## Title

On solution of the inverse problem for confined aquifer flow via maximum likelihood



Journal
Mathematical Geosciences, 18(7)

ISSN
1874-8961

Authors
Loaiciga, Hugo A
Mariño, Miguel A

Publication Date
1986-10-01

DOI
10.1007/bf00897096

Peer reviewed

*Mathematical Geology, Vol. 18, No. 7, 1986*

# On Solution of the Inverse Problem for Confined Aquifer Flow Via Maximum Likelihood[1]

**Hugo A. Loaiciga[2] and Miguel A. Mariño[3]**

*Joint estimation of transmissivity (T) and storativity (S) in a confined aquifer is done via maximum likelihood (ML). The differential equation of groundwater flow is discretized by the finite-element method, leading to equation $\psi\varphi_t + \Gamma x_t = u_t$. Elements of matrices $\psi$ and $\Gamma$, as well as estimated covariance matrix of noise term $u_t$, are functions of T and S. By minimizing the negative log-likelihood function corresponding to discretized groundwater flow equation with respect to T and S, ML estimators are obtained. The ML approach is found to yield accurate estimates of T and S (within 9 and 10% of their actual values, respectively) and showed quadratic convergence in Newton's search technique. Prediction of aquifer response, using ML estimators, results in estimated piezometric heads accurate to $\pm 0.5$ m from their actual, exact values. Statistical properties of ML estimators are derived and some basic results for statistical inference are given.*

**KEY WORDS:** transmissivity, storativity, maximum likelihood, linear regression.

## INTRODUCTION

Historically, two main approaches to estimating transmissivities and storativities have been used: (1) methods based on aquifer tests [see, e.g., Mariño and Luthin, 1982, p. 291–336]; and (2) statistical or mathematical programming techniques [see, e.g., McLaughlin (1975), Neuman (1980), Cooley (1982), Yeh et al. (1983), Kitanidis and Vomvoris (1983), and Aboufirassi and Mariño (1984)]. Proper estimates of groundwater flow parameters are indispensable for simulation and prediction of aquifer response to natural or artificial inputs to a groundwater reservoir.

This paper addresses estimation of groundwater flow parameters in confined aquifers via maximum likelihood (ML). The equation for groundwater flow is discretized via the finite-element method and expressed in a linear form. The negative log-likelihood function of the discretized equation is subsequently derived and minimized (using Newton's method) with respect to unknown pa-

---

[1]Manuscript received 2 December 1985; accepted 20 March 1986.
[2]Department of Land, Air, and Water Resources, University of California, Davis, California 95616.
[3]Department of Civil Engineering, University of California, Davis, California 95616.

rameters. Objectives of this paper are: (1) to develop an ML estimator for transmissivities and storativities in confined aquifers; (2) to apply the proposed methodology and compare it with a linear estimation technique; and (3) to assess statistical properties of ML estimators.

## PROBLEM DESCRIPTION

The governing equation for two-dimensional flow in a confined aquifer is

$$\frac{\partial}{\partial x}\left(T\frac{\partial \phi}{\partial x}\right) + \frac{\partial}{\partial y}\left(T\frac{\partial \phi}{\partial y}\right) - F = S\frac{\partial \phi}{\partial t} \tag{1}$$

in which $\phi$, $T[=T(x, y)]$, $S[=S(x, y)]$, and $F[=F(x, y, t)]$ denote piezometric head, transmissivity, storativity, and a sink/source, respectively. Following Loaiciga and Mariño (1986), eq. 1 can be discretized by the finite-element method within the flow field to yield

$$\underset{(G \times G)}{\psi}\ \underset{(G \times 1)}{\varphi_t} + \underset{(G \times K)}{\Gamma}\ \underset{(K \times 1)}{\mathbf{x}_t} = \underset{(G \times 1)}{\mathbf{u}_t} \qquad t = 1, 2, \ldots, n \tag{2}$$

with $\varphi_0$ assumed known.

