#2
4/24/80

UCID-18575

**PARAMETER IDENTIFICATION IN GROUNDWATER FLOW:
A FEASIBILITY STUDY**

**S. G. AZEVEDO
T. A. DOERR
J. V. CANDY
K. D. PINENTEL**

MASTER

MARCH 11, 1980

**Lawrence
Livermore
Laboratory**

This is an informal report intended primarily for internal or limited external distribution. The opinions and conclusions stated are those of the authors and may or may not be those of the Laboratory.

This work was supported by the United States Nuclear Regulatory Commission under a Memorandum of Understanding with the United States Department of Energy.

ABSTRACT

A set of first-order differential state equations is developed from taking finite differences of the partial differential equation that describes groundwater flow during a pump-drawdown test. The number of states in the model is a direct function of the number of discrete spatial segments used in the approximate model. This number of states is studied to determine a reasonably low-order model that preserves the accuracy of resulting simulations when compared with an available analytical solution. The resulting reduced-order model is then used in an extended Kalman filter for combined state and parameter estimation. This report summarizes work done in a recent pilot study to demonstrate the feasibility of these techniques in groundwater research and points the way toward other approaches to this problem area now under study.

DISCLAIMER

This book was prepared as an account of work sponsored by an agency of the United States Government, neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

TABLE OF CONTENTS

	<u>Page</u>
Abstract	i
List of Figures	iii
List of Tables	iv
I. Introduction	1
II. Groundwater Flow Model Development	3
A. General Model for the Simple System	4
B. Finite Differencing	7
C. Boundary Conditions	11
D. Node Spacing and Quantity	16
III. State Estimator Design	31
IV. Parameter Estimation/Identification	38
V. Conclusions	46
Acknowledgments	47
References	48
Appendix A.	49

LIST OF FIGURES

<u>Figure</u>	<u>Title</u>	<u>Page</u>
1.	Aquifer Pump-Drawdown Test	5
2.	Spatial Discretization (away from boundaries)	9
3.	Uniform Spatial Sampling	12
4.	Drawdown and Model Error Curves for 95 State Model	19
5.	Drawdown and Model Error Curves for 95 State Model	21
6.	New Spatial Sampling Over Total Interval	23
7.	Drawdown and Model Error Curves for 28 State Model	25
8.	Drawdown and Model Error Curves for 28 State Model	26
9.	Drawdown and Model Error Curves for 21 State Model	27
10.	Drawdown and Model Error Curves for 21 State Model	28
11.	Simulated Noisy Measurements	33
12.	State Estimation of Nodes 1 and 10	35
13.	State Estimation (Node 21) and Innovations	36
14.	Drawdown Estimates Reconstructed from Noisy Observation Well Measurements	41
15.	Transmissivity and Storage Coefficient Estimates	42
16.	Parameter Estimate Contour of Constant Probability ($p \approx .99$)	45

LIST OF TABLES

<u>Table</u>	<u>Title</u>	<u>Page</u>
1.	Parameter List	18
2.	Summarized Simulation Results	30
3.	Summary of Continuous-Discrete Extended Kalman Filter	34
4.	Parameter Sample Statistics for an Example of Ten Runs	44

I. INTRODUCTION

The purpose of this report is to detail the development of groundwater flow models and their use in demonstrating the feasibility of applying estimator/identifiers to problems in subsurface hydrology. This feasibility study was recently carried out in the Engineering Research Division of the Electronics Engineering Department on the request and under the support of the NRC-funded Waste Management Program (WMP) at LLL. It was to have been the precursor to a broader study aimed at applying related results and methods from the systems sciences to hydrogeological investigations in the WMP. It has led to studies of measurement error characterization and analysis for the WMP and to more basic research in filtering and parameter identification for multi-dimensional distributed systems.

This study is aimed at applying the extended Kalman filter to a synthesized groundwater aquifer test in order to identify physical aquifer parameters. Suppose a pump-drawdown test is set up with a constant-displacement pump discharging water from a single well that completely penetrates a confined homogeneous isotropic aquifer of constant thickness. Suppose further that a single observation well is drilled some distance away from the discharging well and that all that is available from the experiment is a rather badly noise-corrupted set of measurements of the piezometric surface at the observation well. We desire to process these noisy measurements to determine the transmissivity and storativity for the aquifer. The study involves starting with the partial differential equation model for groundwater flow, transforming it to a system of first-order differential equations by finite differencing, determining a differencing scheme that will yield a model which strikes a balance between computational speed and

accuracy, and applying the reduced-order model in the estimation and parameter-identification problem.

Lawrence Livermore Laboratory has been involved for some time in research sponsored by the Nuclear Regulatory Commission (NRC). The WMP sponsored by the NRC has various goals that support the work of the Commission in licensing and regulating potential nuclear waste repositories. The WMP at Livermore has at its center the study of uncertainty and risk in the design and development of candidate nuclear waste repositories. The study described in this report was done to establish the feasibility of applying modern stochastic systems methods to hydrogeological investigations for the WMP. The original areas of interest embraced a wide variety of possible topics from modern systems analysis; these included developing and implementing formal model-order reduction methods, applying and extending model identification and realization techniques, analyzing state-of-the-art geophysical measurement techniques, suggesting and developing new techniques based on stochastic systems methods, and participating in work aimed at quantifying uncertainties inherent in expert opinion in the field of geophysics and hydrology. The case study chosen for this pilot study was by design rather simple. The problem was chosen because in its original form it is a problem any practicing hydrologist can solve with very little effort. It was a problem for which analytical solutions and tables and charts already exist. But this was the desired experiment; to approach a problem well known in the field, but from a different perspective to see if that new perspective had merit. Although the results we achieved do not shake the earth, they do indicate that there are areas that should be explored in the future to further develop the recent marriage between more traditional hydrology and the systems sciences.

