(n+1)})}{[F(\underline{x}_i,\ \underline{b}^{(n+1)} - \hat{F}(\underline{x}_i)] \cdot P \cdot \underline{S}^{(n-1)}}$$

when the factor $\hat{R}$ lies outside the range $[\rho,0]$, then make the correction factor $\hat{R}$ as follows:

$$\hat{R} = \begin{cases} 2 & \hat{R} < 2 \\ 10 & \hat{R} > 10 \end{cases}$$

The following scheme is proposed as a solution approach to Problem B:

(i) Make an initial guess for the vector parameter $\underline{b}$:

$$\underline{b}^0 = (b_1^0, b_2^0, b_3^0)$$

(ii) Substitute $T = b_1^0 x + b_2^0 y + b_3^0$ into the aquifer model (2.1) and generate $h(\underline{x}_i, \underline{b}^0)$, the model's output, by using the discrete model (2.5) - (2.9) and the Alternate Implicit Direction method.

(iii) Substitute the model's output obtained in step (ii) into equation (3.16) to get the model sensitivity functions. Solve (3.16) the same way as (3.1) - (3.3).

(iv) Solve the quadratic problem (3.11) - (3.12) by solving the equation of characterization (3.14). To do so, use the information generated in steps (ii) and (iii) above.

(v) Check how close the solution of the quadratic problem is to the actual solution, by monitoring the decrease of the correction term $|\underline{S}^{(r)}|$. Specifically, use as a stopping criterion

$$\frac{|Sj^{(r)}|}{\tau + |bj^{(r)}|} \leq \varepsilon \quad , \quad j = 1, 2, 3 \qquad (3.22)$$

where τ is a small positive number which insures that the denominator of (3.22) does not become zero and ε is a convergence factor, usually selected as 10^{-5}. If (3.22) is not met, generate a new quadratic approximation by following

The following scheme is proposed as a solution approach to Problem B:

(i) Make an initial guess for the vector parameter $\underline{b}$:

$$\underline{b}^0 = (b_1^0, b_2^0, b_3^0)$$

(ii) Substitute $T = b_1^0 x + b_2^0 y + b_3^0$ into the aquifer model (2.1) and generate $h(\underline{x}_i, \underline{b}^0)$, the model's output, by using the discrete model (2.5) - (2.9) and the Alternate Implicit Direction method.

(iii) Substitute the model's output obtained in step (ii) into equation (3.16) to get the model sensitivity functions. Solve (3.16) the same way as (3.1) - (3.3).

(iv) Solve the quadratic problem (3.11) - (3.12) by solving the equation of characterization (3.14). To do so, use the information generated in steps (ii) and (iii) above.

(v) Check how close the solution of the quadratic problem is to the actual solution, by monitoring the decrease of the correction term $|\underline{S}^{(r)}|$. Specifically, use as a stopping criterion

$$\frac{|Sj^{(r)}|}{\tau + |bj^{(r)}|} \leq \varepsilon \quad , \quad j = 1, 2, 3 \qquad (3.22)$$

where τ is a small positive number which insures that the denominator of (3.22) does not become zero and ε is a convergence factor, usually selected as 10^{-5}. If (3.22) is not met, generate a new quadratic approximation by following

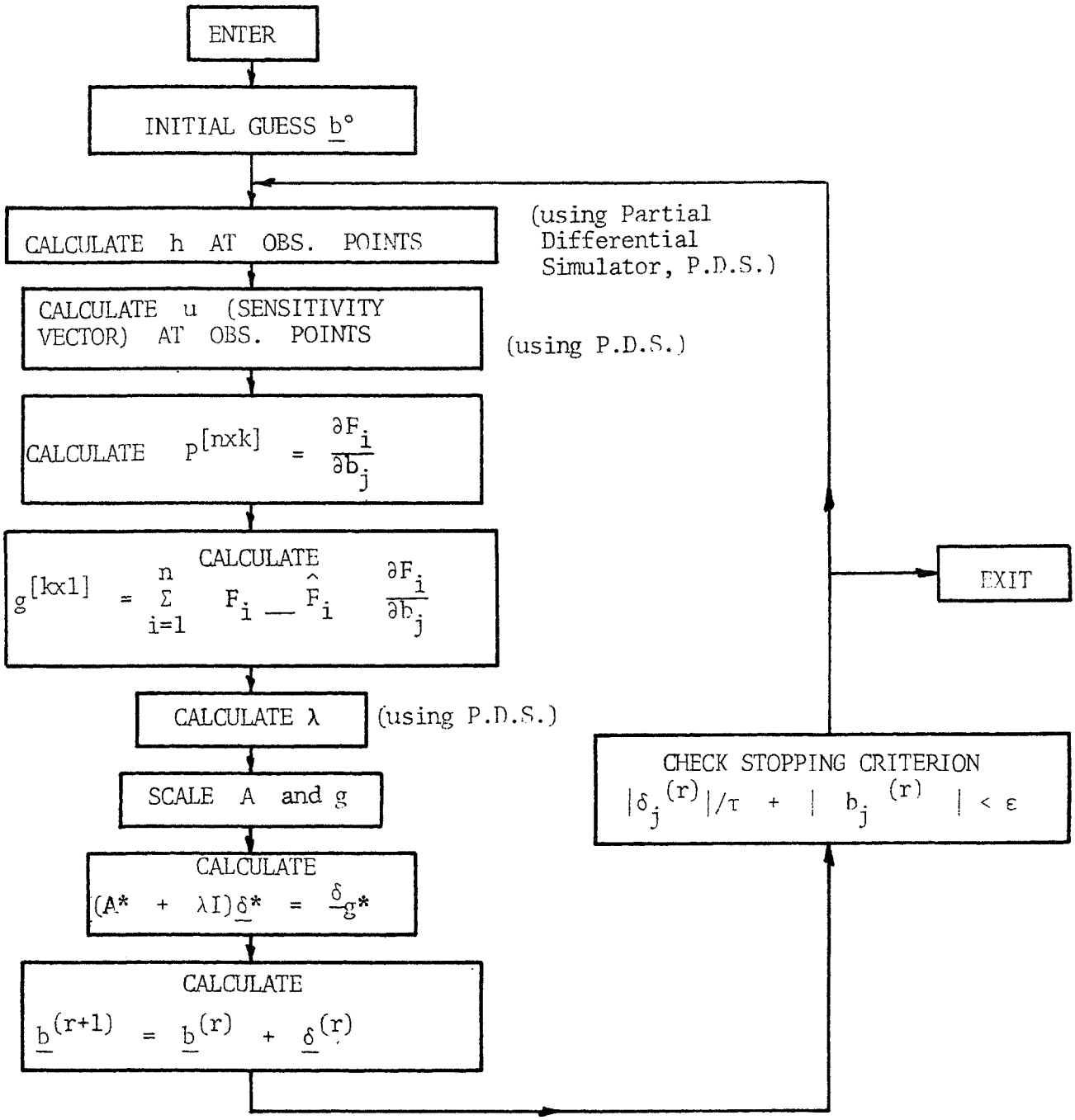
steps (vi) - (ix). Another stopping criterion could be to check the decrease of the gradient $P \cdot (\hat{h} - h)$.

(vi) Set $\underline{b}^{(r+1)}=\underline{b}^{(r)} + \underline{S}^{(r)}$ and repeat steps (ii) and (iii) above.

(vii) Select $\lambda^{(r)}$ using the strategy previously explained. Go to step (iv), (use $\lambda^{(r)}$ rather than $\lambda^{(0)}$).

A flow diagram of the Identification Algorithm is depicted in Figure 3.1. Notice the extra step depicted there, namely the scaling of the matrix A and the vector $\underline{g}$. This is suggested by Marquardt as a means to improve the numerical aspects of the solution of (3.20). The matrix A tends to be ill conditioned, and therefore, inversion of $(A + \lambda I)$ in (3.20) presents numerical difficulties. Common practice is to scale A so that its diagonal elements become unity. The same technique applies to g.

A digital package implementing the Identification Algorithm just described was written in Fortran V language for the Univac 1108 digital computer. Flow diagram and program description are presented in Appendix A. A linear polynomial representation was chosen for T in this program. Reasons for doing so were based on trade-off considerations between accuracy of the T representation vs computer time. For the real example solved in Chapter 4 the linear representation was enough to produce an accurate aquifer model. Extensions of the approach presented in this chapter to more complicated representations of T are discussed in the following chapters.



FLOW DIAGRAM OF THE IDENTIFICATION ALGORITHM

FIGURE 3.1

Chapter 4

COMPUTATIONAL RESULTS AND ANALYSES

4.1 Introduction

To test the performance of the Identification Algorithm developed in Chapter 3, the modeling and calibration of a real unconfined aquifer system is undertaken. The presentation of the real example is organized in four sections which; (i) describe the hydrological as well as the physical aspects of the real system, (ii) propose and implement a mathematical model, (iii) calibrate the model and finally, (iv) validate this model.

The presentation of the real example is preceded by two hypothetical examples brought into consideration for two reasons:

(i) To verify the accuracy of the identification algorithm by checking it with known results.

(ii) Tutorial purposes.

In the following developments, transmissivity is represented with a first order polynomial. This linear transmissivity representation yields an aquifer model accurate enough to avoid the use of higher degree polynomials, yet a comparison between the real and the model transmissivities indicated differences that were already expected. A transmissivity function which is represented by a higher order polinomial is discussed in subsequent sections.

4.2 Hypothetical Examples

A hypothetical aquifer whose characteristics are listed in Table 4.1 is assumed given. It is assumed that all model parameters are known except transmissivity. Thus the identification algorithm is used with the purpose of identifying the transmissivity value. The aquifer being modeled is confined, with constant head and zero flow boundary conditions. It is bounded (hypothetically speaking) by bedrock walls of very low permeability except for one boundary strip which is hydraulically connected with a large source of water (a river) which causes the constant head condition.

The geometric characteristics are such that the discrete model (2.5) - (2.9), used in the numerical computations, is represented by a square grid 5,000 feet wide with 1,000 feet separation between nodes (i.e., there are 25 grid points).

Water head observations are generated by assuming true values for the aquifer transmissivity, and then solving the discrete model using those values plus the conditions in Table 4.1. The solution of the discrete model is obtained by decoupling, in the identification program, the P.D.E. solver (or aquifer simulator as it is called in this work) from the optimization section. Q and Δt are assumed to be 0.1 ft^3/sec, and 2 hours (10^4 secs) respectively. It is hypothesized that a five pumping period (of 2 hr. each) test was performed in the field to assist in determining water level historical record.

Example 1: In example 1 the "real" transmissivity is assumed

Characteristic	Description
Aquifer Type	Confined
Storage Coefficient S	0.0025 (no dimensions)
Transmissivity Coefficient T	Unknown (ft^2/sec)
Initial Head	Constant (ft)
Boundary Conditions	Bound γ_1: Constant head Other bounds: Zero flow (impermeable)
Wells	Flow rate of 0.1 (ft^3/sec) for one recharge well
Approximate Area	One square mile

Aquifer Data: Hypothetical Example

TABLE 4.1

constant on the entire aquifer. Complete output information is available, i.e., the water head changes are known at every point in the aquifer and for every pumping period. Since the grid of the discrete model is made of 25 grid points, and the total number of pumping periods is 5, the number of observation points generated was 125. The assumed true value for transmissivity was $T^* = 0.09\ ft^2/sec$. Using this information the identification problem is formulated as:

$$\min_{T} \phi = \Sigma(h(x_i,T) - \hat{h}(x_i))^2 \tag{4.1}$$

subject to

$$\frac{\partial}{\partial x}(T\frac{\partial}{\partial x} h(x_i,T)) + \frac{\partial}{\partial y}(T \frac{\partial h}{\partial y}(x_i,T)) =$$

$$0.0025 \frac{\partial h}{\partial t} (\underline{x}_i, T) + \frac{(0.1)}{\Delta x^2} \delta(\underline{x}_i - a) \tag{4.2}$$

I.C.:

$$h(x,y,0) = \text{constant}$$

$$h(x,y,t)\big|_{\Gamma_1} = \text{constant}$$

$$\frac{dh}{dn} (x,y,t)\big|_{\Gamma_2,\ \Gamma_3,\ \Gamma_4} = 0$$

where

Γ_i = grid boundaries. The sides of the square grid have been numbered clockwise starting from the left side.

$\vec{n}$ = perpendicular direction to the boundary

Δx = distance between the grid's nodes = 1000 ft.

a = (3, 3, K) = pumping function coordinates

$\delta(\underline{x}_i, -\underline{a})$ = Dirac delta function to indicate that there is only one well at the grid node (3,3)

The problem defined by relations (4.1) - (4.2) was solved using the identification algorithm. Results are presented in Table 4.2. The observed head and the predicted head were compared at the point where the pumping well is located. A close agreement between the two was observed. An initial guess of the transmissivity T , was chosen to be one order of magnitude greater than the "real value" in order to check the performance of the algorithm with starting points far from optimum. Figure 4.1 depicts the minimization of the error function. The algorithm exhibits quadratic convergence, i.e., the error in a given iteration is proportional to the square of the error in the previous iteration. The flatness of the error surface is seen after iteration 7 (see Table 4.2) where the rate of decrease begins to be slower, and the optimal estimate of transmissivity is close to the real value.

Another numerical experiment was carried out to test convergence as a function of the initial guess. This time an initial guess value about one order of magnitude smaller than the real value of the transmissivity was selected. Results are presented in Table 4.3. Convergence was attained in fewer iterations than for the first case.

Example 2: In this example, the aquifer transmissivity is assumed to be a linear function of space. The function is given as follows:

$$T = -(.1875x + .0375y)10^{-4} + 0.1675$$

The coefficient values were selected to be compatible with realistic values for sand and gravel aquifers. A typical range of transmissivity values is $0.01 \leq T \leq 0.1$ ft^2/sec.

The identification problem is to estimate the parameters b_j such that the error function ϕ_1 is minimized, namely

$$\min_{\underline{b}} \phi_1 = \sum_{i=1}^{125} (h(\underline{x}_i, \underline{b}) - \hat{h}(\underline{x}_i))^2 \qquad (4.3)$$

s.t.

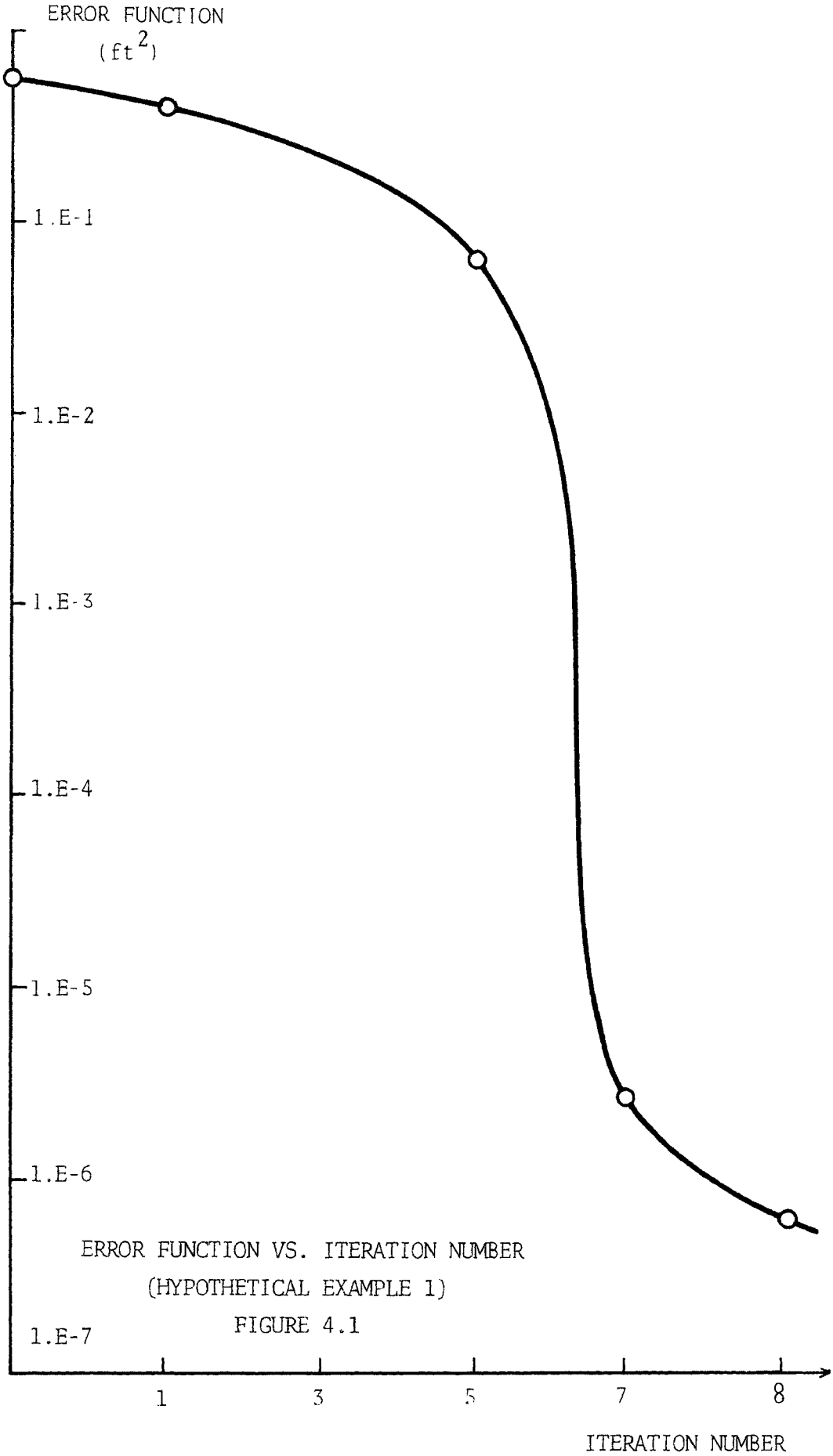
$$\frac{\partial}{\partial x}[b_1x + b_2y + b_3)\frac{\partial h}{\partial x}(\underline{x}_i, \underline{b})] + \frac{\partial}{\partial y}[(b_1x + b_2y + b_3)\frac{\partial h}{\partial y}(\underline{x}_i, \underline{b})] =$$

$$= 0.0025\frac{\partial h}{\partial t} + \frac{(0.1)}{(\Delta x)^2}\delta(\underline{x}_i - a) \qquad (4.4)$$

Solution of problem (4.3) - (4.4) is obtained by using the identification algorithm. The procedure was started with the initial guess

$$T_0 = 10^{-6}x + 10^{-7}y + 0.01$$

where the coefficients were chosen far from the real ones to test the



ERROR FUNCTION VS. ITERATION NUMBER
(HYPOTHETICAL EXAMPLE 1)
FIGURE 4.1

Variable Name	Starting Value	Iteration Number 1	3	5	7	9	11	13	Correct Value
Transmissivity T (ft^2/sec)	0.90	0.6506	0.3676	0.1801	0.08998	0.09005			0.0900
Error Function ϕ (ft^2)	0.4414	0.3778	0.2545	0.09063	1.181E-6	7.086E-7			
Lambda Parameter λ	1.0E-2	1.0E-3	1.0E-5	1.0E-7	1.0E-9	1.0E-10			

Variable Name	Starting Value	Pumping Period 1	2	3	4	5
Observed Head (ft)	0.0	0.1818	0.2250	0.3293	0.3648	0.3905
Calculated Head (ft)	0.0	0.1818	0.2250	0.3293	0.3648	0.3905

TABLE 4.2 - Results hypothetical example 1. Correct T value perturbed by a factor of 10. Optimal water heads are shown in lower table.