II. GROUNDWATER FLOW MODEL DEVELOPMENT

The initial problem to be encountered in this study is the development of a mathematical model which describes water flow in an aquifer (water-bearing geologic zone) during pumping. The model must contain enough detail to adequately characterize the physical hydrogeologic system, while being simple enough for simulation by computer. Within this section, we explore the model development, test several simulators, and present the chosen model used in subsequent parameter estimation.

The particular estimator we plan to utilize requires that the system model be in the state-space (vector) representation. Therefore, in our development of the model, we show how the water flow equation can be decomposed into a number of first-order differential equations using a method of finite differences.

For the purpose of this report, we are interested in a simplified mathematical model to describe the dynamics of groundwater flow in a hypothetical aquifer. Future studies may take into consideration some additional complexities which lead to more accurate modeling of a particular site (i.e., altered boundary conditions, time- and space-varying parameters, etc.). However, the simplified model described here is sufficient to illustrate the concept of parameter estimation.

The topics discussed in this section can be grouped under the following headings:

- General Model for the Simple System
- Finite Difference Method
- Boundary and Initial Conditions
- Node Spacing and Quantity

A. General Model for the Simple System

The model chosen to describe the aquifer system is one which contains several assumptions that make it an ideal case. We are only attempting to explore the concepts of identifying aquifer parameters in a simple test system. Extensions of the model to account for more complicated aquifer structures may be incorporated eventually.

A standard technique currently used to determine the hydrogeologic properties of the aquifer from dynamic measurements is called the pump-drawdown test.¹ In this test, water is pumped out of a well (called the pump or discharge well) and the drawdown, or change in hydraulic piezometric head (water table) is measured at other wells (called measurement or observation wells) located at various distances away (see Figure 1). Our mathematical description of this test includes aquifer parameters whose values are subsequently extrapolated from the measurement data.

We use the following assumptions for developing the simplified aquifer model:

- (i) The groundwater is contained entirely in a homogeneous and isotropic medium of uniform thickness that does not vary with time (i.e., the system parameters are space and time invariant).

Figure 1. Aquifer Pump-Drawdown Test

- (ii) The aquifer is confined; i.e., it is separated from the atmosphere by impermeable material.
- (iii) The aquifer has infinite areal extent.
- (iv) The discharge well penetrates the entire thickness of the aquifer, is fully perforated, and has infinitesimal diameter.
- (v) The pump rate is constant and the flow is radial to the discharge well.

The basic partial differential equation that governs the flow in such an aquifer is the following:^{2,3}

$$\frac{\partial^2 h(r,t)}{\partial r^2} + \frac{1}{r} \frac{\partial h(r,t)}{\partial r} = \frac{S}{T} \frac{\partial h(r,t)}{\partial t} \quad (2.1)$$

where

- $h(\cdot)$ = drawdown from original water level at $t \leq 0$, (m).
- r = radial distance from pump well (m)
- t = time (sec)
- S = coefficient of storage (m/Pa)
- T = transmissivity ($m^3/Pa\text{-sec}$)

with the following initial condition for all $r \geq 0$:

$$h(r,t) = 0 \text{ for } t \leq 0, \quad (2.2)$$

which implies uniform hydraulic head before pumping. The boundary conditions for all $t \geq 0$ are

$$\lim_{r \rightarrow \infty} h(r, t) = 0 \quad (2.3)$$

$$\lim_{r \rightarrow 0} \left[r \frac{\partial h(r, t)}{\partial r} \right] = \frac{q}{2\pi l} \quad (2.4)$$

for a constant discharge rate of q (m^3/sec).

An analytic solution to this problem is given by^{1,2}

$$h(r, t) = \frac{-q}{4\pi l} \left[-E_i \left(-\frac{r^2 S}{4Tt} \right) \right] \quad (2.5)$$

where $-E_i(-u) = \int_u^\infty \frac{e^{-x}}{x} dx$ = the exponential integral.

We will use (2.5) as a verification for the accuracy of the state-space model derived by finite differencing in the next sub-section.

B. Finite Differencing

As an alternative to the analytic solution presented in the last section, we now present a state-space approach for solving the hydraulic equation of (2.1). This is accomplished by spatially discretizing the aquifer region by using first-order finite differences to approximate the partial derivatives. Here we explore how the finite difference method leads to a state-space model. The next section will incorporate boundary conditions to complete the system description.

First, let us spatially discretize a horizontal line extending away from the pump into equally spaced intervals, Δr . Then let r_j (for $j = 1, 2, 3, \dots$) represent the radial distance of each node from the discharge well (see Figure 2).

Equation (2.1) contains partial derivatives with respect to r that can be approximated by first-order differences. For example, let i be some integer greater than 2 and let

$$h'_{i+1/2} = \left. \frac{\partial h(r,t)}{\partial r} \right|_{r=r_{i+1/2}} \approx \frac{h(r_{i+1},t) - h(r_i,t)}{r_{i+1} - r_i} \quad (2.6)$$

This is called a first forward difference and is evaluated for the midpoint of the (r_i, r_{i+1}) interval or at

$$r_{i+1/2} = \frac{r_i + r_{i+1}}{2} \quad (2.7)$$

Then the first backward difference is

$$h'_{i-1/2} = \left. \frac{\partial h(r,t)}{\partial r} \right|_{r=r_{i-1/2}} \approx \frac{h(r_i,t) - h(r_{i-1},t)}{r_i - r_{i-1}} \quad (2.8)$$

for the mid-point
$$r_{i-1/2} = \frac{r_{i-1} + r_i}{2} \quad (2.9)$$

Now we can rewrite (2.1) as

$$\frac{\partial h(r,t)}{\partial t} = \frac{T}{S} \cdot \frac{1}{r} \frac{\partial}{\partial r} \left[r \frac{\partial h(r,t)}{\partial r} \right] \quad (2.10)$$

Figure 2. Spatial Discretization (away from boundaries)

and approximate it at the i th node by taking the second difference

$$\left. \frac{\partial h(r,t)}{\partial t} \right|_{r=r_i} \approx \frac{T}{S} \cdot \frac{1}{r_i} \left\{ \frac{r_{i+1/2} h'_{i+1/2} - r_{i-1/2} h'_{i-1/2}}{r_{i+1/2} - r_{i-1/2}} \right\}.$$

After evaluating, this equation becomes

$$\dot{h}_i \approx \frac{T}{S} \cdot \frac{1}{r_i(r_{i+1} - r_{i-1})} \left[(r_{i+1} + r_i) \frac{(h_{i+1} - h_i)}{(r_{i+1} - r_i)} - (r_i + r_{i-1}) \frac{(h_i - h_{i-1})}{(r_i - r_{i-1})} \right]$$

where $h_i \equiv h(r_i, t)$

$$\text{or} \quad \dot{h}_i = \frac{T}{S} \left[C_{i1} h_{i-1} + C_{i2} h_i + C_{i3} h_{i+1} \right] \quad (2.11)$$

$$\text{where} \quad C_{i1} = \frac{r_{i-1} + r_i}{r_i(r_{i+1} - r_{i-1})(r_i - r_{i-1})}$$

$$C_{i2} = \frac{-2}{(r_{i+1} - r_i)(r_i - r_{i-1})}$$

$$C_{i3} = \frac{r_i + r_{i+1}}{r_i(r_{i+1} - r_{i-1})(r_{i+1} - r_i)}$$

and $i = 1, 2, \dots$

Equation (2.1) has now been transformed to a set of state-space differential equations where the "states" are the drawdown quantities at each node location. All that remains is to implement the boundary conditions and fix the node spacing. These are addressed in the next section.

C. Boundary Conditions

The effects of the physical aquifer boundaries on (2.10) must be incorporated into the model for completeness. To do this, a few more restrictions on the method of spatial sampling must also be added. These are described in this section.

First of all, assume that we have discretized the spatial dimension into equally-spaced intervals. Then the boundary condition at the origin (i.e., at the pump) may be very simply implemented. Because the form of this restriction (2.4) is the same as that in brackets of (2.10), it can be inserted directly into the difference equation for the first node (r_1). Let us assume that the distance between the pump well and node 1 is one-half the standard node spacing (see Figure 3). Then the first forward difference is found as before, but the first backward difference is given by (2.4); i.e.,

$$r_{1-1/2} h'_{1-1/2} = \frac{q}{2\pi T} \quad (2.12)$$

This gives us

$$\left. \frac{\partial h(r,t)}{\partial t} \right|_{r=r_1} \approx \frac{T}{S} \cdot \frac{1}{r_1} \left[\frac{r_{1+1/2} h'_{1+1/2} - \frac{q}{2\pi T}}{r_{1+1/2} - 0} \right]$$

so

$$\dot{h}_1 = \frac{T}{S} \cdot \frac{1}{r_1(r_2 + r_1)} \left[(r_1 + r_2) \left(\frac{h_2 - h_1}{r_2 - r_1} \right) - \frac{q}{\pi T} \right]$$

$$\boxed{\dot{h}_1 = \frac{C_{11} q u(t)}{S} + \frac{T}{S} [C_{12} h_1 + C_{13} h_2]} \quad (2.13)$$

X - Node Locations

• - First Difference Evaluation Points

Figure 3. Uniform Spatial Sampling

where
$$C_{11} = - \frac{1}{\pi r_1 (r_2 + r_1)}$$

$$C_{12} = \frac{-1}{r_1 (r_2 - r_1)}$$

$$C_{13} = \frac{1}{r_1 (r_2 - r_1)}$$

and $u(t)$ is the unit step function in time.

Note that, as we would expect, the C_{11} coefficient provides the source term necessary to drive the system.

The other boundary, the one at $r \rightarrow \infty$ as described by (2.3), is approached in a similar way. It can be shown (Appendix A) that a sufficient condition for (2.3) to hold is:

$$\lim_{r \rightarrow \infty} \left[r \frac{\partial h(r,t)}{\partial r} \right] = 0 \quad (2.14)$$

If we truncate the discretized region at $r = r_N < \infty$ such that $|h_N| \ll |h_w|$ (where h_w is the drawdown at the measurement well, r_w) for all $t > 0$, then we can implement the above boundary condition without significantly adding error. Now we have

$$\left. \frac{\partial h(r,t)}{\partial r} \right|_{r=r_N} \approx \frac{T}{S} \cdot \frac{1}{r_N} \left\{ \frac{0 - r_{N-1/2} h'_{N-1/2}}{r_{N+1/2} - r_{N-1/2}} \right\}$$

so

$$\dot{h}_N = \frac{T}{S} \cdot \frac{-1}{r_N(r_{N+1} - r_{N-1})} \left[(r_N + r_{N-1}) \frac{(h_N - h_{N-1})}{(r_N - r_{N-1})} \right]$$

$$\boxed{\dot{h}_N = \frac{T}{S} [C_{N1} h_{N-1} + C_{N2} h_N]} \quad (2.15)$$

where

$$C_{N1} = \frac{r_N + r_{N-1}}{r_N(r_{N+1} - r_{N-1}) (r_N - r_{N-1})}$$

$$C_{N2} = \frac{-(r_N + r_{N-1})}{r_N(r_{N+1} - r_{N-1}) (r_N - r_{N-1})}$$

and r_{N+1} is a hypothetical node beyond the N^{th} node. Figure 3 summarizes this discretization scheme.