Variable Name	Starting Value	Iteration Number 1	2	3	4	5	6	7	Correct Value
Transmissivity T (ft^2/sec)	0.01	0.03198	0.06053	0.08272	0.08941	0.09002	0.09005		0.0900
Error Function Ø (ft^2)	2.114	0.4211	5.013E-2	2.000E-3	1.357E-5	7.080E-6	7.000E-7		
Lambda Parameter λ	1.0E-2	1.0E-3	1.0E-4	1.0E-5	1.0E-6	1.0E-7	1.0E-8		

Variable Name	Starting Value	Pumping Period 1	2	3	4	5
Observed Head (ft)	0.0	0.1818	0.2250	0.3293	0.3648	0.3950
Calculated Head (ft)	0.0	0.1818	0.2250	0.3293	0.3648	0.3905

TABLE 4.3 - Results hypothetical example 1. Correct T value perturbed by a factor of 1/10. Optimal water heads are shown in lower table.

convergence of the algorithm. Results of the parameter search are displayed in Table 4.4. The error function decreases monotically as shown in Figure 4.2. In the first two steps the initial error was decreased 99%. The reduction is about one order of magnitude per iteration.

Several numerical experiments were performed to identify the region of convergence of the algorithm for this particular problem. This region was roughly determined to be $\frac{b^*}{1000} \leq b \leq 1.56b^*$ by the following procedure:

(i) Perturb the optimal solution by increasing the parameter values of (b_1, b_2, b_3) by one order of magnitude.

(ii) Perturb the optimal solution by decreasing the parameter values of (b_1, b_2, b_3) by one order of magnitude.

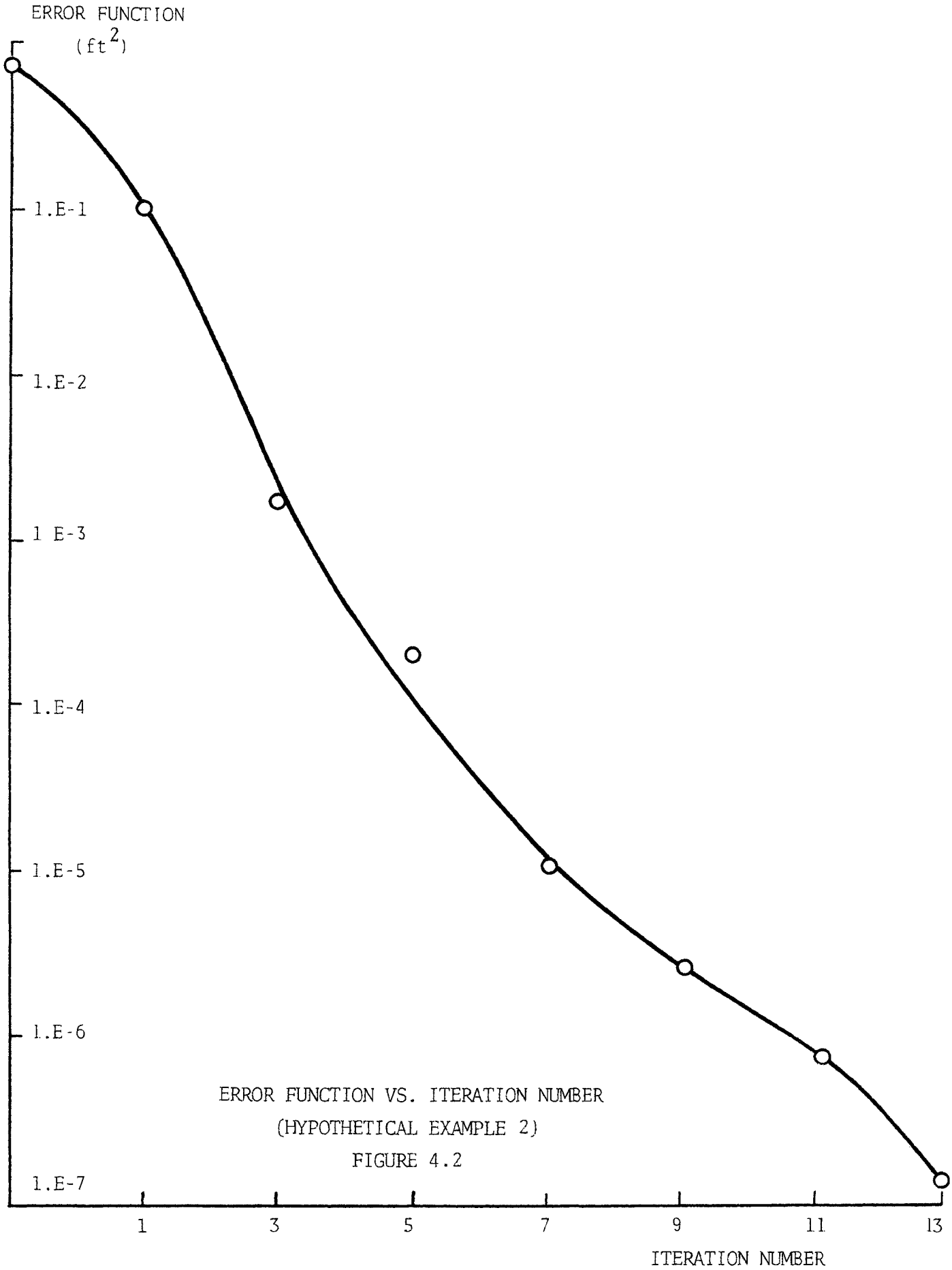
(iii) Perturb each parameter individually from its optimal value.

It was observed that during the minimization process, transmissivity may sometimes become partially negative, especially when the starting point is not within the convergence range specified above. Negative transmissivity caused instability problems (oscillatory solutions) to the alternate implicit direction algorithm which solved the aquifer equation. The AID has been shown to be stable, (Birkoff and Urvga, 1959], for any time step or space increment when applied to parabolic equations of the type considered here. However, the proofs of stability impose constraints on the geometry of the aquifer as well as on the coefficients of the equation. In particular, it is required that the coefficients be greater than or equal to zero.

To alleviate the problem, the identification algorithm was modified so that detection of the oscillatory condition could be made and then

Parameters	Starting Value	Iteration Number 1	3	5	7	9	11	13	Correct Values
$b_1 \times 10^{+6}$ (ft/sec)	1.0	1.996	-12.92	-17.26	-18.18	-18.48	-18.62	-18.70	-18.75
$b_2 \times 10^{+6}$ (ft/sec)	0.1	1.269	-6.591	-6.332	-4.970	-4.281	-3.989	-3.800	-3.850
b_3 (ft/sec)	0.01	0.01766	0.1556	0.1720	0.1708	0.1688	0.1678	0.1678	0.1675
Error Function ϕ (ft^2)	0.7942	0.1059	1.495E-3	1.761E-4	1.109E-5	2.744E-6	6.736E-7	1.229E-7	
λ Parameter	1.0E-2	1.0E-3	1.0E-5	1.0E-7	1.0E-9	1.0E-11	1.0E-13	1.0E15	

TABLE 4.4 - Results hypothetical example 2. The initial guess is one order of magnitude (greater for b_1 and b_2) smaller than the correct values.



ERROR FUNCTION VS. ITERATION NUMBER
(HYPOTHETICAL EXAMPLE 2)
FIGURE 4.2

corrected. The abnormal situation was corrected by providing the program with a built-in mechanism to select a new starting point. The idea is to penalize the error function by a factor of 10 (arbitrarily chosen) so that the algorithm be forced to look for a better λ which in turn implies the choice of a new point $\underline{b}$. This is an empirical method which proved to be useful for some cases. In extremely bad cases a new starting point will have to be externally supplied to the program.

Results of the sensitivity analysis are presented in Tables 4.5 and 4.6 The algorithm seems to be more sensitive to b_1 than to b_2 and b_3 since the maximum increase allowed for b_1 is 1.5 before the algorithm diverges or becomes unstable. Notice that the optimal solution $T = -(18.75x + 3.85y)10^{-6} + 0.167$, becomes negative at x = 7000 ft, y = 7000 ft, for values of b_1 greater than 1.1 x 18.75. It appears reasonable that the maximum increase allowed for b_1 be 1.5.

4.3 Discussion

The performance of the identification algorithm in the solution of example problem 1 was satisfactory. The accurate convergence to the optimum, while starting with poor initial guesses of transmissivity, is very encouraging. The first run took 25 seconds and 10 iterations. The second run took 20 seconds and only 6 iterations on the UNIVAC 1108 digital system.

On the other hand, solution of example 2 pointed out some limitations of the algorithm. The instability produced in the aquifer simulator when

Parameters	Starting Value	Iteration Number 1	3	5	7	9	11	13	Correct Values
b_1 X 10^{+6} (ft/sec)	-0.1	1.034	-16.67	-16.63	-18.00	-18.40	-18.55	-18.70	-18.75
b_2 X 10^{+6} (ft/sec)	-0.01	1.064	-11.37	-6.770	-5.302	-4.437	-4.042	-3.878	-3.875
b_3 (ft/sec)	0.01	0.01647	0.2004	0.1714	0.1715	0.1692	0.1679	0.1678	0.1675
Error Function $\emptyset$ (ft^2)	1.475	0.2405	0.00356	1.320E-4	1.826E-5	3.817E-6	1.070E-6	1.587E-7	
λ-Parameter	1.0E-2	1.0E-3	1.0E-5	1.0E-7	1.0E-9	1.0E-11	1.0E-13	1.0E-15	

TABLE 4.5 - Results hypothetical example 2. Sensitivity Analysis. Perturbation of b_1, b_2, and b_3.

Parameters	Starting Value	Iteration Number 1	2	3	4	5	6	7	Correct Values
$b_1 \times 10^{+6}$ (ft/sec)	-30.00	-19.47	-19.26	-19.09	-18.85	-18.75			-18.75
$b_2 \times 10^{+6}$ (ft/sec)	-3.75	-3.416	-3.445	-3.505	-3.567	-3.68			-3.75
b_3 (ft/sec)	0.1675	0.1688	0.1685	0.1680	0.1671	0.1675			0.1675
Error Function ϕ (ft^2)	8.232E-3	7.000E-5	2.690E-6	9.923E-7	2.386E-7	2.000E-7			

TABLE 4.6 - Results hypothetical example 2. Sensitivity Analysis. Perturbation of b_3.

transmissivity becomes partially negative could be a limitation for applications where there is no previous knowledge at all regarding transmissivity values. For these cases several initial points will have to be tried.

In example 2 , the range, $\frac{b^*}{100} \leq b \leq 1.5\ b^*$, still allows a wide variation for T. As far as computer time is concerned, the different experiments performed in example 2 averaged 43 sec/iteration. The average number of iterations was 14. These times are reasonable particularly when taking into account the fact that each iteration requires the solution of four partial differential equations. In the hypothetical examples as well as in the real one below, b_3 was selected as a large number (between 0.1 and 0.5) while b_1 and b_2 were chosen smaller (10^{-6}, 10^{-8}). This selection of b's locates the transmissivity plane almost parallel to the aquifer area and enough above it to avoid the transmissivity to be driven to a negative value at the first algorithm's corrections. Usually, a good starting value for b_3 was the arithmetic mean of the available information on transmissivity.

4.4 A Real Example

The construction and validation of an aquifer model for the Fairfield-New Baltimore area is an important step in this project since no prediction of the real system behavior can be made without such a component. The purpose of this section is to show how the modeling and identification of the Fairfield-New Baltimore Hamilton aquifer can be implemented using the identification algorithm.

4.4.1 Description of the Real System

The area modeled in the present study which is shown in Figure 4.3 was discussed in details in section 1.2. In particular, it consists of 32 square miles of the Great Miami River Valley southwest of Hamilton, Ohio.

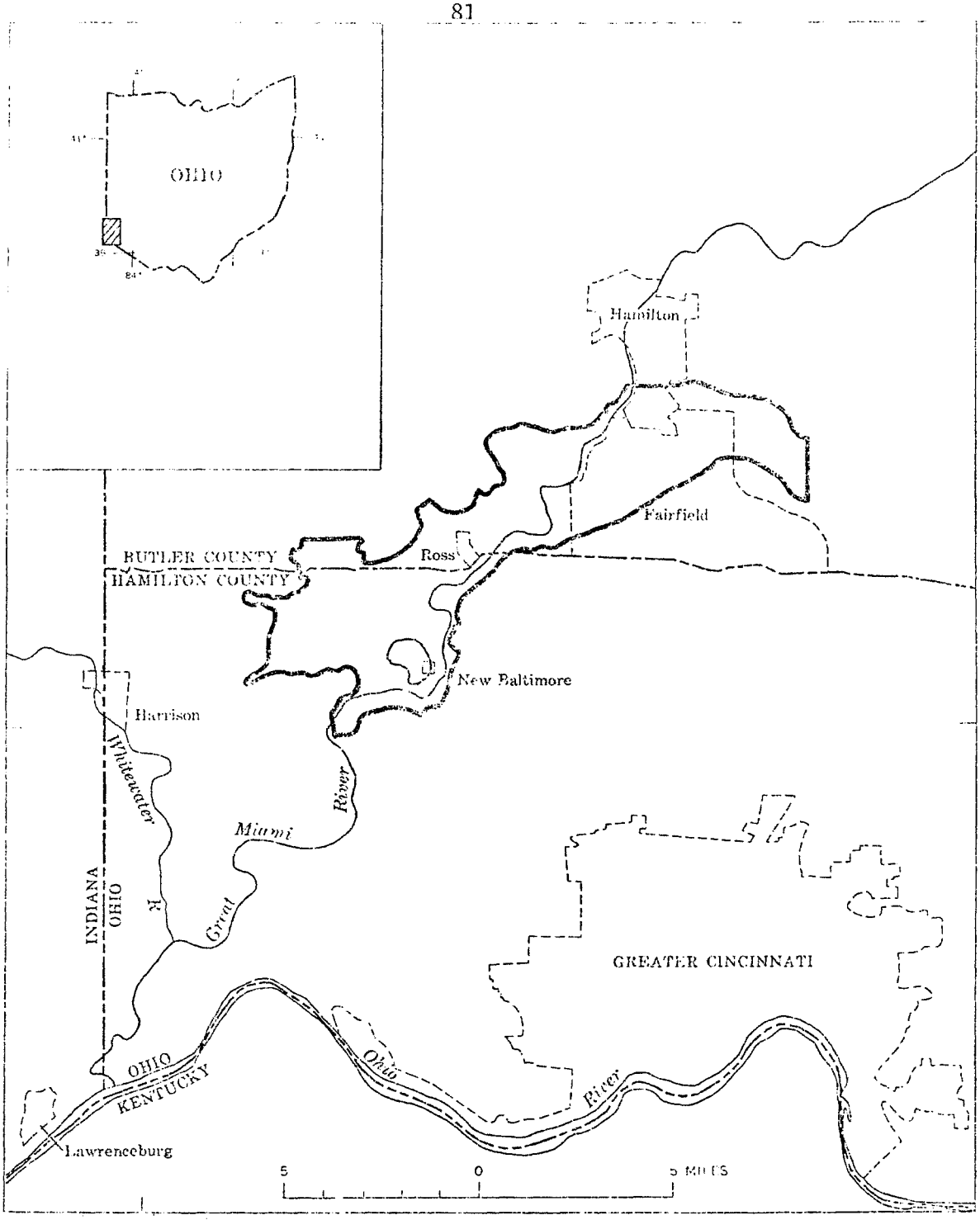
The basic hydrologic properties of the aquifer are essentially described by the coefficients of transmissivity and storage. A perfect simulation of the system requires these parameters to be known at every point. Since this is impossible, the hydrologist must use the available data to interpolate those figures in some optimal manner. The identification algorithm developed for this project constitutes one possible interpolation scheme.

For the Fairfield-New Baltimore area only four reliable, [Spieker, 1968], tests have been performed to determine the aquifer transmissivity. The test points are shown in Figure 4.4 as T_1, T_2, T_3, T_4. The average storage coefficient has been estimated to be 0.145 in this study, based on figures suggested by Spieker, [Spieker, 1968].

4.4.2 The Aquifer Model

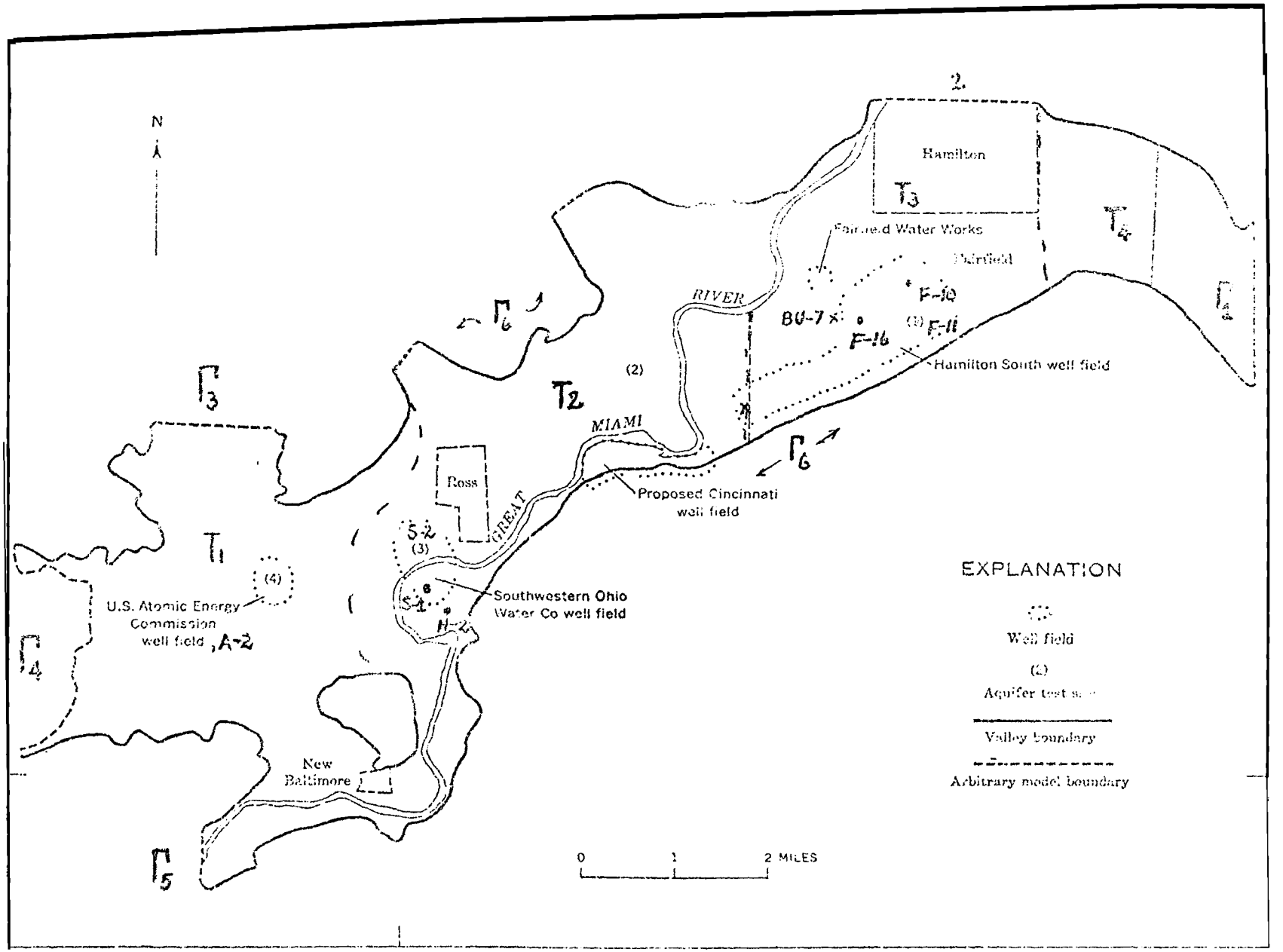
The modeling of the real system described in the previous sections is implemented in this section. The model assumptions are:

1. The aquifer drawdown is and will remain small compared with its saturated thickness.
2. The initial conditions of the system are uniform throughout the aquifer.



LOCATION OF THE FAIRFIELD-NEW BALTIMORE AREA.
LOWER GREAT MIAMI RIVER VALLEY (after Spieker).

FIGURE 4.3



THE FAIRFIELD-NEW BALTIMORE AQUIFER. Γ_i DENOTES CONSTANT HEAD BOUNDARIES. Γ_6 DENOTES CONSTANT FLOW BOUNDARIES. T_i LOCATES AREAS WHERE TRANSMISSIVITY HAS BEEN MEASURED.

FIGURE 4.4

3. The east and west aquifer boundary conditions can be represented by constant head boundaries.
4. All flow within the aquifer is two dimensional (no vertical component); recharge from the boundaries is one dimensional.
5. The stream is modeled as a constant head boundary condition.
6. Storage coefficient is constant.
7. Time is divided into equal intervals of one year each. Within each interval, called pumping period, wells pump at constant rates. All wells fully penetrate the aquifer.

In the present work, initial conditions of the area were averaged to 550 ft. based on Spieker's figures, [Spieker, 1968]. Assumption 1 introduces some error since it does not allow the model to take into account dewatering of the system. The error becomes serious only when 20 to 25 percent of the aquifer's saturated thickness has been dewatered. This thickness has been estimated between 140 and 200 feet. Therefore, the model can be used to evaluate drawdowns of as much as 35 feet. Beyond this point equation (2.1) does not accurately represent the system.

Assumptions 2 through 4 and 7 may introduce errors of local extent but are not expected to introduce serious errors in the regional potential distribution.

Under conditions of heavy pumping assumption 5 would have to be modified since the reduction in the streamflow may reach levels, such that there would not be enough river water to maintain the water level constant beneath the streambed. Actual pumping conditions are low (by comparison with the stream recharge), and therefore assumption 5 is justified.

Finally, assumption 6 is based on the fact that most of the aquifer extent (except for small marginal areas which are of the semiconfined type) are unconfined and the coefficient of storage of an unconfined aquifer is generally about 0.1-0.2 [Spieker, 1968].

Now, the aquifer model is defined using the information above as follows:

$$\frac{\partial}{\partial x}\left(T(x,y)\frac{\partial h}{\partial x}\right) + \frac{\partial}{\partial y}\left(T(x,y)\frac{\partial h}{\partial y}\right) =$$

$$0.145\,\frac{\partial h}{\partial t} + \frac{P(x,y,t)}{(\Delta x)^2}\,\delta(\underline{x}_i - \underline{a}) \qquad (4.5)$$

I.C. $h(x,y,0) = 550$ ft.

$h(x,y,t)\big|_{\Gamma_i} = 550$ ft. $i = 1, 2, \ldots, 5$ (see Fig. 5.4)

$h(x,y,t)\big|_{\Gamma_6} = 550$ ft. River

$\frac{dh}{d\vec{n}}(x,y,t)\big|_{\bar{\Gamma}}$ = Recharge from boundaries (see Table 5.7)

where

$P(x,y,t)$ = Pumping function. See Table 4.8

Δx = 2,640 ft. = grid's nodes distance

$\vec{n}$ = perpendicular direction at the boundary

$\delta(\underline{x}_i - \underline{a})$ = Dirac delta function. It indicates the location of the pumping well

$\underline{a}$ = Pumping coordinates = (x_i, y_i, t_k)

Boundary Points (B.P.)	12,5	12,6	11,6	11,8	12,8	12,9	11,10	10,11	9,11	9,12	9,13	8,14
	(I,J) (See figure 5.5 for location of this coordinates)											
Infiltration Rate (I.R.)	7	5	5	7	3	3	10	8	5	5	6	8
B.P.	7,15	6,15	5,15	5,16	6,17	6,18	5,19	4,20	4,25	4,26	4,27	5,28	6,28
I.R.	8	5	8	6	5	5	8	9	6	5	5	7	3
B.P.	8,28	7,27	6,26	6,25	7,24	8,23	9,22	10,21	10,20	10,18	10,19	10,17	
I.R.	8	5	5	5	8	5	8	7	7	5	5	5	
B.P.	10,16	11,15	12,14	13,13	14,13	15,13	16,13	17,12	18,11	18,10	19,8		
I.R.	5	6	7	8	5	5	6	5	6	5	7		
B.P.	17,7	16,8	16,7	16,6	16,5								
I.R.	5	5	5	6	6								

TABLE 4.7 - Infiltration Rates Fairfield-New Baltimore Aquifer. Units: $ft^3/sec.*100$

Well Coordinates	Well Name	PUMPING PERIODS										
		1952	1953	1954	1955	1956	1957	1958	1959	1960	1961	1962
13,8	A-2	155	155	155	155	155	155	155	155	155	155	155
13,11	S-1 S-2	1512	1835	1762	2155	2031	2260	2019	2298	2223	2004	1951
7,21	F-10	0000	0000	500	0000	338	377	381	372	356	354	357
8,21	F-11	0000	500	0000	0000	423	471	477	465	445	443	446
8,20	F-16	500	0000	0000	500	338	377	381	372	356	354	357

TABLE 4.8 - Pumping history Fairfield-New Baltimore aquifer. Figures are given in $ft^3/sec.$ * 100. Data from 1957-62 were not used in the identification of T.

Next step in the modeling of the Fairfield-New Baltimore aquifer is the identification of the transmissivity function in equation (4.5).

4.4.3 Identification

In order to pursue the calibration of the equation model (4.5), input output information describing the evolution of the system for a given period of time is required. The time period between years 1952 to 1962 was chosen for the identification and validation process. This period has the best available records for these purposes. The period 1952 - 1962 was divided in two subperiods: data from 1952 to 1956 was used in the model identification and from 1957 to 1962 in the model validation.

Infiltration rates and the complete pumping history of the region from 1952 to 1962, which were obtained from the Miami Conservancy District, are presented in Tables 4.7 and 4.8. A breakdown per month can be obtained from, [Spieker, 1968]. Location of the pumping wells is displayed in Figure 4.4.

Unfortunately, drawdown records were not as abundant as pumping records and it was only possible to get water level information for two wells, named BU-7 and H-2, whose location is shown in Figure 4.4.

In spite of the scarce water level information available, it was decided to run the identification algorithm for two reasons: (1) Scarcity of information is a usual problem the hydrologist has to face. (2) No minimum requirements on the amount of information needed

to calibrate the model has been made in this study. It is felt that the more information available on the system's evolution, the more accurate the model calibration can be performed. The present problem seems a good case for testing this hypothesis. As a consequence, two experiments were planned.

In experiment 1 the model calibration was performed by using the only two water head records available. In experiment 2 the model calibration was performed by using results from the analog model by Spieker [Spieker, 1968] for the same problem area. Specifically, his model equation (same as[2.1]) was solved using the parameters and conditions that he determined for the area, thus the additional water head information generated via Spieker's mathematical model was used as "observation' water head data in experiment 2. This provided additional water head estimates for the six pumping wells of the region plus their surroundings (4 adjacent nodes). A total of 25 observations per year or 125 observations for five years were generated and used in the identification process. The main purpose of the two experiments was to compare the performance of the algorithm under scarce vs. non-scarce information conditions. Presentation of results follows.

Table 4.9 summarizes the characteristics of the aquifer under study. Table 4.10 presents drawdown observations based on hydrographs of observation wells BU-7 and H-2. Figure 4.5 indicates the constant head and recharging boundaries. Tables 4.11a - 4.11b present the additional information generated based on Spieker's model, and used in experiment 2. This information was input into the identification algorithm. The presen-

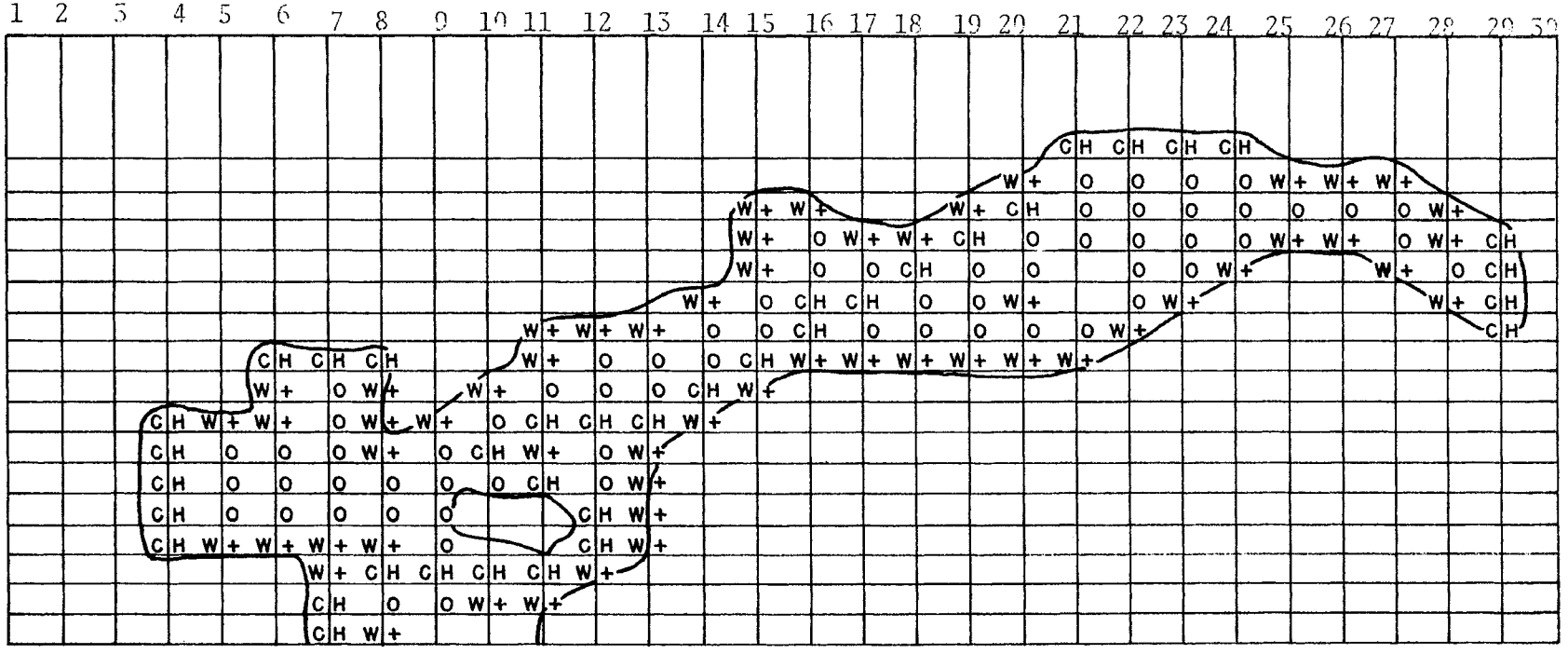
Characteristic	Description
Aquifer Type	Unconfined. Small marginal areas are of semi-confined type
Storage Coefficient S	0.145 average (dimensionless)
Transmissivity Coefficient T	Unknown (ft^2/sec)
Initial Head	550 (ft) average water table height
Boundary Conditions	East & West bounds: Constant Head North & South bounds: 6.8 (mgd) flow on the average
Wells	Six pumping wells distributed in three areas: Southwestern Co. (13.8 mgd) Hamilton South (7.5 mgd) Atomic Energy Commission (1.0 mgd)
Approximate Area	30 square miles (a 20 x 30 node grid has been used to approximate this area)

Acquifer Data: Fairfield-New Baltimore

TABLE 4.9

Well Coordinates	Well Name	WATER HEADS: OBSERVED VS. PREDICTED						
		1952	1953	1954	1955	1956	1957	1958
8,19 Observ.	BU-7	0.00	1.50	2.00	1.00	3.50	--	--
8,19 Predict.	BU-7	2.56	1.99	1.60	3.15	4.50	--	--
14,12 Observ.	H-2	5.00	6.00	6.50	3.50	6.00	--	--
14,12 Predict.	H-2	2.77	4.41	4.77	5.61	5.67	--	--

TABLE 4.10- Comparison of water heads, experiment 1. Water head is measured in feet.



COMPUTER DISCRETIZATION OF THE AQUIFER AREA. CH DENOTES CONSTANT HEAD BOUNDARIES, W^+ RECHARGING BOUNDARIES, AND W^- PUMPING WELLS.

FIGURE 4.5

YEAR = 1952			YEAR = 1953		
OBSERVATION POINT		DRAWDOWNS (ft)	OBSERVATION POINT		DRAWDOWNS (ft)
6	19	0	6	19	0
6	20	-511	6	20	-606
6	21	-611	6	21	-1043
6	22	-534	6	22	-1060
7	19	-745	7	19	-648
7	20	-1459	7	20	-1385
7	21	-1187	7	21	-2085
7	22	-879	7	22	-1768
8	19	-1514	8	19	-1165
8	20	-3459	8	20	-2199
8	21	-1853	8	21	-4183
8	22	-1219	8	22	-5298
9	19	-1311	9	19	-1284
9	20	-1950	9	20	-2003
9	21	-1633	9	21	-2737
9	22	-1337	9	22	-2581
12	8	-503	12	8	-586
12	10	-41	12	10	-51
13	7	-556	13	7	-629
13	8	-1338	13	8	-1425
13	9	-462	13	9	-513
13	11	-6575	13	11	-8086
13	12	-2071	13	12	-2063
14	8	-566	14	8	-641
14	12	-726	14	12	-942

TABLE 4.11a - Water Head Observations Used in Experiment 2.

YEAR = 1954			YEAR = 1955		
OBSERVATION POINT		DRAWDOWNS (ft)	OBSERVATION POINT		DRAWDOWNS (ft)
6	19	0	6	19	0
6	20	-840	6	20	-646
6	21	-1669	6	21	-861
6	22	-1436	6	22	-831
7	19	-656	7	19	-871
7	20	-1702	7	20	-1715
7	21	-3751	7	21	-1558
7	22	-2187	7	22	-1278
8	19	-887	8	19	-1728
8	20	-1575	8	20	-3788
8	21	-2304	8	21	-2275
8	22	-2033	8	22	-1676
9	19	-917	9	19	-1576
9	20	-1380	9	20	-2314
9	21	-1767	9	21	-2074
9	22	-1859	9	22	-1818
12	8	-575	12	8	-576
12	10	-53	12	10	-53
13	7	-640	13	7	-642
13	8	-1436	13	8	-1438
13	9	-520	13	9	-521
13	11	-7793	13	11	-9499
13	12	-2522	13	12	-3067
14	8	-652	14	8	-654
14	12	-920	14	12	-1117

TABLE 4.11a - Continued.