Now we have the entire discretized state representation of the groundwater flow equation:

$$\dot{h}_i = \begin{cases} \frac{T}{S} [C_{i2} h_i + C_{i3} h_{i+1}] + \left(\frac{1}{S}\right) C_{i1} q u(t), & i = 1 \\ \frac{T}{S} [C_{i1} h_{i-1} + C_{i2} h_i + C_{i3} h_{i+1}], & 1 < i < N \\ \frac{T}{S} [C_{i1} h_{i-1} + C_{i2} h_i], & i = N \end{cases}$$

In matrix form this becomes:

$$\begin{bmatrix} \dot{h}_1 \\ \dot{h}_2 \\ \vdots \\ \dot{h}_i \\ \vdots \\ \dot{h}_{N-1} \\ \dot{h}_N \end{bmatrix} = \left(\frac{T}{S} \right) \underbrace{\begin{bmatrix} C_{12} & C_{13} & & & & \\ C_{21} & C_{22} & C_{23} & & & \\ & \ddots & \ddots & \ddots & & \\ & & C_{i1} & C_{i2} & C_{i3} & \\ & & & \ddots & \ddots & \ddots \\ & & & & C_{(N-1)1} & C_{(N-1)2} & C_{(N-1)3} \\ & & & & & C_{N1} & C_{N2} \end{bmatrix}}_F \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_i \\ \vdots \\ h_{N-1} \\ h_N \end{bmatrix} + \left(\frac{1}{S} \right) \begin{bmatrix} C_{11}q \\ 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} u(t)$$

$$\dot{\hat{x}}(t) = \left(\frac{T}{S} \right) F x(t) + \left(\frac{1}{S} \right) G u(t)$$

or

$$\dot{\hat{x}}(t) = \left(\frac{T}{S} \right) F x(t) + \left(\frac{1}{S} \right) G u(t) \quad (2.16)$$

where

$$x(t) = \begin{bmatrix} h_1 & h_2 & \dots & h_N \end{bmatrix}^T.$$

With this dynamic state representation of the system, we can perform simulations of the aquifer drawdown test. The outcomes of these simulations can then be compared to the analytic solution (2.5) as a measure of modeling accuracy. In the next section, we utilize this comparison to determine the value of N (number of states) needed for our system model. In the process of finding N , we define an alternative method of spatial discretization.

D. Node Spacing and Quantity

We have not yet addressed the problem of determining some reasonable values for N , the total number of nodes, and r_N , the boundary distance. This decision is closely coupled with the selection of the spatial sampling interval (Δr). Thus we have an engineering trade-off; for the best approximation to the analytic solution, the spacing between nodes should be small giving us large N , but because of storage and speed limitations on the computer, we wish to reduce the number of calculations being performed by minimizing N . For this conceptual study, we are more concerned with computational speed, so we reduce N while keeping the accuracy as high as possible at all nodes. We will describe several attempts made in search of a suitable number of nodes.

Many of the difficulties in state reduction of this system lie in the implicit constraint of (2.14) that the head at the $(N+1)^{st}$ node is set to zero (i.e. $h_{N+1} = 0$). Since this node is necessarily placed at a finite position, we notice a discrepancy in this boundary condition from that of (2.3). Reflections from this erroneous boundary may result in poor modeling of other nodes.

In order to study the effects of selected N and boundary node distances, a drawdown simulation program was developed on the CDC 7600 computer. This simulator utilizes the finite-difference model (2.16) with user-selected input parameters to produce drawdown values for a hypothetical aquifer system. Also, these drawdown curves are compared to those produced by the analytic model (2.5) for the same system parameters and the relative error is plotted for all the nodes.

We are most interested in modeling the states that represent the observation well locations. To simplify this initial study, we will assume only one measurement well is penetrating the aquifer at a radial distance of 100 meters from the discharge well. As a criteria for suitable modeling of the system, we specify that the steady-state error of the measurement well drawdown from the analytic values must be less than 2%. This will put restrictions on the minimum boundary distance and the minimum N. Standard system parameters used in all simulations are listed in Table 1. The T, S, and q parameter values were suggested for this pilot study by L. D. Thorson, a hydrological systems modeler at LLL. Only the boundary distance, N, and the spacing scheme are changed for each run of the simulator. Each simulation produces node drawdown and error values for 100 equi-spaced discrete time points in one-tenth of a year (i.e., one sample every 8 3/4 hours).

As a first pass, the simulator is run with uniform spacing to a boundary at 1000 meters with 10 states between the pump and measurement wells. This gives us a total of 95 states in the interval. As suspected, these simulation results, compared to the analytically derived values, are quite good at all nodes except near the boundary (N^{th}) node. The simulated drawdown curves and the associated relative error curves from the analytic model versus time are shown in Figures 4 and 5 for three different node locations; the state nearest to the pump, the measurement well node, and the final (N^{th}) node.

In Figure 4, for the node near the pump, we see reasonable modeling (within 3% error) in spite of the fact that this is the most rapidly changing state and the one of largest magnitude. The observation well drawdown

TABLE 1. PARAMETER LIST

Transmissivity (T)	$1 \times 10^{-8} \text{ m}^3/\text{Pa-sec}$
Storativity (S)	$5 \times 10^{-7} \text{ m/Pa}$
Pump Rate (q)	$.01 \text{ m}^3/\text{sec}$
Measurement Well Distance (r_w)	100 m
Number of Nodes between Pump and Well (w)	10

Figure 4. Drawdown and Model Error Curves for 95 State Model

Figure 5. Drawdown and Model Error Curves for 95 State Model

(Figure 4), as with most of the central nodes in the interval, falls well within our criterion of 2% error at steady state. The large initial relative error can be attributed to the very low magnitude of drawdown at that time and to poor modeling of the time delay. For that matter, neither model, analytic nor state-space, adequately portrays the delay region before the pump effects reach a certain node. That is, both models compute an instantaneous drawdown after starting the pump for nodes far away from the pump, however a pure time delay would occur in a physical aquifer system before the wave of groundwater movement approaches distant nodes. Hydrogeologists, in fact, ignore this initial part of the drawdown curve because of potential errors.