YEAR = 1956

OBSERVATION POINT		DRAWDOWNS (ft)
6	19	0
6	20	-1427
6	21	-2437
6	22	-2274
7	19	-1496
7	20	-3306
7	21	-5105
7	22	-3609
8	19	-2623
8	20	-5271
8	21	-6351
8	22	-4441
9	19	-2643
9	20	-3992
9	21	-4664
9	22	-4434
12	8	-576
12	10	-53
13	7	-642
13	8	-1438
13	9	-521
13	11	-8987
13	12	-2916
14	8	-654
14	12	-1069

TABLE 4.11a - Continued

YEAR = 1952			YEAR = 1953		
OBSERVATION POINT		DRAWDOWNS (ft)	OBSERVATION POINT		DRAWDOWNS (ft)
6	19	-.000000	6	19	-.000000
6	20	-.530857	6	20	-.623984
6	21	-.619037	6	21	-1.039144
6	22	-.543416	6	22	-1.054990
7	19	-.756806	7	19	-.657858
7	20	-1.467938	7	20	-1.386740
7	21	-1.186875	7	21	-2.051189
7	22	-.883659	7	22	-1.735517
8	19	-1.511098	8	19	-1.161544
8	20	-3.421866	8	20	-2.171297
8	21	-1.823468	8	21	-3.992072
8	22	-1.214679	8	22	-2.514910
9	19	-1.282507	9	19	-1.266367
9	20	-1.904663	9	20	-1.970619
9	21	-1.601156	9	21	-2.651348
9	22	-1.329860	9	22	-2.506758
12	8	-.512047	12	8	-.576988
12	10	-.043959	12	10	-.054620
13	7	-.566957	13	7	-.640974
13	8	-1.320580	13	8	-1.405682
13	9	-.468210	13	9	-.518913
13	11	-7.214109	13	11	-8.879487
13	12	-2.087315	13	12	-2.620361
14	8	-.539650	14	8	-.607309
14	12	-.633808	14	12	-.814830

TABLE 4.11b - Water Heads Predicted by Model from Experiment 2.

YEAR = 1954			YEAR = 1955		
OBSERVATION POINT		DRAWDOWNS (ft)	OBSERVATION POINT		DRAWDOWNS (ft)
6	19	-.000000	6	19	-.000000
6	20	-.862288	6	20	-.672268
6	21	-1.655571	6	21	-.870826
6	22	-1.416339	6	22	-.840091
7	19	-.669419	7	19	-.887468
7	20	-1.707635	7	20	-1.729666
7	21	-3.681129	7	21	-1.555635
7	22	-2.132637	7	22	-1.275206
8	19	-.893308	8	19	-1.729634
8	20	-1.573498	8	20	-3.753715
8	21	-2.253234	8	21	-2.238414
8	22	-1.988275	8	22	-1.661059
9	19	-.913395	9	19	-1.548088
9	20	-1.371126	9	20	-2.268211
9	21	-1.736283	9	21	-2.032237
9	22	-1.825383	9	22	-1.798155
12	8	-.585905	12	8	-.587250
12	10	-.056273	12	10	-.056523
13	7	-.651824	13	7	-.653544
13	8	-1.415879	13	8	-1.417400
13	9	-.525412	13	9	-.526371
13	11	-8.559774	13	11	-10.431515
13	12	-2.538707	13	12	-3.085838
14	8	-.616953	14	8	-.618459
14	12	-.793580	14	12	-.964532

TABLE 4.11b - Continued.

YEAR = 1956

OBSERVATION POINT		DRAWDOWNS (ft)
6	19	.000000
6	20	1.467000
6	21	-2.428633
6	22	-2.252329
7	19	-1.520619
7	20	-3.313951
7	21	-5.029217
7	22	-3.539490
8	19	-2.621887
8	20	-5.215648
8	21	-6.138299
8	22	-4.328899
9	19	-2.605573
9	20	-3.921900
9	21	-4.550408
9	22	-4.333839
12	8	-.587463
12	10	-.056562
13	7	-.653824
13	8	-1.417642
13	9	-.526521
13	11	-9.871262
13	12	-2.633577
14	8	-.618703
14	12	-.920827

TABLE 4.11b - Continued.

tation of results is divided in two parts:

Experiment 1:

The identification process was stared with the initial guess

$$T_o = (x + y)10^{-8} + 0.5$$

The initial guess was based on transmissivity information of the area (see Figure 4.4). After 3 iterations the algorithm could not find any better solution than the one presented in Table 4.12. The initial error was reduced from 109 ft^2 to 27.6 ft^2. Notice that the error function is composed of 10 square error terms since five years are being considered for the identification process (1952-1957) and only two observations are available per year. A comparison between the observation data and the optimal solution is made in Table 4.10. It is noted that the observations of well H-2 are matched better than those of BU-7. A maximum difference of 2.8 feet between calculated and observed heads of the following years match better. On average, the maximum differences registered are of 1.5 feet.

There are two sources of error that could explain the lack of output matching between the real system and the model. First are the observation heads themselves. The hydrographs of observation wells BU-7 and H-2 are the resultant of two components: (i) The river fluctuations, and (ii) pumping. Since it is impossible to track the river fluctuations, it is not possible to know accurately how much of the drawdown indicated by the hydrographs corresponds to pumping effects. However, a long term trend can be detected and based on this trend the observation data were

Parameters	Starting Value	Iteration Number 1	2	3	4	5	6	7	Correct Values
b_1 X 10^{+6} (ft/sec)	0.01	0.490	4.80	5.34					UNKNOWN
b_2 X 10^{+6} (ft/sec)	0.01	-0.564	-10.54	-9.82					UNKNOWN
b_3 (ft/sec)	0.5	0.413	0.346	0.345					UNKNOWN
Error Function $\emptyset$ (ft^2)	109.0	67.76	27.60	42.20					
λ- Parameter	0.000	1.00	0.10	27.60					

TABLE 4.12 - Results identification of real aquifer.. Experiment 1 .

chosen. As a consequence, a $\pm$ 2 feet difference is to be expected between the model and the system. Second is the small amount of information available. An aquifer whose areal extension is 30 square miles is being identified with only two observation points. The matching of water heads seems reasonably good, yet nothing can be said until a general drawdown map of the area can be obtained for the same time periods considered in this study. The fact that the model matches a pair of points reasonably well does not mean it is emulating the total system well. This is why the validation stage in any modeling process is so important. The validation results are presented in the next section and they shed some light on the characteristics of the identification process as well as on the minimum information requirement levels.

Experiment 2:

In this experiment the additional information presented in Table 4.11a, and based on the solution of Spieker's model equation, is used in the model calibration process. The identification algorithm was initialized with

$$T = (x + y)10^{-6} + 0,1$$

where the transmissivity coefficients were selected based on previous performance of the algorithm with small values of b_1 and b_2 rather than on any information on transmissivity values.

The algorithm performed quite well in this case. The initial error was reduced from 6065 square feet to 3,5 square feet in six iterations. A comparison of the real drawdown values (Table 4,11a) and

the model's predicted drawdown (Table 4.11 b) shows generally good agreement between them, in some cases, up to the fourth decimal place. Again, perfect matching can not be expected since the model does not represent exactly the real system.

The results obtained here seem to be good, yet nothing can be said until it is proved that the model is in fact calibrated. In other words, if the model is well calibrated, then it should be capable of predicting future behavior of the system under any given conditions. This is the validation step which is taken care of in the next section.

The results of the identification are displayed in Table 4.13. Notice that between the starting and the optimal parameter values there is a difference of about one order of magnitude.

4.4.4 Model Validation

The main goal of this section is to test whether the models identified in the previous section are capable of predicting the real system behavior or not. The relevance of this verification is emphasized in dealing with the results of experiment 1.

Since the calibration process was performed using historical records from 1952 to 1956, the validation test consists of checking the performance of the model for conditions other than the ones used in the identification. In 1962, the hydrologic conditions of the real system were determined, [Spieker, 1968]. A drawdown map of the aquifer describing the state of

Parameters	Starting Value	Iteration Number 1	2	3	4	5	6	7	Correct Values
b_1 X 10^{+6} (ft/sec)	0.1	1.002	3.700	6.929	9.575	10.98	11.14		UNKNOWN
b_2 X 10^{+6} (ft/sec)	0.1	0.3322	2.597	4.930	7.169	8.625	9.062		UNKNOWN
b_3 (ft/sec)	0.1	0.2485	0.2160	0.1440	0.04644	-0.03625	-0.03524		UNKNOWN
Error Function Ø (ft^2)	6,065.	1,716.	251.0	50.72	6.951	6.311	3.570		
λ- Parameter	0.000	0.000	0.000	0.000	0.000	1.0E-4	2.000		

TABLE 4.13 - Results identification of real aquifer.Experiment 2.

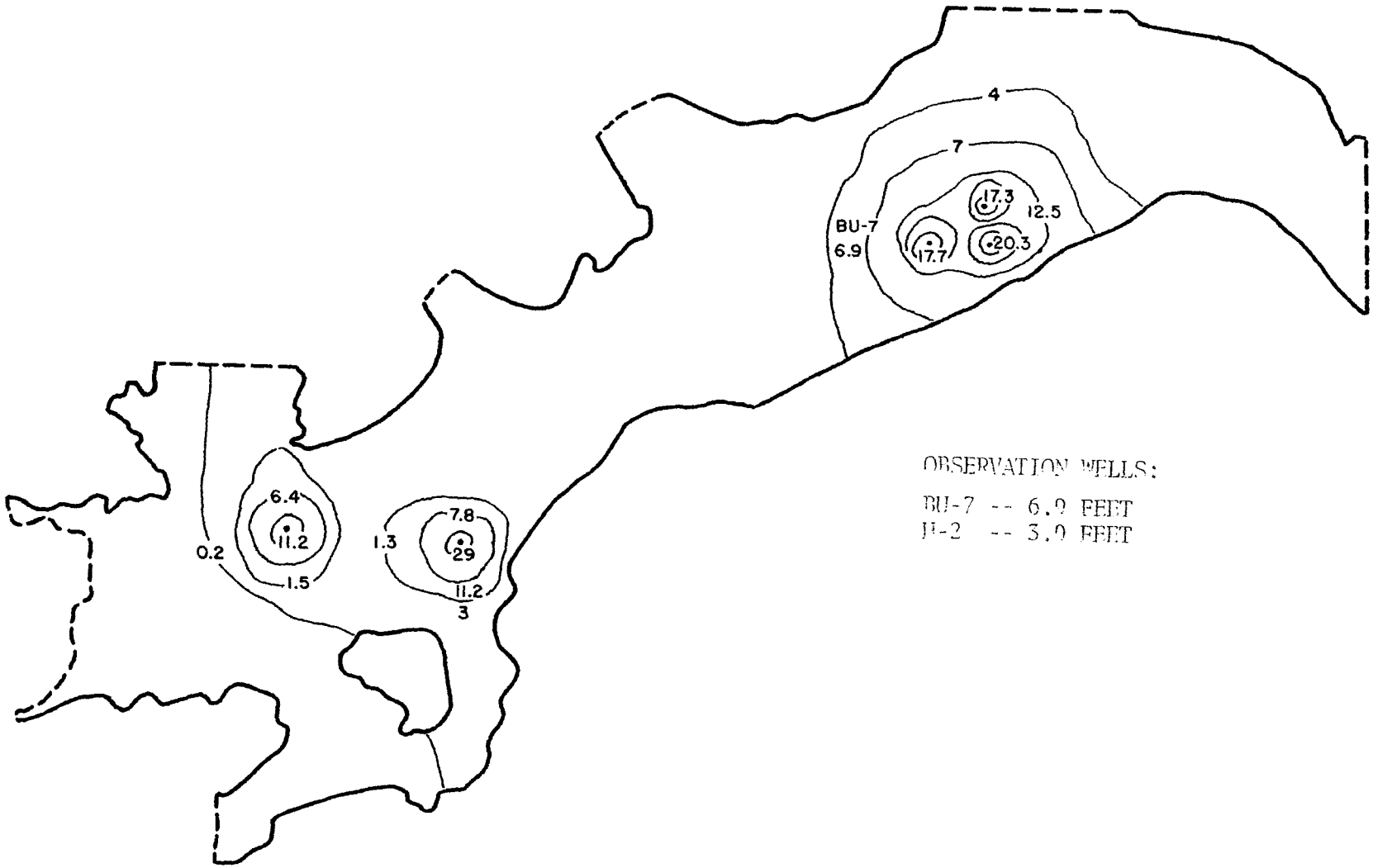
the system in November 1962 is shown in Figure 4.6. The model verification test consists of driving the aquifer model from 1952 to 1962 (using the same input conditions as in the real system) and comparing the model's output in 1962 with the conditions displayed in Figure 4.6. If the outputs match each other, the model is said to be validated.

Model From Experiment 1:

The model calibrated in the previous section showed a reasonably good matching of drawdowns at the observation wells BU-7 and H-2 yet the results obtained in the verification test, and presented in Figure 4.7, shown a very poor prediction of the actual conditions of the systems in 1962. Notice, that the drawdowns at BU-7 and H-2 are matched accurately though. This points out how misleading it is to attempt to identify a large system with so little information. Local conditions were reflected in the identification process accurately. Global aquifer conditions were not reflected in the model since the observation points BU-7 and H-2 did not convey enough information about those general characteristics.

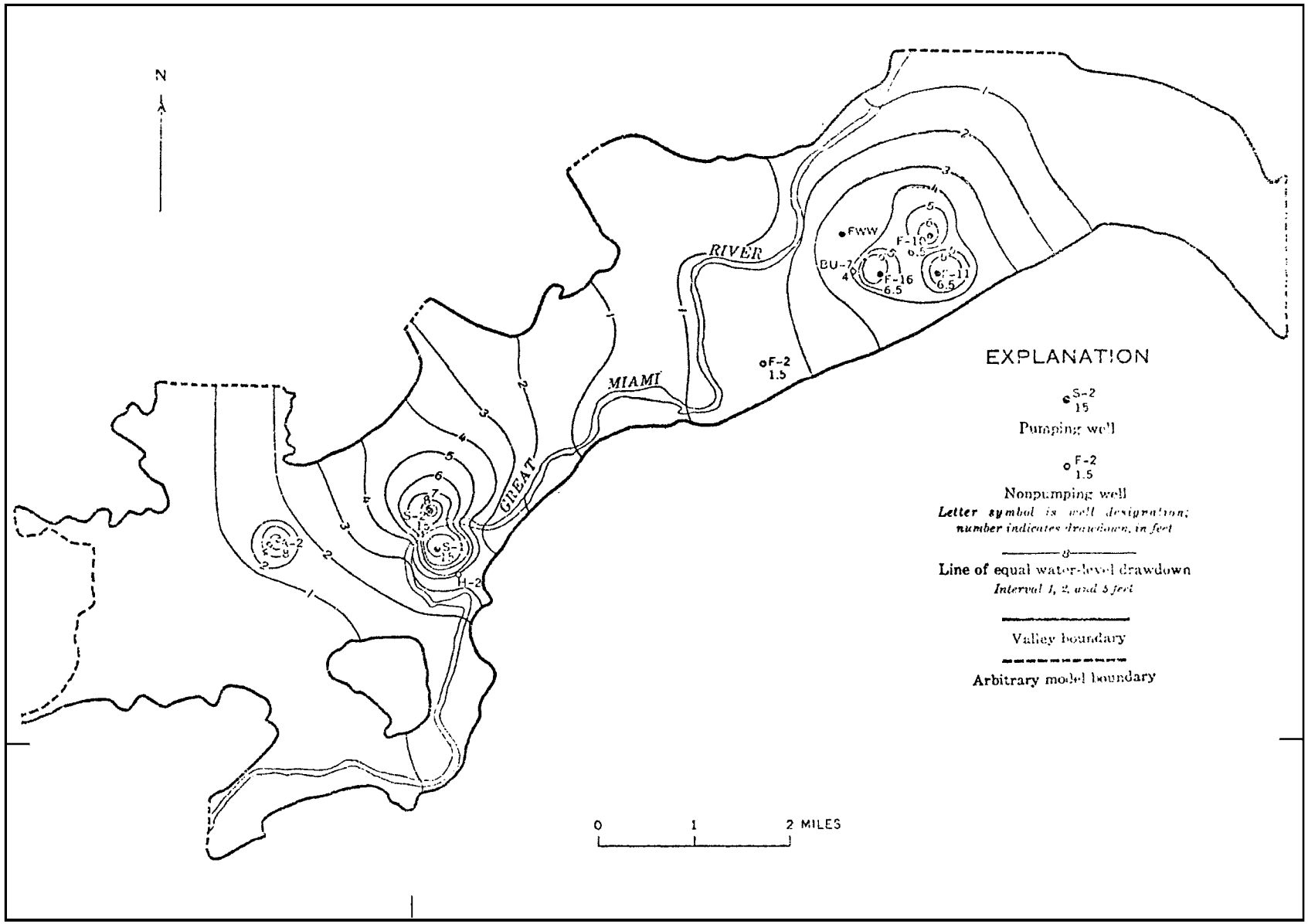
The worst prediction was made at the Southwestern well field where the real drawdown is 15 feet and the model predicted 29 feet. The results of this experiment are very important for two reasons:

(i) Show that matching of the observations does not guarantee a well calibrated model,



DRAWDOWNS CAUSED BY PUMPING FOR THE PERIOD 1952-62,
BASED ON DIGITAL MODEL DERIVED FROM EXPERIMENT 1.