At the boundary node, Figure 5, we see that the relative error is quite high, but notice that we only slightly emerge from the delay region over the time interval. Also, the magnitude of drawdown is very small compared to that at the measurement well, which is necessary to reduce boundary effects. As we mentioned before, we are not as concerned with the boundary node dynamics as we are with the measurement well.

Diagrams of the drawdown and the relative error versus radial distance at the final time ($t_f \approx 7$ month) are also shown on the bottom of Figure 5. We see again graphically that the relative errors at the boundaries are larger than in the central portion of the interval.

This simulation satisfies the criterion established earlier in that the relative error at the measurement well node is less than 2% at steady-state. However, we wish to reduce the number of states in the model to ease the

burden on the computer. For this purpose, an alternate means of spatial sampling is used. This sampling scheme involves uniform spacing as before to a point beyond the measurement well, followed by exponentially increasing spacing to the boundary. In this way, we provide a means of rapidly reaching the boundary without requiring excessive numbers of states. Also, because the drawdown decreases with increasing distance from the pump, modeling errors caused by nonlinear spacing should be minimal.

Additional parameters to be specified for this type of discretization include: 1) number of uniformly spaced samples to be placed beyond the measurement well, and 2) multiplicative factor for exponential spacing. A diagram of the spacing in the full interval is shown in Figure 6, where r_w and r_L are measurement well and last uniformly-spaced nodes respectively, and d is the multiplicative factor.

A simulation run of the same aquifer system using the new discretization scheme with the following different input information:

Number of Linear Samples (L)	- 16
Multiplier (d)	- 2.0
Boundary Location (r_{N+1})	- > 50,000 m

All other input parameters remained the same from the previous simulation attempt. The boundary is placed far away (50,000 m) to reduce the effects of boundary reflections. The multiplier of 2.0 means that, after L nodes from the pump, each inter-node space is twice the length of the previous one.

Figure 6. New Spatial Sampling Over Total Interval

For these conditions, a 28-state simulation model is produced. This is a large savings, in number of nodes, from the previous attempt. We can expect some decrease in performance because of the model order reduction, and indeed see this to be the case in Figures 7 and 8, which are analogous to Figures 4 and 5. The boundary node is much further from the pump than in the previous case, so it has a much lower magnitude of drawdown. The states near the pump and measurement wells (Figure 7) show only slightly higher relative error values than with complete linear spacing, and we see that the criterion for modeling ability ($< 2\%$ steady-state error at the measurement well) is still satisfied. Figure 8 again shows the drawdown and relative error curves versus distance at the final time and we see the same modeling problems at the boundaries as before. We have arrived at some very desirable results: the number of states has been drastically reduced while the mathematical description remains acceptable.

Suppose, as a second order reduction step, we simply move the boundary from 50,000 meters to 1,000 meters with all other parameters identical. In this case, the number of states is 21 and the model still remains acceptable. In fact, the drawdown and error curves for the pump and measurement well nodes are identical to the last model attempt (Figure 9). Only the drawdown at the boundary (Figure 10) is different than before, but due to the small magnitude at that point, the boundary effect is not propagated back to the other states.

Figure 7. Drawdown and Model Error Curves for 28 State Model

Figure 8. Drawdown and Model Error Curves for 28 State Model

Figure 9. Drawdown and Model Error Curves for 21 State Model

Figure 10. Drawdown and Model Error Curves for the 21 State Model

A summary of the simulator results we have seen is shown in Table 2. The simulation has been successful and reducing the complexity of the model has been beneficial. The next section describes how this state-space model can be incorporated into an estimator algorithm for reconstructing the drawdown pattern.

TABLE 2
SUMMARIZED SIMULATION RESULTS

	Type of Spacing	Boundary Location (m)	Final Error (%) at Distance		
			5 m	100 m	1000 m
99 State Model	Uniform Spacing	> 1,000	3.2	0.2	36.0
28 State Model	Uniform and Exponential	> 50,000	3.2	1.5	10.5
21 State Model	Uniform and Exponential	> 1,000	3.2	1.5	~100

III. STATE ESTIMATOR DESIGN

As a next step toward the identification of the system parameters (T and S from the previous section), we develop a linear state estimation algorithm assuming knowledge of the parameters. Given noisy measurement data from simulated observation wells, we estimate the drawdown at all nodes. In this section we show that estimation is possible for the previously developed aquifer system and we describe the method.

Recall from the last section that the dynamic model of the aquifer system is

$$\dot{x}(t) = \left(\frac{T}{S}\right) Fx(t) + \left(\frac{1}{S}\right) Gu(t) \quad (3.1)$$

where F and G are given in (2.16). Now the observation well represents the discrete-time measurement system that is of the form⁴:

$$\begin{aligned} z_k &= H x_k + v_k \\ &= h(r_w, t) + v_k \end{aligned} \quad (3.2)$$

where

z_k is the measurement at discrete-time k,

H is the measurement matrix given by $H = e_w^T$ (a unit row vector with a 1 at the w^{th} element),

$x_k = x(k\Delta t)$ where Δt = sampling interval of measurements,

v_k is random noise representing errors in measurements.

We assume the noise to be Gaussian with zero mean and variance σ^2 , i.e., $v_k \sim N(0, \sigma^2)$ for all time k . This allows us to model the uncertainties in the measurement instrumentation. A graph of the simulated drawdown measurement from the 28 state model (in Section II-D) with additive Gaussian noise of variance 10^8 meters² (standard deviation is 10% of full scale) is shown in Figure 11.