FIGURE 4.6



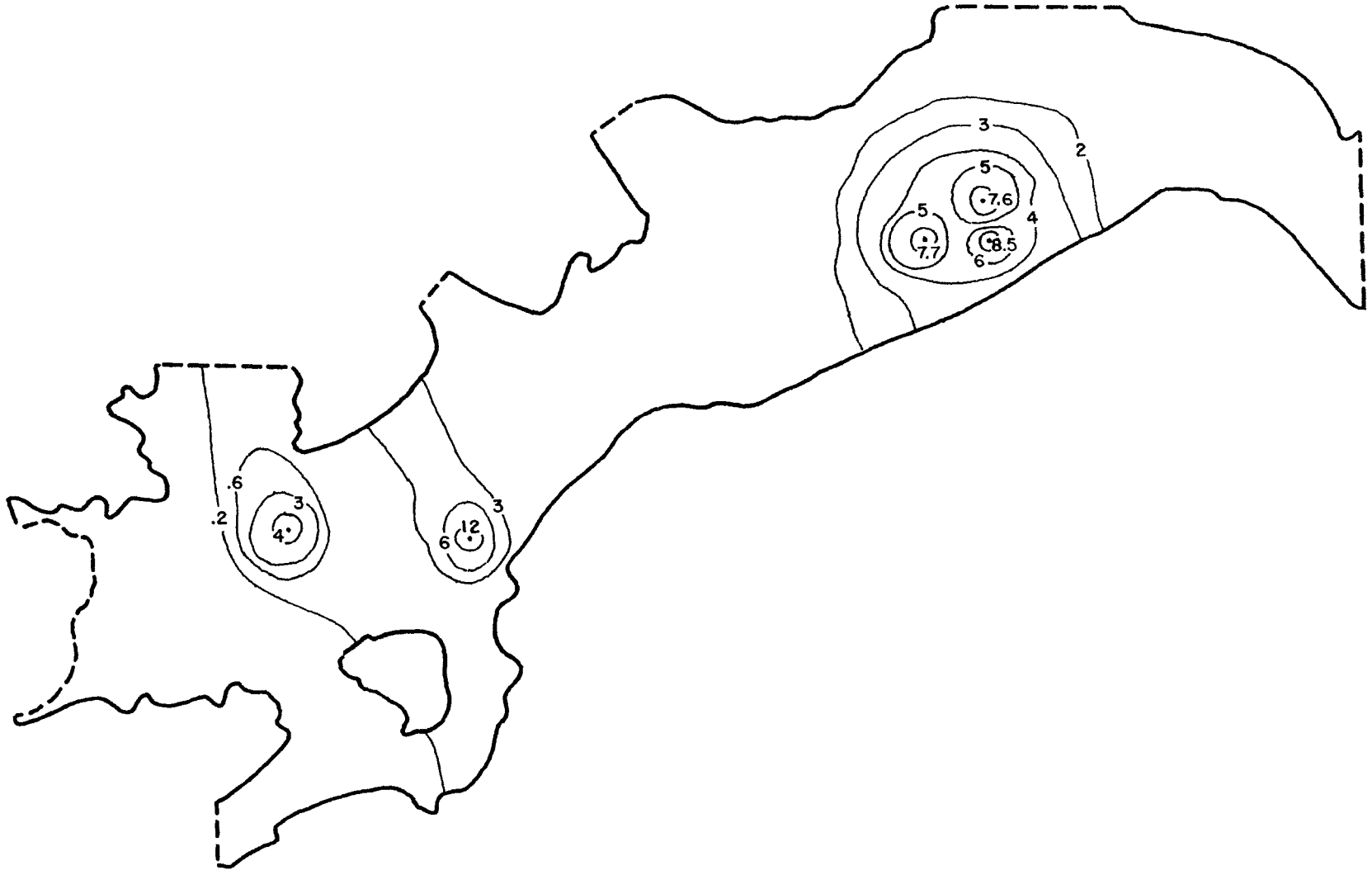
DRAWDOWNS CAUSED BY PUMPING FOR THE PERIOD 1952-62. REAL SYSTEM OBSERVATIONS MADE ON NOVEMBER 1962 (after Spieker).

FIGURE 4.7

(ii) Indicate that observation points far from the location of the pumping sources are not advisable for the identification of the system since these points do not reflect properly the system's characteristics of the pumping areas.

Model from Experiment 2:

The verification of this model was carried out in the same way described above. Figure 4.8 presents the drawdown map as predicted by the model. Comparison of this map with the actual drawdown map of Figure 4.6 shows good agreement between them except for the southwestern area where the comparison becomes difficult because the two main wells were lumped as one in the model. There were two reasons for lumping these wells: (i) the pumping history of these wells is aggregated, (ii) the dimension of the grid used in the discretization of the model equation are such that the wells appear as aggregated. A refined grid will produce better resolution of this area if the need for a more accurate answer arises. However, the drawdown predicted by the model at the well's location is close to the actual. The comparison of drawdowns reports a good model matching of local and global conditions of the system. Therefore, the model from experiment 2 is assumed valid and ready to be used in further studies. The model representation is finally given by



DRAWDOWNS CAUSED BY PUMPING FOR THE PERIOD 1952-62,
BASED ON DIGITAL MODEL DERIVED FROM EXPERIMENT 2

FIGURE 4.8

$$\frac{\partial}{\partial x}([1.14x + 0.96 - 0.035]\frac{\partial h}{\partial x}) + \frac{\partial}{\partial y}([1.14x + 0.9y - 0.035]\frac{\partial}{\partial y}) =$$

$$0.145\,\frac{\partial h}{\partial t} + \frac{P(x,y,t)}{(\Delta x)^2}\,\delta(\underline{x}_i - \underline{a})$$

I.C. $h(x,y,0) = 550$ ft.

$h(x,y,t)\big|_{\Gamma_i} = 550$ ft $\quad i = 1, 2, \ldots, 5$ (see Figure 4.4)

B.C. $h(x,y,t)\big|_{\Gamma_6} = 550$ ft River condition

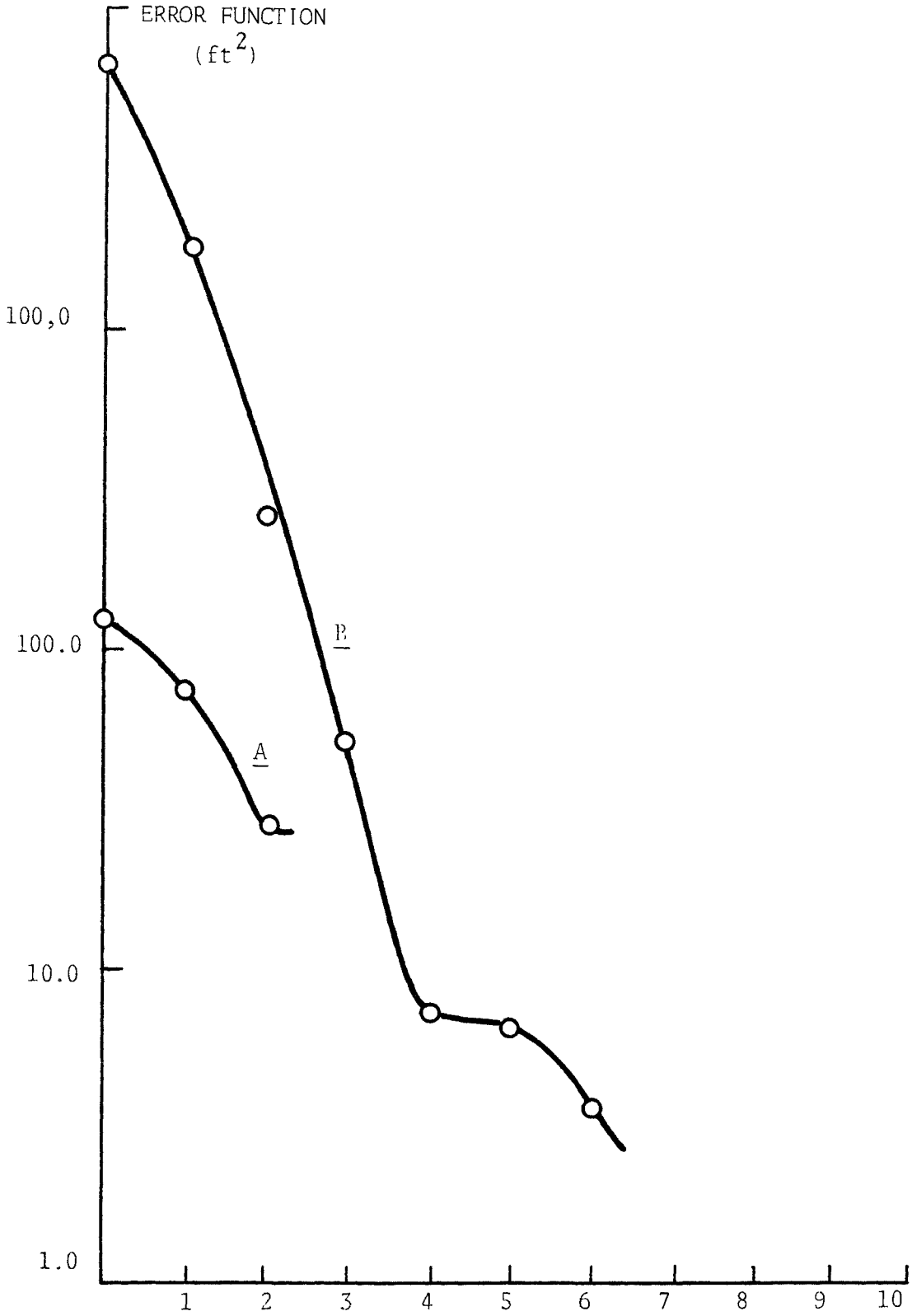
$\frac{dh}{d\bar{n}}(x,y,t)\big|_{\Gamma} =$ Recharge from boundaries (see Table 4.7)

A comparison of the error functions for experiment 1 and 2 is given in Figure 4.9. Validation results for the aquifer digital models from experiments 1 and 2 are given in Table 4.14.

4.4.5 Validity of Results

In order to evaluate the validity of the real example study made in this chapter, it is important to emphasize the limitations of the present analysis and to determine possible sources of error.

Concerning boundaries of the area modeled, pumping history, and drawdowns for the year 1962, the information is well documented and highly reliable, therefore the error source expected from the data is minimized. Regarding the storage coefficient, Spiekers considered the figure of 0.2 to be valid for the unconfined area partly because of the close agreement between the drawdowns determined from his analog model



A, ERROR FUNCTION EXPERIMENT 1.
B, ERROR FUNCTION EXPERIMENT 2.

FIGURE 4.9

Well Coordinates	Well Name	Observed Head(ft)	Predicted Head(ft)	Error %
13,8	A-2	6.0	11.0	83
13,11	S1-S2	15.0	29.0	93
8,20	F-16	6.5	17.7	154
8,21	F-11	6.5	20.8	220
7,21	F-10	6.5	17.3	154
8,19	BU-7	4.0	6.9	72.5
14,12	H-2	4.0	3.0	25

*Model from Experiment 1. Water heads compared on November 1962.

Well Coordinates	Well Name	Observed Head(ft)	Predicted Head(ft)	Error %
13,8	A-2	6.0	4.0	33
13,11	S1-S2	15.0	12.0	20
8,20	F-16	6.5	7.7	18.4
8,21	F-11	6.5	8.7	30
7,21	F-10	6.5	7.5	15.3
8,19	BU-7	4.0	3.13	21.7
14,12	H-2	4.0	1.0	75

Model from Experiment 2. Water Heads compared on November 1962.

TABLE 4.14 - Results Validation Aquifer-Digital Models from Experiments 1 and 2.

*Note that the model from Experiment 1 was identified with only two observation points. Since the validation process pointed out large percentage errors between this model and the real system, the former was rejected.

and those of Figure 4.8. For the semiconfined area he uses values from 0.02 to 0.1. The aquifer digital model used in this study does not take variable storage coefficients. As a consequence the storage coefficient for the model was obtained by averaging the weighted (using the area percentage as a weight) sum of storage coefficients for the aquifer. This introduces some error in the computations which does not seem to affect seriously the final results since a good agreement between calculated vs. actual drawdowns was obtained. In future extensions of this work the relevance of the storage coefficient will be evaluated more carefully. The methodology presented in this report can be readily applied to handle the variable storage coefficient case providing the aquifer model be modified.

Concerning the interaction aquifer-river and aquifer-boundaries, the rates of induced recharge are the most difficult variables to estimate in this study and the ones which may produce the largest errors. An important factor in the present analysis is the rate of recharge of the river to the aquifer. Since the river is a large source of water and the infiltration rates determined at two different points on the aquifer are high, it is assumed the river can be handled as a constant head boundary condition. The consequences of this assumption are observed in Figure 4.8. The model results for the Southwestern area show smaller drawdowns than the actual system, indicating that the recharge contribution of the river to the aquifer is overestimated. Notice, however, that the error induced by this overestimate does not affect seriously the prediction at the well itself. A more accurate

result demands a more accurate modeling of the river. This would require more knowledge of the infiltration rates. At the present moment this information is not available.

Regarding the interaction aquifer-boundaries, it is noted that the rate of leakage from the bedrock valley walls has not been determined. The figures used in this study are based on those estimated by Spieker [Spieker, 1968]. He considers those estimates correct, partly because the results of his analog model simulation seem to match well the real conditions observed in the system, and also because of the fact the river is the largest recharger of the region; so an error in the estimation of these infiltration rates would not affect seriously the drawdown distributions. Another source of error comes from the discretization of the aquifer area in order to solve equation (2.1). A comparison of Figures 4.4 and 4.5 reveals the gross approximations made in the representation of the real area's geometry. Again, a trade-off between accuracy of the model's solutions and computer time is the main reason for adopting this representation (i.e., the one in Figure 4.5).

Finally, the largest source of error in this study may come from the fact that in the identification process observation data was generated by a model rather than from the real system itself (since this was not possible). However, there is confidence in the data generated for two reasons: (i) the model derived with this data showed a good agreement with the real system. (ii) the model source of information has been verified and found completely acceptable, [Spieker, 1968].

Chapter 5

EXTENSION OF AQUIFER MODEL DEVELOPMENT: A MULTICELL MODEL

5.1 Introduction

A technique has been proposed [Maddock, 1972] to model groundwater systems via a simulation model for which an "Algebraic Technological Function" (A.T.F.) may be calculated. The principal idea underlying the A.T.F. technique is a linearization of the system model. This makes possible a linear superposition of the system's response to water head elevation charges, which can be effectively coupled with a management model.

The A.T.F. has several disadvantages, however. While it may be appropriate for systems governed by a single partial differential equation, application to systems comprising portions governed by effectively different equations may make the modeling process difficult. Another disadvantage occurs for the case in which the system consists of combined unit aquifers. Although each unit is affected by the others, an input from within a unit has a greater influence than an input from outside units. Thus, points within and outside a given unit deserve different weightings in the construction of the model. Finally, for any real system, it is likely that detailed analysis will require extensive computer capacity and consequently may prove to be an important restriction. For this case

direct use of the A.T.F. technique may prove to be inadequate.

Decomposition of the groundwater system into different cells combined with the A.T.F. approach constitutes a new approach. It appears that it may provide an improved solution to the problems discussed above. An accurate analysis and modeling procedure is made possible for each cell, and the dimensionality of the model is reduced.

5.2 Model Formulation

The multicell approach to modeling groundwater makes use of a set of balance equations, of which each represents a mass balance applied to a particular cell. For a single cell representing an area, within an aquifer and surrounded by impervious boundaries, the balance equation takes the form [Bear et al, 1972],

$$Q \cdot \Delta t = (h(t + \Delta t) - h(t)) \cdot A \cdot S$$

where

Δt = Period for which the balance is written,

Q = Net inflow into the cell

A = Area of cell

$h(t)$ = Average water level elevation in the cell at time t

S = Aquifer storativity at the cell (averaged).

Applying the same principle of water balance to a multicell system, taking into account the interflow between adjacent cells, leads to a set of difference equations. The form of these equations is identical to the form of those which result from the discretization of an ordinary differential equation when used to approximate the aquifer system.

The thickness of an aquifer usually is small compared with its lateral dimensions. For an unconfined flow in nonhomogeneous medium, in which the storage coefficient is assumed to be independent of water table elevation while transmissivity is not, the following difference equation may be used [Yu and Haimes, 1973].

$$\sum_j \{\bar{C}_{ji}(h_{jm} - h_{im}) - \bar{k}_{ji}[(h_{jm})^2 - (h_{im})^2]\} = \bar{S}_i (h_{i\,m+1} - h_{im}) - Q_{im} \qquad (5.1)$$

where

$$\bar{C}_{ji} \triangleq \frac{W_{ji}C_{ji}}{L_{ji}} \qquad \bar{k}_{ji} \triangleq \frac{W_{ji}k_{ji}}{2L_{ji}} \qquad \bar{S}_i \triangleq \frac{A_iS_i}{\Delta t}$$

$$C_{ji} \triangleq k_{ji}(E_{ji} - F_{ji})$$

h_{im} = Water table elevation at the i-th cell during the m-th time step.

Q_{im} = Net inflow into the i-th cell during the m-th time step.

w_{ji} = Length of the perpendicular sector associated with the segment between cells i and j.

L_{ji} = Distance between the centers of nodes i and j.

K_{ji} = Hydraulic conductivity averaged between cells j and i.

E_{ji} = Effective aquifer depth averaged between cells j and i.

F_{ji} = Elevation at the top of the aquifer averaged between cells j and i.

A_i = Area of i-th cell.

S_i = Storage coefficient averaged over the i-th cell.

One should note that the multicell approach is very much simplified by comparison with the real system. Boundary conditions must be simplified as well. Constant flow may be handled through inflow to a particular cell. Constant head requires a fixed head for the cell at all times. No flow requires that the hydraulic conductivity be set at zero between cells and the construction of an imaginary neighboring cell.

The multicell model provides approximate inflows and outflows for each cell in the modeling procedure. These values may be computed for each time step together with averaged water heads. To obtain an accurate estimate of drawdown at a particular point of interest one can isolate the cell in which the point of interest is located. This cell may then be modeled in greater detail using a mathematical model which takes into account the particular boundary conditions which relate to its adjacent cells, as functions of time.