Design of the state estimator is straightforward using the extended Kalman filter (EKF) algorithm for a linear system (Table 3) and the DYNEST code developed at Lawrence Livermore Laboratory⁵. In order to implement the algorithm, we must supply both system and measurement Jacobians that are given by

$$F_j(x) \equiv \frac{\partial}{\partial x} \left[\begin{pmatrix} T \\ S \end{pmatrix} Fx(t) \right] = \begin{pmatrix} T \\ S \end{pmatrix} F \quad (3.3)$$

$$H_j(x) \equiv \frac{\partial}{\partial x} [Hx_k] = H \quad (3.4)$$

The linear Kalman filter was implemented using DYNEST with the 21 state reduced order model to see how well the estimator could track the states. The 28-state model was used as a "truth" model to construct error plots for the estimator. Thus, we expect a slight mismatch between the models, especially at the nodes approaching the boundary.

Estimator output curves and the associated estimation error curves for the three representative states are shown in Figures 12 and 13. We see good tracking of states 1 and 10 (Figure 12) indicating a fairly good match between

Figure 11. Simulated Noisy Measurements ($\sigma^2 = 10^8$) at the Observation Well (Node 10; 100 Meters from Pump)

TABLE 3.

SUMMARY OF CONTINUOUS-DISCRETE EXTENDED KALMAN FILTER

Prediction:

$$\hat{x}_{k+1|k} = \hat{x}_{k|k} + \int_k^{k+1} [f(\hat{x}_{\alpha|k}) + g(u_{\alpha})] d\alpha$$

$$\tilde{x}_{k+1|k} = \int_k^{k+1} [F(\hat{x}_{\alpha|k})\tilde{x}_{\alpha|k} + \tilde{x}_{\alpha|k}F^T(\hat{x}_{\alpha|k}) + Q_{\alpha}] d\alpha$$

Innovation:

$$e_{k+1} = z_{k+1} - h(\hat{x}_{k+1|k})$$

$$R_{k+1|k}^e = H(\hat{x}_{k+1|k})\tilde{x}_{k+1|k}H^T(\hat{x}_{k+1|k}) + \sigma_{k+1}^2$$

Correction:

$$K_{k+1} = \tilde{x}_{k+1|k}H^T(\hat{x}_{k+1|k})(R_{k+1|k}^e)^{-1}$$

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1}e_{k+1}$$

$$\begin{aligned} \tilde{x}_{k+1|k+1} = & [I - K_{k+1}H(\hat{x}_{k+1|k})]\tilde{x}_{k+1|k}[I - K_{k+1}H(\hat{x}_{k+1|k})]^T \\ & + K_{k+1}R_{k+1}K_{k+1}^T \end{aligned}$$

where

$$F(\hat{x}_{k|k}) := \left. \frac{\partial f(x)}{\partial x} \right|_{x=\hat{x}_{k|k}}, \quad H(\hat{x}_{k|k}) := \left. \frac{\partial h(x)}{\partial x} \right|_{x=\hat{x}_{k|k}}$$

and

$$\tilde{x}_{k+1|k} := \text{Cov}(\tilde{x}_{k+1|k}) \text{ for } \tilde{x}_{k+1|k} := x_{k+1} - \hat{x}_{k+1|k}$$

$$x_0 \sim N(\hat{x}_{0|0}, \tilde{x}_{0|0})$$

Figure 12. State Estimation of Nodes 1 and 10 using Simulated Noisy ($\sigma^2 = 10^8$) Data; Dotted Lines are 2σ bounds

Figure 13. State Estimation (Node 21) and Innovations using Simulated Noisy ($\sigma^2 = 10^8$) Data

models. Quantitatively, the tracking or estimation error is depicted in Figures 12 and 13, along with the 2σ bounds (dotted lines on plots) predicted by the filter. The error curve for state 21 (Figure 13) shows very poor estimation for this particular state since the curve extends well outside of the 2σ lines. This is due to differences between the simulator and estimator models.

Model mismatch is indicated by the statistical properties of the innovations sequence, i.e., $\{\epsilon_k\}$ is a zero mean, white sequence when the filter is operating properly. Figure 13 shows whiteness in the innovations sequence, but also a high mean value indicating that the model in the filter is slightly wrong producing the bias. In spite of this, the drawdown estimation at the state of interest (the observation well) is very good; i.e. steady-state error variance of ~ 180 (meter²).

In the next section, we will use the extended Kalman filter for the nonlinear problem of state and parameter estimation.

IV. PARAMETER ESTIMATION/IDENTIFICATION

The final step of parameter estimation from the simulated aquifer data is described in this section. In this step, the system model is nonlinear due to the parameter identification procedure and the evaluation of the Jacobians is more involved. The complexity increases because the parameters (T and S) which were previously constant, are now assumed to be unknown and must be estimated.

In the same way that a model was constructed for the state estimator, we must establish a dynamic model for the new system. This is merely an extension of the simulator model of Section II and will be developed in this section.

The two parameters of interest for this system, T and S, are now assumed to be unknown. They can therefore be "augmented" to the states of the system.^{6,7} The unknown parameter vector θ may be written as

$$\theta = \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} = \begin{bmatrix} T \\ S_R \end{bmatrix} \quad (4.1)$$

where $S_R = 1/S$. The reciprocal of the storage coefficient is used for simplification of the Jacobian calculation. Since we have a time-invariant system, these two parameters have no process dynamics; i.e.,

$$\frac{d\theta}{dt} = 0 \quad (4.2)$$

If we treat these parameters as states of the system, they can be included in a new augmented form of the state vector.

Now the general form of the model for this nonlinear dynamical system is

$$\dot{s}(t) = f(x, \theta) + g(u, \theta) \quad (4.3)$$

with
$$s(t) = \begin{bmatrix} x(t) \\ \theta(t) \end{bmatrix} = \begin{bmatrix} h_1 \dots h_N | \theta_1 \theta_2 \end{bmatrix}^T = \text{augmented state vector}$$

$$f(x, \theta) = \begin{bmatrix} \theta_1 \theta_2 Fx(t) \\ 0 \end{bmatrix} \quad (4.4)$$

and
$$g(u, \theta) = \begin{bmatrix} \theta_2 G u(t) \\ 0 \end{bmatrix} \quad (4.5)$$

The discrete-time measurement system is now given by

$$z_k = H_{\theta} s_k + v_k \quad (4.6)$$

where
$$H_{\theta} = \begin{bmatrix} H & | & 0 & 0 \end{bmatrix}$$

In spite of the fact that the new augmented system is nonlinear, the EKF algorithm presented earlier can estimate the drawdown and parameter values. We must first evaluate the system and measurement Jacobians. These can be directly determined in continuous time as follows:

$$F_J(s) = \frac{\partial f(x, \theta)}{\partial s}$$

$$= \frac{N}{2} \left\{ \begin{array}{c|c} \overbrace{\theta_1 \quad \theta_2 \quad F}^N & \overbrace{\theta_2 \quad Fx(t) \quad \theta_1 \quad Fx(t)}^2 \\ \hline 0 & 0 \end{array} \right\} \quad (4.7)$$

and

$$H_J(s) = \frac{\partial}{\partial s} [H_\theta s] = H_\theta \quad (4.8)$$

The estimator/identifier with these Jacobians was implemented using the same 21-state model from the last section augmented by two parameter states, $\hat{T}$ and $\hat{S}_R$. The nonlinear stochastic differential equations of (3.3) can now be solved recursively with the EKF algorithm (Table 3).

Initial estimates of the states must be supplied to the EKF algorithm, so the parameter states are "guessed" with large initial errors (100%) from true value for the simulator. The noisy data of Figure 12 is used for input to the EKF.

Typical output from the estimator/identifier for this set of data is shown in Figures 14 and 15. In Figure 14, we see the "reconstructed" drawdown estimates with associated estimation error curves ($\tilde{x} \equiv \hat{x} - x_{TRUE}$) and 2 sigma bounds predicted by the estimator, i.e., the 95% confidence limits. We see that initially the estimates are in large error due to the uncertainty in the parameters, but eventually good tracking is observed. In fact, in steady-state the error standard deviations are $2 \times 10^3 m$ and $7 \times 10^2 m$ which are quite reasonable.

Figure 14. Drawdown Estimates Reconstructed from Noisy ($\sigma=10^4$ m) Observation Well Measurements

Figure 15. Transmissivity and Reciprocal Storage Coefficient Estimates from Noisy ($\sigma=10^4 \text{ m}$) Observation Well Measurements

Simultaneous parameter estimates of transmissivity and reciprocal storage coefficients ($S_R=1/S$) are also shown for this run in Figure 15. Here we see the estimates with associated estimation errors. For this run the estimator performed quite well with relative errors of approximately 0.3 and 8.5 percent respectively for T and S_R .

To study the performance of the identifier, an ensemble of EKF runs were made. The estimator/identifier was run ten times with the same input data except that the initial parameter estimates were selected from a Gaussian distribution around the true value with a 25% error standard deviation, i.e.,

$$\hat{\theta}_0 \sim N \left[\theta_{\text{TRUE}}, (.25 \times \theta_{\text{TRUE}})^2 \right].$$

The results of the ten runs are tabulated in Table 4. Notice that the estimated means of both parameters are very close to the true value (< 8%) yet the standard deviation of the estimates is very small.

This implies that the precision of the estimator is very high (i.e., it converges to nearly the same value each time), but the accuracy is not as good (i.e., the converged value is not equal to the true value). This fact implies that the estimator is biased.

The same conclusion can be derived from the constant probability contours (see [9] for details) of the estimates shown in Figure 16. The ellipse has collapsed visually to a line centered at the sample mean due to scaling. The ellipse contains the sample estimates ($\hat{T}$, $\hat{S}_R$) with a probability of 0.99. The true value of these parameters is also plotted and the bias of the estimator, though not large, is apparent.

PARAMETER	TRUE VALUE (x_{TRUE})	ESTIMATED MEAN ($\bar{x}$)	STANDARD DEVIATION ($\sigma_{\bar{x}}$)	MEAN ERROR (BIAS) ($\bar{x}$)	% MEAN ERROR ($\bar{x}$, %)	ERROR STD ($\sigma_{\bar{x}}$)	MAX ERROR ($\bar{x}_{\text{max}}$)	MIN ERROR ($\bar{x}_{\text{min}}$)
T ($\times 10^{-8}$)	1.00000	1.00510	.00070	-.00510	.510	.00070	.00579	.00376
$S_R (=1/5)$ ($\times 10^6$)	2.00000	2.07380	.00319	-.07380	7.380	.00319	.07662	.06754

TABLE 4. PARAMETER SAMPLE STATISTICS FOR AN EXAMPLE OF TEN RUNS

Figure 16. Parameter Estimate Contour of Constant Probability ($p \approx 0.99$)

V. CONCLUSIONS

A simplified aquifer model has been introduced and implemented in the general extended Kalman filter computer code DYNEST in order to estimate heads and physical aquifer parameters. For the specific problem addressed in this report, the pump-drawdown test, we found reasonable success ($< 10\%$ relative error) in estimating the aquifer parameters. In solving this problem, we also found that:

- Partial differential equation models can be utilized in the EKF by transforming them to a set of first-order ordinary differential equations (state space model) using the method of finite differences.
- Boundary conditions can easily be incorporated in the state space model.
- Inter-nodal distances may be altered to reduce the model size while retaining acceptable accuracy.
- A linear stochastic estimator using the reduced order model provides good state estimates given noisy measurements.
- The state-space model is easily expanded for the nonlinear parameter identification problem by augmentation of the state vector and reasonable parameter estimates results.