To illustrate the usefulness of the proposed modeling technique,

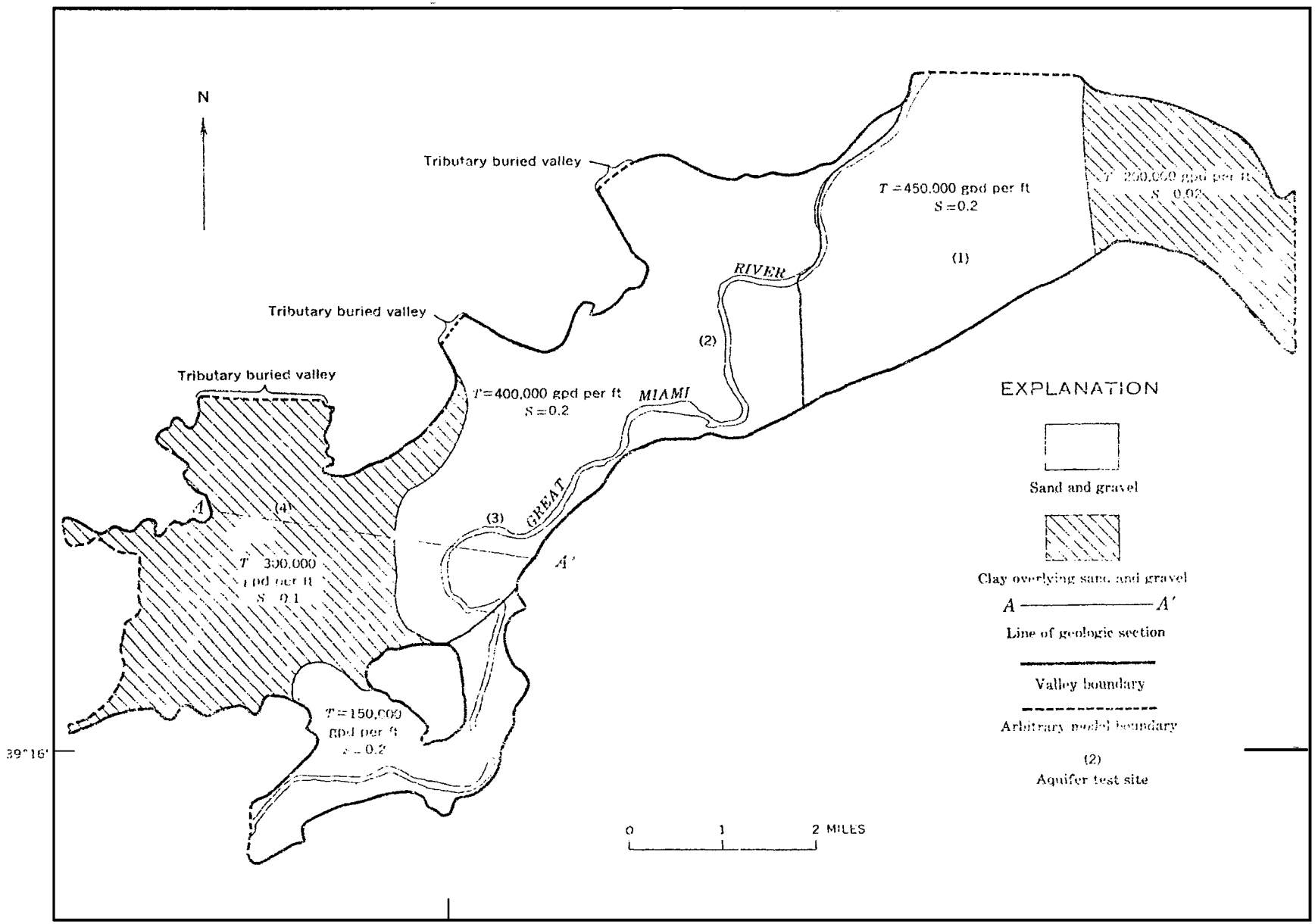
the Fairfield-New Baltimore aquifer in the lower Great Miami River Valley was modeled. A computer program was written to simulate the aquifer. The system was divided into cells with differing characteristics (see Figure 5.1). The data utilized included pumpage, water elevations, and cell boundary conditions, and were taken from Spieker, [Spieker, 1968] (as also used in chapter 4). An explicit computation scheme can be used, if care is taken to avoid the stability problem by choice of an appropriately small time step. The semi-pervious bedrock which forms the natural boundaries for the groundwater system can be handled as part of the water balance for each cell (constant inflow). The river can be handled as constant head cells. Initial water head values in all cells are part of the input to the program. For each time period (one year) the forcing function (pumpage) at each cell is given.

The simulation model can produce two types of output:

(i) For each time period, the interflow between adjacent cells is provided.

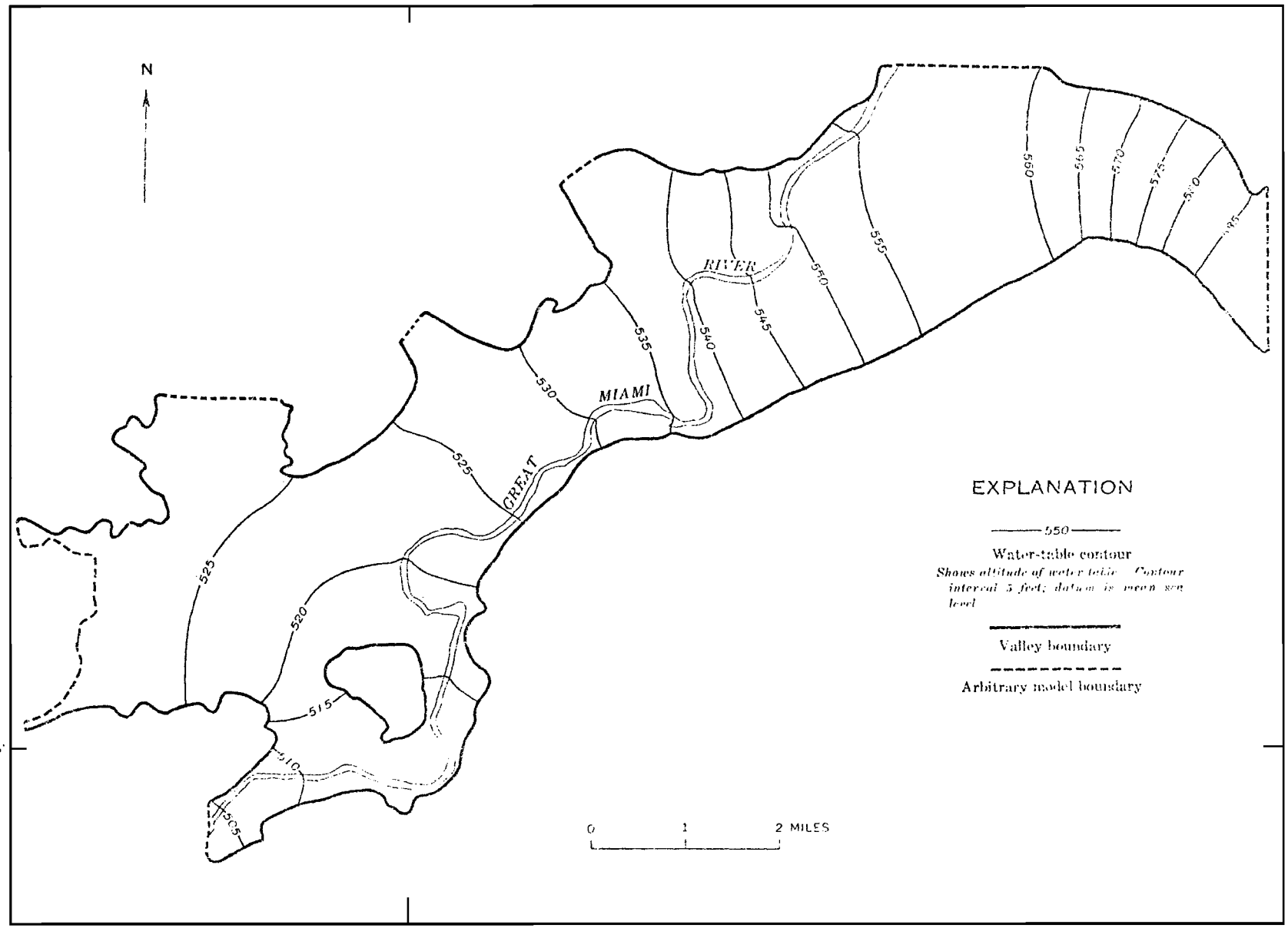
(ii) For each time period the averaged water level is predicted in all cells.

Cell 4 (see Figure 5.2) was selected to be simulated by means of a particular model. Maddock's program [Maddock, 1972] was used for this purpose. The square grid used in chapter 4 was implemented for cell 4 only. The results from the multicell model for boundary conditions between cells provide the data for the related nodes. The program is described in detail in Maddock [Maddock, 1972]. The method used for



GENERALIZED GEOLOGY AND COEFFICIENTS OF TRANSMISSIBILITY (T) AND STORAGE (S) OF THE FAIRFIELD-NEW BALTIMORE AREA.

FIGURE 5.1



CONTOURS SHOWING ALTITUDES OF THE WATER TABLE ASSUMED TO HAVE PREVAILED BEFORE PUMPING WAS BEGUN IN THE REPORT AREA.

FIGURE 5.2

solving the problem is the alternating direction implicit technique.

Ten years of pumpage imposed on the system were simulated. The steps for doing this were:

(i) The multicell model was used to determine the initial state befor pumpage began. Running the model over a long time provided a steady state head distribution for the system regardless of the initial values chosen on a random basis, and with no pumpage imposed.

(ii) The steady state calculated at the previous step was used for calibrating the model's parameters (equation 5.1). This was achieved by comparing the results with a map of initial head given by Spieker (Figure 5.2).

(iii) The system was simulated by the multicell model to produce the system response averaged at each cell for a ten year period.

(iv) The cell model using outputs from the previous step was used to simulate cell 4 over the same ten years and so to produce drawdown at wells. Table 5.1 summarizes the results. Three pumping wells are located in cell 4. An analog model (Spieker) and digital model results (see chapter 4) for which the entire aquifer was modeled are compared as are the two-stage results, to the real drawdowns as measured in the aquifer at the end of the simulated time period (see Figure 4.7):

WELL	REAL VALUE	ANALOG	DIGITAL	2-STAGE
F-16	6.50	9.0	7.8	7.2
F-11	6.50	9.0	7.8	8.5
F-10	6.50	9.0	7.8	6.8

DRAWDOWN IN FEET

TABLE 5.1

From the Table 5.1 one may conclude that the suggested modeling technique provides a procedure by which to predict an accurate map of drawdown at different parts of a groundwater system. One can expect that the advantage of this technique is even greater if it is applied to a more complex system for which a single stage model is impractical.

In this study we are primarily interested in the coupling of a groundwater system with a desired management model. Using the proposed technique, it is possible to extend the A.T.F. approach toward handling a more complex system. The basis used is that the drawdown at a point in a aquifer due to input at some other point within the groundwater system may be approximated by a linear combination of responses predicted by the multicell model, and by those predicted by the more detailed model of the unit aquifer to which the point belongs.

Let $\beta(k,j,i)$ be the drawdown at point k due to one unit input at point j occuring i time periods previously during one unit time with k and j located at the ℓ-th unit aquifer. Let $\gamma(\ell,r,i)$ be the average drawdown at unit aquifer 1 due to one unit input in cell r occuring i time periods previously during one unit of time. Let $D(\ell,t)$ be the areal averaged drawdown at the ℓ-th cell at the end of t time periods due to the aggregated input to the whole system:

$$D(\ell,t) = \sum_{r=1}^{R} \sum_{m=1}^{t} \gamma(\ell,r,t-m+1) * Q(r,m)$$

where

R = total number of cells r = 1, ..., R

and

Q(r,m) = Aggregated input of water to the r-th cell area during the m-th time period.

Let $D_\ell(k,t)$ be the drawdown at the k-th point located at the ℓ-th cell at the end of the t time period due to the aggregated input in the ℓ-th cell

$$D_\ell(k,t) = \sum_{j=1}^{M_\ell} \sum_{m=1}^{t} \beta(k,j,t-m+1) * Q_\ell(j,m)$$

where

M_ℓ = total number of points in the ℓ-th cell where input may be introduced,

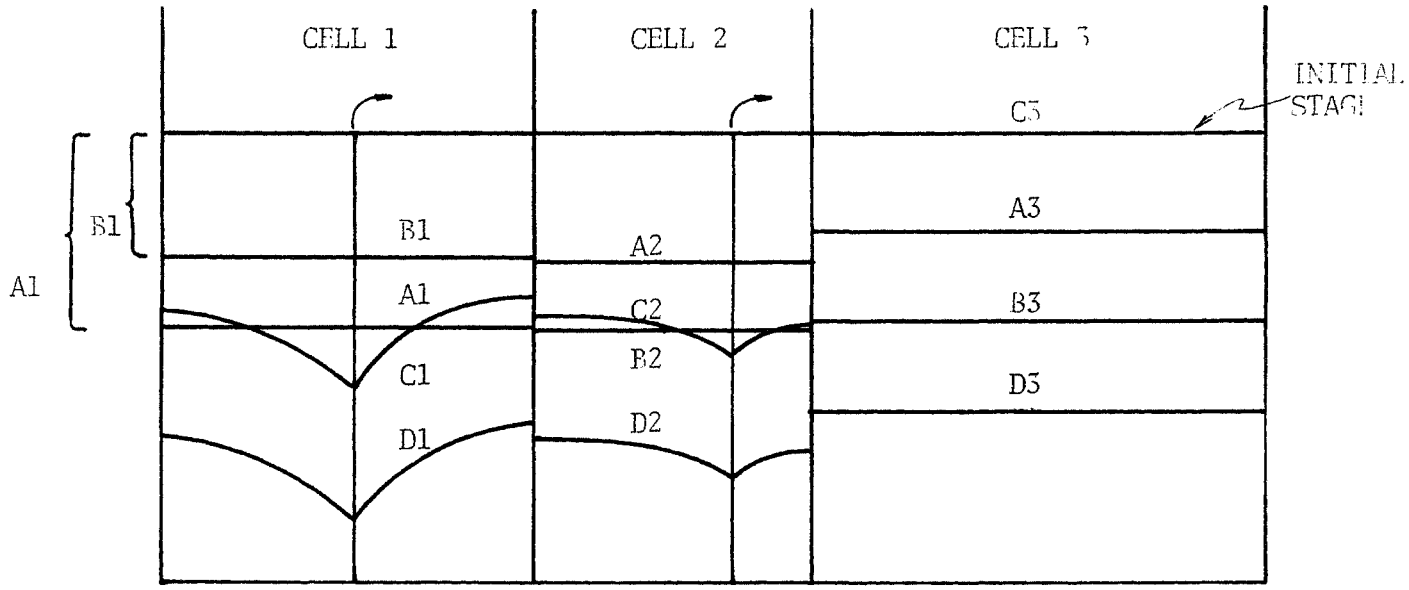
and

$Q_{\ell}(j,m)$ = Aggregated input of water in the j-th point located at the ℓ-th cell during the m-th time period.

Let $DT_{\ell}(k,t)$ be the total drawdown at the k-th point located at the ℓ-th cell at the end of the t time period due to the aggregated input in the whole groundwater system so that

$$DT_{\ell}(k,t) = D(\ell,t) + D_{\ell}(k,t) - \sum_{m=1}^{t} \gamma(\ell,\ell\ t-m+1) * Q(\ell,m)$$

The significance of this last algebraic equation may be illustrated graphically on the schematic section of three adjacent cells. A pumpage is imposed on cells 1 and 2. Figure 5.3 shows the decomposition of the drawdown at each cell due to these wells.



AL - DRAWDOWN DUE TO PUMPAGE AT 1-ST CELL

BL - DRAWDOWN DUE TO PUMPAGE AT 2-ND CELL

CL - PARTICULAR DRAWDOWN DUE TO PUMPAGE AT THE PARTICULAR CELL

DL - COMPOSITION OF DRAWDOWN:

CELL 1: D1 = B1+C1

CELL 2: D2 = A2+C2

CELL 3: D3 = A3+B3

REALIZATION OF THE ALGEBRAIC DECOMPOSITION

FIGURE 5.3

REFERENCES

1. Bear, J., Dynamics of Fluids in Porous Media, American Elsevier, New York, 1972.

2. Bear, J., "Numerical Modeling of Groundwater Systems," Report P.N. 180, Dept. Civil Engineering, Technion-Institute of Technology, Haifa, Israel, June, 1972.

3. Bear, J. & J. Schwartz, "Electric Analog for Regional Groundwater Studies," TAHAL, Consulting Engineers, Ltd., Tel-Aviv, Israel, P.N. 609, 1966.

4. Bear, J., U. Shamir & E. A. Hafez, "Numerical Modeling of Groundwater Systems," P.N. 180, June, 1972.

5. Bellman, R. E. & R. E. Kalaba, Quasilinearization and Non-Linear Boundary Value Problems, American Elsevier, New York, 1965.

6. Birkhoff, G. & S. Varga, "Implicit Alternating Direction Methods," Trans. Amer. Math. Soc., 92, 13-24 (1959).

7. Bredehoeft, J. D. & G. F. Pinder, "Digital Analysis of Areal Flow in Multi-Aquifer Groundwater System. A Quasi Three Dimensional Model," Water Resources Research; in print.

8. Cesari, L., & D. Cowles, "Existence Theorems in Multidimensional Problems of Optimization with Distributed and Boundary Controls," Archive for Rational Mechanics and Analysis, V. 46, No. 5, P. 321-55, 1972.

9. Falkenburg, D. R., "The Identification of Distributed Parameter Systems," Ph.D. Dissertation, Case Western Reserve University, 1971.

10. Fletcher, R., "A Modified Marquardt Subroutine for Non-Linear Least Squares," A.E.R.E. report, R6799, H750, 1972.

11. Haimes, Y. Y., "Identification of Aquifer Parameters by Decomposition and Multilevel Optimization," M.S. Thesis, U.C.L.A., 1967.

12. Haimes, Y. Y., R. L. Perrine & D. A. Wismer, "Identification of Aquifer Parameters by Decomposition and Multilevel Optimization," Israel Journal of Technology, Vol. 6, No. 5, 1968.

13. Jacob, C. E., "Flow of Groundwater," in Rouse, H., (ed) Engineering Hydraulics, John Wiley, New York, 321-386, 1950.

14. Karplus & Kawameto, "Identification Problem in Distributed Parameter Systems," notes, Short Course at U.C.L.A., January 9, 1966.

15. Kleinecke, D. R., "Use of L.P. for Estimating Geohydrological Parameters of Groundwater Basins," Water Resources Research, Vol. 7, No. 2, 1971.

16. Lions, J. L., Optimal Control of Systems Governed by Partial Differential Equations, Springer Verlag, New York, 1971.

17. Lopez, H. "Modeling and Identification of Groundwater Systems," M.S. Thesis, Case Western Reserve University, 1973.

18. Maddock, T., III, "A Program to Simulate an Aquifer Using Alternating Direction Implicit-Iternative Procedure," U. S. Dept. of Interior Geological Survey, 1972.

19. Marino, M. A., & W. G. Yeh, "Identification of Parameters in Finite Leaky Aquifer System," American Society of Civil Engineers Hydraulic Division, Feb., 1973.

20. Marquardt, D. W., "An Algorithm for Least Squares Estimation of Non-Linear Parameters, J. Soc. Indus. Appl. Math., Vol. II, No. 2, June, 1963.

21. Peaceman, D. W., & Rachford, H. H., "The Numerical Solution of Parabolic and Elliptic Differential Equations," J. Soc. Indust. Appl. Math., Vol. 3, March, 1955.

22. Phillipson, G. A., Identification of Distributed Systems, American Elsevier, New York, 1971.

23. Reid, J. K., "Least-Squares Solution of Sparce Systems of Non-Linear Equations by a Modified Marquardt Algorithm," 1972.

24. Ritchmayer & Morton, Difference Methods for Initial Value Problems, John Wiley; New York.

25. Rubin, J., "Theoretical Analysis of Two Dimensional Transient Flow of Water," Proc. Soil Sci. Soc. Am., 32(5), 607-615, 1968.

26. Seinfeld & Chen, "Estimation of Parameters in Partial Differential Equations," Chemical Engineering Science, 1971, Vol. 26, pp. 753-766.

27. Schenke, R. T., & Y. Y. Haimes, "Distributed System Identification by Quasilinearization," SRC Tech. Memo #47, Case Western Reserve University, 1973.

28. Spieker, A. M., "Effect of Increased Pumping of Groundwater in the Fairfield-New Baltimore Area, Ohio," U. S. Geological Survey Professional Paper, 605-C, Washington, 1968.

29. Vemuri, V., & W. J. Karplus, "Identification of Non-Linear Parameters of Groundwater Basins by Hybrid Computation," Water Resources Research, Vol. 5, No. 1, pp. 172-185, 1969.

30. Wismer, D. A., R. L. Perrine, & Y. Y. Haimes, "Modeling and Identification of Aquifer Systems of High Dimension," Automatica, Vol. 6, pp. 77-86, 1970.

31. Yeh, W. G., & G. W. Tauxe, "Quasilinearization and Identification of Aquifer Parameters," Water Resources Research, Vol. 7, No. 2, April, 1971.

32. Yu, W., & Y. Y. Haimes, "Multilevel Optimization for Conjunctive Use of Ground and Surface Water," to be published in Water Resources Research.

APPENDIX A

COMPUTER PROGRAM

A.1 Introduction

A program, written in FORTRAN V for the UNIVAC 1108 digital computer, implements the identification algorithm developed in this report. The program is composed of two main parts: One part includes the optimization procedure and the other includes the procedure for solving the aquifer equation and the sensitivity equations of the system being modeled. To make use of the program the user has to completely define the characteristics of the aquifer under study and the input information needs to be properly added according to the rules defined in this appendix. An input-output historical record will be also required for the system identification phase.

The program is divided into three subprograms, namely: MAIN (optimization section), MAD (aquifer simulator), and MATCH. The latter coordinates the communication between the optimization section and the aquifer simulator. MAIN plays the role of the main program while MAD and MATCH are merely subprograms.

The identification program has been designed in order to calibrate the transmissivity function of confined aquifers models. It is assumed that the aquifer under study can be modeled by a linear parabolic partial differential equation. The mathematical model is already contained within the program and the user has only to specify the aquifer characteristics. Once this is done, the program finds the best transmissivity which is represented by a linear polynomial function for the given model. The resulting output

is a calibrated aquifer model.

The aquifer simulator subroutine (MAD) solves the aquifer equation many times during the optimization process. It also solves the sensitivity equations of the system which provide gradient information used in the optimization. The model equation built in the program handles different pumping pattern conditions, leakage from confining regions with storage capabilities, leakage from lakes and/or streams, irregular-shaped boundaries, and nonhomogeneous transmissivity. The transmissivity function structure has been approximated in the program by a linear polynomial in x and y, the polynomial coefficients being given by BET(1,IKC), BET(2,IKC), and BET(3,IKC). The program allows for future extensions to quadratic or higher polynomial representations for the transmissivity function.

The aquifer simulator requires the boundary conditions of the aquifer be either constant head or constant flow conditions. As for initial conditions, a constant head is assumed throughout the aquifer.

The model calibration is carried out over a time period which is usually divided into equal intervals called pumping periods. The size of the time horizon is determined by the input-output historical record available for the identification process.

MATCH has the main task of setting proper conditions in the aquifer simulator section so that the sensitivity equations can be solved. For example, calculation of the forcing function for the sensitivity equations is one of the tasks performed by MATCH. MAIN is the master control and optimization section of the program which determines the direction and size of the corrections to be made on the initial transmissivity guess.

MAIN operates based on information supplied by the aquifer simulator section (MAD) and the input-output historical record. An operation description of the total program is presented in the following section.

A.2 Application

The identification program can be used for estimating coefficients embedded in systems described by a second order linear parabolic equation. In this project the program has been applied to the identification of an aquifer system, therefore the following description only uses terminology related to the water resource field. However, it is easy to make analogies with any other systems (not necessarily water resource) which can be described by parabolic equations, thus allowing for more general application to the program.

The model equation (2.1) used in this program has been derived under certain assumptions which are rewritten for convenience of the user below. Also the optimization subroutine has requirements which are specified below.

1. It is assumed that the aquifer being modeled can be treated as a two dimensional flow system.
2. Drawdowns relative to the saturated thickness are small.
3. The storage coefficient is constant throughout the aquifer.
4. Boundary conditions can be accurately approximated by constant head or constant flow.
5. The time horizon considered for the transmissivity identification must be divided into equal intervals called pumping periods.

Within a pumping period wells are assumed to pump at constant rates.

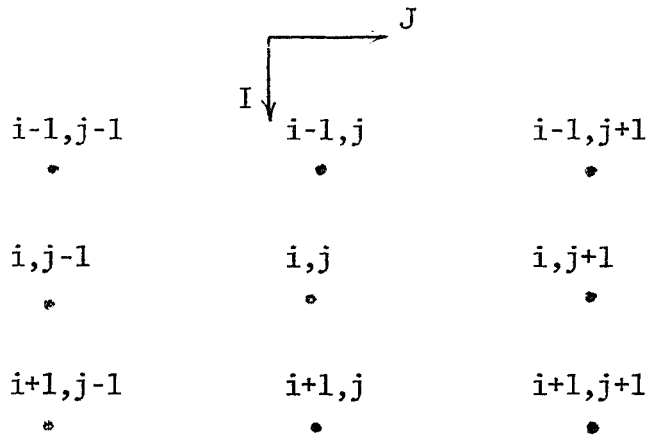
6. The aquifer area is discretized by superimposing a grid on the plan view of the aquifer. Node points represent an area segment. If more than one well exists in a given area segment, their pumping rates are added and treated as single well. A restriction exists on the number og grid points, namely, it can not be larger than 20 x 30 nodes due to computer storage limitations.
7. Input-output historical records covering the system behavior throughout the time horizon considered in the identification are required. No constraints are imposed on the number and location of the observation points. However, it is desirable to select water head measures of as many well fields as possible and for different time periods.
8. Starting values of the transmissivity parameters BET(1,1), BET(2,1), BET(3,1) have to be guessed by the user to initialize the optimization process. If the user has some information about the aquifer transmissivity he may attempt to draw a plane passing through some of the known values. If no or very little information is available on transmissivity, the user is advised to set BET(1,1) and BET(2,1) to 10^{-6} and to choose BET(3,1) between the range [0.2, 0.8].

A.3 Program Description

The identification program was designed following a subroutine approach. The break down of the program into independent subroutines coordinated by a main program greatly facilitates the development as well as the understanding of the program by potential users. There are three major subroutines which make up the program. From the users viewpoint the one which requires more attention is MAD(the aquifer simulator) since it requires a great deal of data and imposes conditions on input operations for other subroutines. Preliminary operations to activate the aquifer simulator subroutine are described first.

A.3.1 Preliminary Operation

First, discretization of the aquifer area by superimposing a square grid on the plan view of the aquifer is performed. This operation defines the frame of reference to be used throughout the program. Then, water head observations, storativity values initial conditions, pumping rates, recharge sources, and leakage at each of the grid nodes are defined. The nodal array of the grid is pictured as follows:



The geometry of the aquifer is defined by means of the matrix of coordinates T. If a node point (I,J) lies otside the boundary of the aquifer then T(I,J) is made equal to zero, otherwise T(I,J) is set equal to one. The dimensions of the matrix T can not exceed 20 x 30 as explained before.

Once the grid is defined, rules for assigning hydrologic values characterizing the aquifer can be defined as follows:

1. The storage coefficient is a constant everywhere including areas outside the boundaries of the aquifer. The user has to give only the value of the constant of storativity and the program assigns this value to the different grid nodes.

2. Similarly, the initial conditions of the system are defined by a constant value throughout the aquifer. Note that equation (2.1) can be decomposed in two equations,namely, the dynamic and the steady state(or static) equations. The assumption that the initial head is a constant throughout the aquifer is used to by-pass the problem of solving the steady state equation of the system(in addition to

the dynamic). For aquifers where the initial water head can not be approximated by a constant, the present program is only applicable provided the steady state solution of the aquifer equation is found externally to the program and superimposed onto the dynamic response supplied by the program.

3. The aquifer simulator recognizes leakage from lakes or streambeds by assigning artificial pumping values,ranging from 2 to 10, to the node points where the leakage occurs. The values 2 to 10 are merely signals. Associated with these values corresponding hydraulic conductivities, HYDCON(K) must be defined. For example, a pumping value of 2 at node (I,J) indicates to the program that there is leakage at this node. To calculate the leakage value the program requires the value of HYDCON(K), where K = P(I,J) - 1 , which has to be defined by the user.

4. At node points where there is a constant head condition a signal value of 1 is assigned to the pumping rate. For example, assuming there is a constant head condition at node(I,J) , then a pumping rate P(I,J) = 1 is assigned at node (I,J).

5. Constant flow boundary points are defined as pumping (or recharge) wells with pumping (or recharge) rate values equal to the influx (recharge) or outflux (pumping) of water through the area represented by the node. The pumping values are given in cubic feet per second.

8. Since each grid node represents an area segment of the total aquifer area, a number of wells in the same area associated with

a given node, should be combined into a single well whose pumping rate is the sum of the individual pumping rates.

9. Location of pumping wells, boundary points, constant head points etc. , is always referenced to the grid superimposed on the aquifer.

10. The input-output historical record needed in the identification process is input to the program by giving the spatial coordinates and the time of the observation measure as well as the value in feet of the water head observed.

11. The initial guess for the transmissivity function is defined in the program by means of the program parameters BET(1,1), BET(2,1) and BET(3,1). Rules to select those values were already discussed in section A.2.8.

A.3.2 Structure

The program is composed of twenty three subroutines which perform the following tasks:

OPTIMIZATION SECTION

1. MAIN: Program control and optimization are performed. The flow diagram depicted in figure 4.1 sumarizes the tasks performed by MAIN.
2. INPT: Initial guess of transmissivity parameters, water head observations (i.e. input-output historical record) ,and parameter values needed in the optimization scheme are read in.
3. MAD: Aquifer simulator. Solution of the aquifer and sensitivity equations is implemented.

4. MATCH: Coordinates information exchange between MAIN and MAD. sets proper boundary and initial conditions to solve sensitivity equations. Calculates the forcing functions of these equations.

5. MATMPY: Matrix multiplication subroutine used in the calculation of A ($A = P^T P$ and P is the matrix of solutions of the sensitivity equations), and in the estimation of Lambda, the parameter which controls the size of the incremental corrections in the optimization process.

6. PHIC: Evaluates the error function PHI which is given by the sum of squares of the residuals (Observed water heads - Predicted water heads).

7. RSCAL: Scales the matrix of derivatives A (by transforming its diagonal elements into 1's), and the gradient of the error function PHI.

8. VECT: Calculates the magnitude and the direction of the correction to the initial transmissivity parameters.

9. CHECK1: Evaluates the Stoping criterion.

AQUIFER SIMULATOR SECTION

10. INF: Signals for suppresion or activation of subroutines related to the printing of drawdown distributions, well locations or transmissivity maps are read in. Indications of the type of boundary conditions, leakage, river and lake infiltrations are input. Grid dimensions, number of pumping periods are considered in the identifica-

tion process; the time span of a pumping period, initial conditions, storage coefficients, and coefficients specifying the leakage conditions(if any) are read in. Finally, the geometry of the aquifer is also input.

11. PARAM: Calculation of the iteration parameter $I^n_{i,j}$ for the implicit iterative method is performed.

12. INF2: Constants, whose initial values are the same regardless of what aquifer model is being identified, are initialized.

13. WELL: The value of pumping and recharge wells, and their location nodes are read in. Also nodes where constant head and/ or constant flow boundary conditions exist as well as leakage from rivers, lakes or confined layers are read in along with the corresponding flow or infiltration values.

This information is stored in the matrix PS(I,J,K) and called in whenever the aquifer simulator is used in the solution of the aquifer equation. Note that for the sensitivity equations PS changes from one iteration of the optimization process to the next one.

14. ITRATE: Within a pumping period the alternate implicit direction method is cycled until either convergence is achieved or the maximun number of iterations allowed within a pumping period is exceeded. The cycling within iterations is achieved by using the iterative parameter I^n_{ij} . Detection of oscillatory conditions in the solution of the aquifer equation produced by possible negative transmissivity values (the optimization procedure may drive

the transmissivity function negative), is performed.

15. CYCLE: Calls a new well field at the end of each pumping period.

16. LEAKY: Determines the pumping rate of wells, at a given node, in cubic feet per second. Also, identifies whether a node is a constant boundary point or a point of leakage, by checking the value of the pumping rate P(I,J). If P(I,J) is greater than 10 or less than zero the node is a pumping or recharging well. If P(I,J) is 1, a constant boundary point is located at node (I,J), and a vertical permeability of 10^{50} is assigned along with a pumping rate of zero ft^3/sec. If P(I,J) is between 2 and 10, leakage from a stream or lake bed is present and a hydraulic conductivity value of HYDCON (P(I,J)) is is assigned to the node.

17. LEAKY1: Calculates the discharge of water in ft^3/sec. which leaks from a confining layer with storage.

18. WCHECK: Prints the water head at a desired node and at a particular time. The location of these nodes is read in.

19. MATCAL: The alternate implicit direction method is implemented in this subroutine which is divided in three sections: INITL which initialize the variables used in the alternate implicit direction technique. ROW which solves the equation in the X direction, i.e., calculates a solution of the aquifer equation for a particualr time, keeping column values fixed (see equation 7). COLUMN which solves equations in the Y direction,i.e., calculates a solution of the aquifer equation keeping row values fixed (see equation 8).

20. PRINT: It prints (when desired) the transmissivity initial distribution, well distribution and drawdown maps.

21. INTERV: The time interval within a pumping period can be broken down into subintervals of different time lengths The purpose of so doing is to obtain a more accurate solution of the aquifer equation. As the length of these subintervals increases, more iterations will be needed to achieve convergence within a pumping period. On the other hand, as the length of the subintervals decreases, fewer iterations of the alternate implicit method may be needed to achieve convergence within a subinterval but more subintervals will be necessary to reach the end of a pumping period. INTERV gives the option to the user of adjusting the length of the subintervals by increasing their length geometrically according to the expression

$$\Delta t_K = \Delta t_{K-1} [\ \text{ALPHA}^{\,k} - 1\ ] / [\ \text{ALPHA} - 1\ ]$$

This expression is based on the fact that the response of an aquifer behaves approximately as an error function. The user should select a value for ALPHA ranging from 1.1 to 10.0. There is no rule for the selection of ALPHA except experience with the program and the problem under study.

22. TIME: Calculates time passage in seconds, minutes, hours, days, and years.

23. WRITER:Coordinates the print out of head distribution and maps.

A.3.3 Flow Chart

The program flow chart is presented in figures A.1 and A.2 .

A.4 Input Requirements

Input to the program consists mainly of data cards related to the aquifer simulator and to the optimization process. The sequence of data input as well as a detailed description of the variables involved follows.

A.4.1 Data Cards

First, data cards related to the optimization section, MAIN, are read in using the subroutine INPT. The data deck for this section is described below. The program variables are :

NIT = Maximun number of iterations allowed to the optimization routine MAIN.

NG = Number of water head observations used in the identification.

KB = Number of parameters to be identified.

BET(1,1) = Initial guess trasmissivity coefficient b_1. Under scarce information conditions set BET(1,1)= 10^{-6}.

BET(2,1) = Initial guess transmissivity coefficient b_2. Under scarce information conditions set BET(2,1)= 10^{-6}.

BET(3,1) = Initial guess transmissivity coeffcient b_3. Under scarce information conditions choose BET(3,1) between [0.2,0.8].

ELMDA(1) = Initial value of λ.

NU = Number of pumping periods considered. NU defines the time context

of the identification process.

JSCALE = Scales water head observations to integer values in feet.

IDIM = Total number of rows making up the grid.

JDIM = Total number of columns making up the grid.

FHEAD = Observation data format.

I,J,K = Space coordinates (I,J) and time (K) of the observation measure.

PS(I,J,K) = Matrix of observation points.

IFLAG = A signal flag. If IFLAG=0 another observation point is read. If IFLAG = 1 no more observation cards are read.

EPS1 = If the stoping criterion at a given iteration becomes less than EPS1 , the program stops.

The above data is input according to the following card sequence:

CARD	VARIABLE	FORMAT
1	NIT, NG, KB	3I4
2	BET(1.1),BET(2,1),BET(3,1)	3F13.10
3	ELMDA(1),NU,JSCALE,IDIM,JDIM	F10.6,4I10
4	FHEAD	3A4
5	I (row), J (column), K (time), PS(I,J,K) (observation), IFLAG (signal)	20I4
6	EPS1	F10.6

The card number above indicates the order in which data cards are being read. For example, the observation data are read in the fifth

place.

Now data cards related to the aquifer simulation section of the program are read in. Most of the variables describing the aquifer are input through subroutine INF. Program variables are:

PCHAR(I) = A chart containing 160 symbols to be used in the printing of transmissivity,drawdown, and well location maps. A list of this symbols can be seen in Section A.5 .

IPT = If a map of the initial aquifer transmissivity distribution is desired, set IPT = 1 , otherwise IPT = 0

IPW = If a map showing the aquifer well location and constant head boundary location is desired, set IPW = 1 , otherwise IPW = 0

IPH = If a map showing water head distributions throughout the aquifer is desired set IPH = 1, otherwise IPH = 0. This feature is not applicable for the case where the aquifer simulator is used for the solution of the sensitivity equations.

IBOUND = If IBOUND= 1 , then there exists somewhere in the aquifer a constant head boundary. This IBOUND is a signal which indicates that constant head points will be input to the program. If there is no constant head condition set IBOUND = 0

ILEAK1 = Leakage signal: if ILEAK1 = 1 there is somewhere in the aquifer leakage from a confining region with storage capabilities.

IDRAW = Set IDRAW = 0 . In future extensions of the program, IDRAW= 1

will activate a drawdown printout option.

ICHECK = Set ICHECK = 0, in future extensions of the program ICHECK =1 will activate a mass balance option to check accuracy of the solutions produced by the alternate implicit direction technique.