Although the aquifer system presented in this paper was a rather simple one by current geological standards, the research performed with the system model suggests that stochastic methods of estimation are well suited to further applications in this field. Current activity in the area of large-scale systems at LLL will allow us to incorporate higher-order models of more complex hydrogeological structures. Further statistical methods (including maximum likelihood identification, kriging techniques, etc.) are also being studied for possible applications in the geoscience area.

ACKNOWLEDGMENTS

The author would like to sincerely thank D. T. Gavel, D. R. Dunn, N. K. Madsen, L. D. Thorson, and D. G. Dudley for their technical assistance and their willingness to discuss ideas. Also, many thanks to Ms. Deborah Koehler for typing the manuscript.

REFERENCES

1. S. N. Davis and R. J. M. DeWiest, Hydrogeology, John Wiley & Sons, Inc., New York, 1966.
2. C. V. Theis, "The Relation Between the Lowering of the Piezometric Surface and the Rate and Duration of Discharge of a Well using Groundwater Storage", Trans. Am. Geophys. Union, Vol. 16, pp. 519-524, 1935.
3. R. J. M. DeWiest, "On the Storage Coefficient and the Equations of Groundwater Flow", Journal of Geophys. Research, Vol. 71, no. 4, pp. 1117-1122, February, 1966.
4. A. Gelb, et al., Applied Optimal Estimation, M.I.T. Press, Cambridge, Massachusetts, 1974.
5. R. Castleton and J. V. Candy, DYNЕСТ - A Dynamic Estimator Calculation Program, Lawrence Livermore Laboratory Report, UCRL-52573, February 1979.
6. J. V. Candy, On-Line Structural Response Analysis: Using the Extended Kalman Estimator/Identifier, Lawrence Livermore Laboratory Report, UCID-18175, May 1979.
7. J. Wilson, P. Kitanidis, and M. Dettinger, "State and Parameter Estimation in Groundwater Models", Proceedings of AGU Chapman Conference, University of Pittsburgh, Pittsburgh, Pennsylvania, pp. 657-679, May 1978.
8. L. Ljung, "Asymptotic Behavior of the Extended Kalman Filter as a Parameter Estimator for Linear Systems", IEEE Trans. on Autom. Contr., AC-14, pp. 36-50, 1979.
9. S. G. Azevedo and D. T. Gavel, A Joint Gaussian Probability Plot Program: ECP2D, Lawrence Livermore Laboratory Report, UCID-18513, January, 1980.

APPENDIX A

BOUNDARY CONDITIONS

We will show here the derivation of the boundary condition at the pump well (Equation 2.4) and also the equivalence of the boundary conditions

$$\lim_{r \rightarrow \infty} h(r, t) = 0 \quad (A.1)$$

and

$$\lim_{r \rightarrow \infty} \left[r \frac{\partial h(r, t)}{\partial r} \right] = 0 \quad (A.2)$$

for the hydraulic equation (2.1).

The analytic expression for the solution to the drawdown equation was given in Section II as

$$h(r, t) = \frac{-q}{4\pi T} \int_{\frac{r^2 S}{4Tt}}^{\infty} \frac{e^{-x}}{x} dx, \quad (A.3)$$

which implicitly contains the constraint of (A.1).

Now let us calculate the variation of drawdown with respect to r :

$$\frac{\partial h}{\partial r} = \frac{\partial h}{\partial \alpha} \frac{\partial \alpha}{\partial r} \quad (A.4)$$

where $\alpha = \frac{r^2 S}{4Tt}$.

It is known that, in general,

$$\frac{d}{da} \int_a^b f(u) du = -f(a)$$

so

$$\begin{aligned} \frac{\partial h}{\partial \alpha} &= \frac{-q}{4\pi T} \frac{-e^{-\alpha}}{\alpha} \\ &= \frac{q(4Tt)}{4\pi T(r^2 S)} e^{-\frac{Sr^2}{4Tt}} \\ \frac{\partial h}{\partial \alpha} &= \frac{qt}{\pi r^2 S} e^{-\frac{Sr^2}{4Tt}} \end{aligned} \quad (A.5)$$

Also,

$$\frac{\partial \alpha}{\partial r} = \frac{\partial}{\partial r} \left(\frac{r^2 S}{4Tt} \right) = \frac{rS}{2Tt} \quad (A.6)$$

Now combining (A.5) and (A.6) into (A.4) we get

$$\begin{aligned} \frac{\partial h}{\partial r} &= \left(\frac{rS}{2Tt} \right) \left(\frac{qt}{\pi r^2 S} \right) e^{-\frac{Sr^2}{4Tt}} \\ \frac{\partial h}{\partial r} &= \frac{q}{2\pi r T} e^{-\frac{Sr^2}{4Tt}} \end{aligned} \quad (A.7)$$

From this we can directly find the boundary conditions:

$$r \frac{\partial h}{\partial r} = \frac{q}{2\pi l} e^{-\frac{Sr^2}{4Tt}}$$

so

$$\lim_{r \rightarrow 0} \left(r \frac{\partial h}{\partial r} \right) = \frac{q}{2\pi l}$$

(Boundary near the pump)

$$\lim_{r \rightarrow \infty} \left(r \frac{\partial h}{\partial r} \right) = 0$$

(Boundary at infinity)