IRIV = If IRIV = 1 a river or lake is present in the aquifer and leakage occurs, otherwise IRIV =0

IFLOW = If IFLOW = 1, a constant flow boundary condition is present, otherwise IFLOW = 0

FTRAN = Format to read the geometric pattern of the aquifer

FPUMP = Pumping rate and constant head boundary format

KTH = The frequency of printing maps, related to the aquifer simulation, is controlled by KTH

ITER = Number of time steps within a pumping period

NUMKT = Total number of time steps considered in the identification. The number of pumping periods,NU, equal NUMKT/ITER.

INFT = Set INFT = 25 . This variable is related to the number of terms considered in the formula for infiltration from a confining layer.

LENG = Maximum number of iterations in the water head solution, within a pumping period , can be cycled. LENG should be assigned a value of 60 in this program since the aquifer simulator is going to be used to solve different sensitivity equations (at each iteration of the optimization procedure a new set of equations is produced) . The value LENG = 60 is found to cover

most situations.

STRTH = Initial water head conditions in the aquifer.

DX = The distance in feet between grid nodes.

ALPHA = If the pumping periods are divided into subintervals, ALPHA controls, through a geometric increase, the size of these subintervals. For example, if ALPHA = 2.0, then the time steps within a pumping period are doubled after completion of a row (solution of equation 7) and column (solution of equation 8) calculations. ALPHA must always be greater than one.

S = Storage coefficient of the aquifer

RANGE = Time in seconds of the duration of a pumping period

BEDTHK = If there is a stream or lake bed through which water leaks, BEDTHK is the thickness of the bed. Set BEDTHK = 1.0 when there is no leakage.

HYDCON(K) = Hydraulic conductivity for the stream or lake bed. The aquifer model allows definition of nine different values of hydraulic conductivities. The user has to combine or coordinate the index of HYDCON with the value of P(I,J) as explained before (see section A.3.1 point 4). For example, if at node (I,J) a stream or lake passes by, the user must indicate this to the program by assigning a value to P(I,J) between 2 and 10 (integer values only). Associated with each P(I,J) a value HYDCON(P(I,J) - 1) is defined. Nine values usually give enough flexibility to the user for modeling

river or lake interactions with the aquifer.

SCALE 1 = Scales decimal pumping values into integer values in cubic feet per second.

SCALE 2 = Set Scale 2 = 1.0. In the optimization procedure this scale will be redefined. This scale is related to transmissivity

DT = The value of the time increment in seconds. If the time geometric increase feature is to be used, then set DT = 1.0 Otherwise, the value of DT equals the size of the pumping period or of one of the subintervals.Usually one year is considered a normal value for DT (or 31,557,600 seconds).

EPS = If within a given time step (a pumping period is made of time steps) two consecutive iterations of the alternate implicit direction method produce head values with a difference less than EPS (EPS = 10^{-3} is a good value), the program proceeds to the next time step.

ZPERM = If there is leakage from storage of a confining layer, then ZPERM is the permeability of that layer. Otherwise set ZPERM = 1.0

ZLENG = Thickness of the confining layer. If no leakage is present, set ZLENG = 1.0

The above data is input according to the following card sequence:

CARD	VARIABLE	FORMAT
7,8,9,10	PCHAR(I)	40A2
11	IPT, IPW, IPH, IBOUND, ILEAK1, IDRAW, ICHECK, IRIV, IFLOW	10I2
12	FTRAN, FPUMP	2(3A4)
13	KTH, NUMKT, INFT, LENG, ITER	10I8
14	STRTH, DX, ALPHA, S	4F20.5
15	RANGE, BEDTHK, HYDCON(1), HYDCON(2)	4F20.5
16	HYDCON(3), HYDCON(4), HYDCON(5),HYDCON(6)	4F20.5
17	HYDCON(7), HYDCON(8), HYDCON(9), SCALE 1	4F20.5
18	SCALE 2, DT, EPS	4F20.5
19	ZPERM, ZLENG, SC	4F20.5
20	T(I,J) (Aquifer Geometry)	20I4

Aquifer Geometry.- The geometry of the aquifer is defined by the grid superimposed on the aquifer area. The grid is made up of rows and columns. The geometry of the aquifer is input to the program by reading a one grid's row at a time. At each node of a given row an integer value, either 1 or zero, is assigned to the dummy varible T(I,J). If the node is inside or on the boundary set T(I,J) = 1, otherwise set T(I,J) = 0. The first card of a row (a grid's row may need more than 1 card to be defined) is occupied by the row number. Cards will continue to be read untill all the nodes of the row are read in. Then, a new row along with its row number

is read. The process continues until all the rows have been entered. The last card of the set is a check card. In columns 2 through 26 the check words "AQUIFER GEOMETRY READ IN," should be punched. This completes the description of the information being input in subroutine INF. It remains to describe how the boundary conditions as well as the river or lake interactions with the aquifer are defined to the program. This is the task of subroutine WELL. The information is input as follows:

Constant Head Boundary Cards.- Node points which lie on the constant head boundary (if any) are read one point per card. Each of the cards provides the row (variable I) and the column (variable J) number of the node, and a signal flag (JSIG) which determines if another constant boundary card is to be read. If JSIG = 0 , the next constant head card is read in, if JSIG = 1 the reading process is ended. The format for the constant head nodes is supplied by FPUMP.

Constant Flow Boundary Cards.- Node points which lie on the constant flow boundary (if any) are read one node point per card. Each of the cards provides the row (I) and the column (J) number of the node along with the representative pumping value P(I,J) for the node to simulate inflow (pumping rate negative) or outflow(pumping rate positive). The value of P(I,J) must be an integer scaled in ft^3/sec. A signal flag,JSIG , determines if another constant flow point is to be read. If JSIG = 1, the reading process ends. If JSIG = 0, the next

constant flow card is read in.

River, Lake Cards.- If a lake or stream has leakage into the aquifer from its bed, then those node points on the lake or stream are given integer P(I,J) values greater than 1 but less than or equal to 10. The P(I,J) value is not an actual pumping rate but a signal, and also an index for the hydraulic conductivity bed. Each of the river or lake cards provides the row and column number of the node (I,J) , the signal P(I,J), and the next card signal JSIG.

Pumping (Recharge) Well Card.- For each well in the aquifer a well card is required. The card provides the well's nodal position (I,J), and its rate of pumping P(I,J) which has to be an integer number in ft^3/sec(use SCALE 1 for this purpose) .The well card also carries a signal flag, JSIG , which works as explained in previous paragraphs. For each pumping period exists an associate set of pumping data cards. The last card of each set must have the signal JSIG = 1 which marks the end of a pumping period data set.

The card sequence is as follows:

CARD	VARIABLE	FORMAT
21	I , J , JSIG (row) (column) (signal)	3I5
22	I, J, P(I,J) , JSIG (flow or index)	4I5
23	I, J, P(I,J) , JSIG (pumping-recharge)	4I5

The card number above indicates the order in which the boundary cards should be arranged to be read in. Card 21 refers to the set of constant head cards and the number 21 indicates the position of the set in the sequence of data cards of the program. Card 22 refers to the set of constant flow and/or river-lake cards while card 23 refers to the set of pumping or recharge rates.

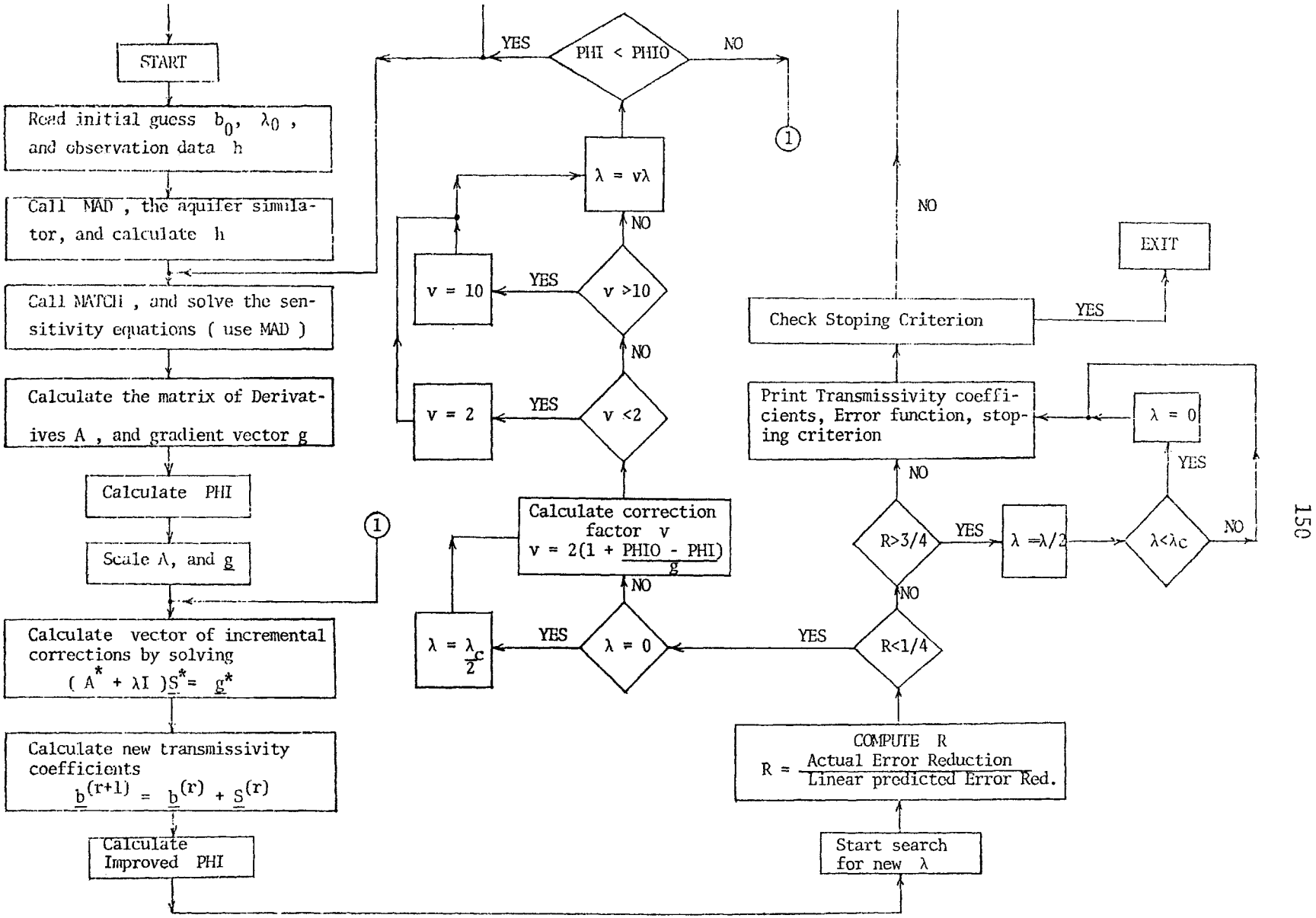
A.5 Output

The output of the program is mainly composed of error messages, input data, maps, and results of the identification process.

1. ERROR MESSAGES

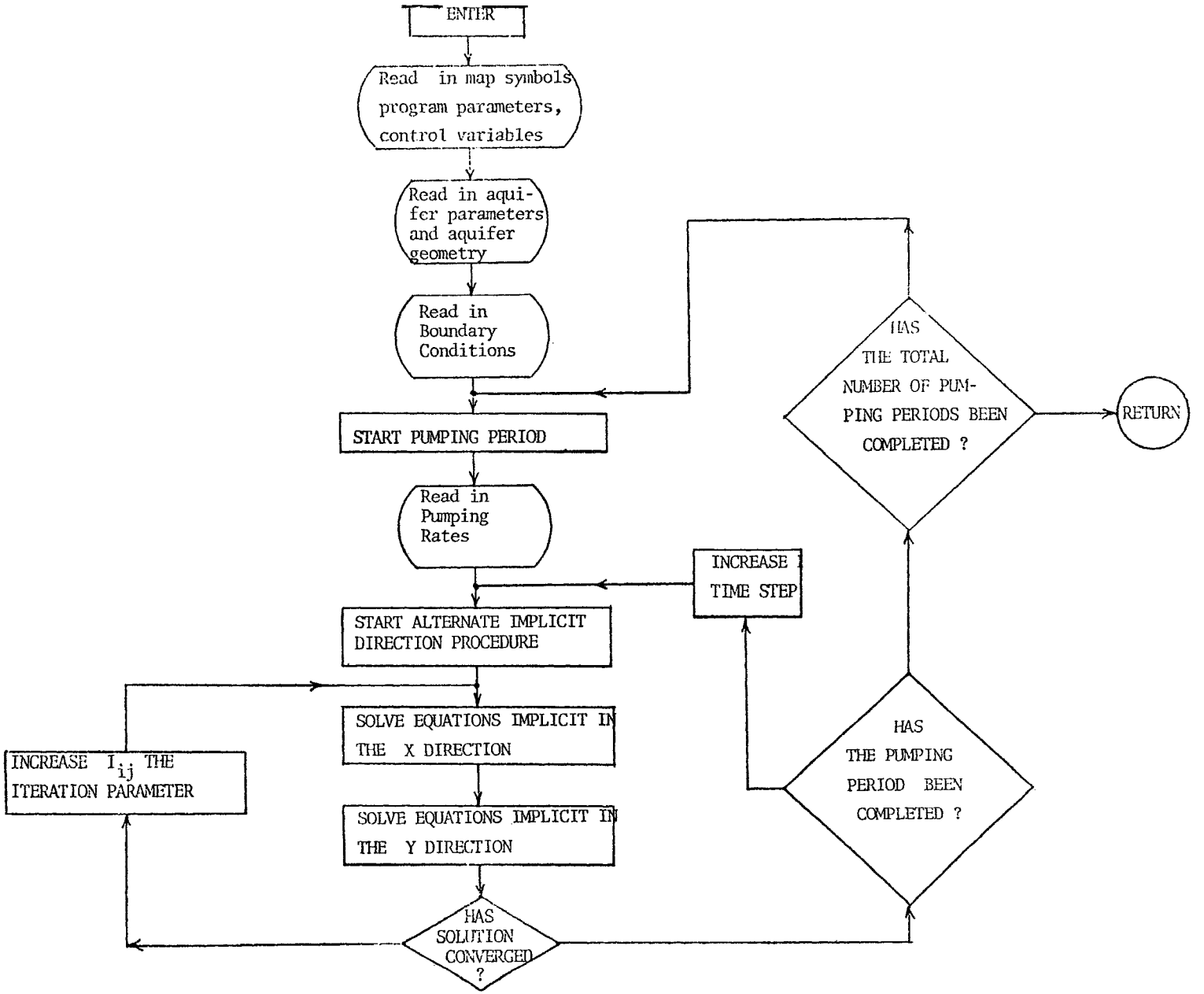
The error message - THE SOLUTION FAILED TO CONVERGE - is printed to indicate that either the solution of the aquifer equation or the sensitivity equations was not attained. Some possible causes for this are the use of negative transmissivity values in the equations above, the use of large time increments in discretizing these equations, coarse grids, or maldefinition of the aquifer geometry and/or boundary conditions. The user can identify the problem area by means of the additional error message - TROUBLE IN MATCH - which is printed whenever the convergence failure is related to the sensitivity equations. The nature of the problem is identified by the error message - OSCILLATIONS - which is printed when a negative transmissivity value drives unstable the alternate implicit direction method.

Another error message printed by the program is - OVERFLOW DETECTED IN MATRIX INVERSION - which indicates that ill-conditioning of the matrix A is causing trouble in the inversion procedure. The problem may be



FLOW DIAGRAM OPTIMIZATION SECTION

FIGURE A.1



FLOW DIAGRAM AQUIFER SIMULATOP SECTION

FIGURE A.2

caused by round off error in the calculation of the sensitivity equations compounded with ill conditioned characteristics of the minimization surface PHI.

2. MAPS

There are only two maps which are printed:

a- INITIAL TRANSMISSIVITY MAP which depicts the geometry of the aquifer as well as the initially guessed transmissivity values (in hundredths of feet square per second).

b- WELL LOCATION and BOUNDARY CONDITIONS MAP which shows those nodes where there are wells discharging water from the aquifer (^{-}W) or wells recharging the aquifer (^{+}W), constant head conditions (CH) , river or stream conditions (R), and constant flow conditions which are indicated in the map by a ^{-}W (outflow) or ^{+}W (inflow) symbol.

3. INPUT DATA

The following input data is printed for checking purposes; observation data, maximum number of iterations allowed to the optimization process, number of grid points, number of unknown parameters, initial guess of transmissivity coefficients, initial value of λ, and information related to the aquifer simulator and read in INF (see section A.4.1, p. Pumping rates and infiltration rates are not printed, however the program can be easily modified to do so.

4. IDENTIFICATION RESULTS

At each iteration of the optimization procedure the following information is printed; the value of the coefficients being identified, the error function

PHI (observation head - predicted head) , the iteration number, the stopping criterion CHEK (for each transmissivity coeffcient), the incremental correction vector $\underline{S}$, and the new value of λ . For cases where the λ procedure requires more than one iteration , the transmissivity coefficients are printed for each iteration. The last printing before the program stops contains the optimal solution.

A.6 Concluding Remarks

Concerning point 2 section A.3.1 , the steady state solution can be obtained by using the aquifer model independently of the optimization section. Then, solution of the aquifer equation considering a long time horizon (say 100 years) will give the steady state response. To disconnect the aquifer simulator from the optimization section introduce in MAIN the statement GO TO 250 after 12 CONTINUE . The same data deck used in the identification process should be used for the above implementation.

The convergence of the identification algorithm depends on the initial guess of the transmissivity parameters. Should the optimization process fail to produce a solution, the user will have to supply a new starting point. The information generated in unsuccessful runs can be used to make better choices of the initial guess. For example, a severe correction on one (or several) parameter (s) indicates overestimation(or underestimation) of the parameter value.

Concerning core requirements the program as now written, requires about 47 K words on the UNIVAC 1108 computer, a word being equivalent

to 36 bits. As for computer time , with 20 x 30 grid nodes (see Chapter 4 , the Fairfield-New Baltimore aquifer) and a period of five years, and with yearly changes in pumping rates, the program takes 120 seconds .This time involves 7 iterations of the optimization procedure , each iteration requiring the calculation of the solution of 4 parabolic partial differential equations.