Matrices $\psi$ and $\Gamma$ contain unknown elements that are functions of transmissivities and storativities; $\varphi_t$ contains unknown nodal heads in the flow domain; $\mathbf{x}_t$ contains, as its first $G$ elements, values of nodal heads at time $t - 1$ (i.e., $\varphi_{t-1}$) and remaining $K - G$ $(K \geq G)$ elements are functions of boundary conditions and/or sink/source distribution throughout the field (an example is given subsequently); $\mathbf{u}_t$ is a vector of random disturbances that account for errors in approximating eq. 1 by eq. 2 and that is assumed to have the following properties

$$E(\mathbf{u}_t) = 0 \tag{3}$$

$$E(\mathbf{u}_t\mathbf{u}_t^T) = \Sigma \tag{4}$$

$$E(\mathbf{u}_t\mathbf{u}_s^T) = 0, \qquad s \neq t \tag{5}$$

Equations 3–5 specify that $\mathbf{u}_t$ is a white noise disturbance vector.

Under the assumption that $\mathbf{u}_t$ has a multivariate normal distribution, the likelihood function associated with eq. 2 is

$$L = \frac{|\psi|^n}{(2\pi)^{nG/2}} |\Sigma|^{-n/2} \exp\left\{-\frac{1}{2}\sum_{t=1}^{n} (\psi\varphi_t + \Gamma\mathbf{x}_t)^T \Sigma^{-1}(\psi\varphi_t + \Gamma\mathbf{x}_t)\right\} \tag{6}$$

For estimation purposes, it is convenient to use the logarithm of eq. 6 to obtain the following log-likelihood function

$$\ln L = -\frac{nG}{2} \ln (2\pi) - \frac{n}{2} \ln |\Sigma| + n \ln |\psi| - \frac{1}{2} \sum_{t=1}^{n} (\psi \boldsymbol{\varphi}_t + \Gamma \mathbf{x}_t)^T \Sigma^{-1} (\psi \boldsymbol{\varphi}_t + \Gamma \mathbf{x}_t) \tag{7}$$

Unknown matrices $\psi$ and $\Gamma$ have elements that are functions of transmissivities and storativities. Therefore, one must maximize eq. 7 with respect to a parameter vector $\boldsymbol{\theta}$ $(q \times 1)$ whose elements are unknown transmissivities and storativities, which vary within subdomains of the finite-element spatial discretization. The problem at hand consists of maximizing eq. 7 with respect to $\boldsymbol{\theta}$.

## NUMERICAL ESTIMATION ALGORITHM

In order to adhere to standard mathematical programming conventions, the negative log-likelihood function eq. 7 is minimized. Minimization is numerically via Newton's method. By letting $f = -\ln L$, search for an optimum starts at some specified value $\boldsymbol{\theta}_0$. Subsequent iteration values, say the $(k + 1)$st, are obtained by first taking a second-order Taylor expansion of $f$ about the current ($k$th) point, i.e.

$$f(\boldsymbol{\theta}_k + \mathbf{p}_k) = f(\boldsymbol{\theta}_k) + \nabla f(\boldsymbol{\theta}_k)^T \mathbf{p}_k + \tfrac{1}{2} \mathbf{p}_k^T G(\boldsymbol{\theta}_k) \mathbf{p}_k \tag{8}$$

in which $\nabla f(\boldsymbol{\theta}_k)$ and $G(\boldsymbol{\theta}_k)$ are gradient and matrix of second derivatives of $f$ at the current point $\boldsymbol{\theta}_k$, respectively. The right-hand side of eq. 8 is minimized by $\mathbf{p}_k$ (the step to the next point) given by

$$G(\boldsymbol{\theta}_k)\mathbf{p}_k = -\nabla f(\boldsymbol{\theta}_k) \tag{9}$$

Upon solution of eq. 9 for $\mathbf{p}_k$, the next iteration point is

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k + \mathbf{p}_k \tag{10}$$

In order to guarantee convergence to a local optimum, a steplength factor $\alpha_k$ customarily is introduced such that

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k + \alpha_k \mathbf{p}_k \tag{11}$$

where $\alpha_k$ is selected to minimize $f(\boldsymbol{\theta}_k + \alpha_k \mathbf{p}_k)$ with respect to $\alpha_k$. Equation 11 is used iteratively until convergence is reached.

Conveniently, eq. 7 is simplified by taking its derivative with respect to $\Sigma$, equating to zero and solving for $\Sigma$, to obtain estimator $\hat{\Sigma}$

$$\hat{\Sigma} = \frac{1}{n} \left[ \sum_{t=1}^{n} (\psi \boldsymbol{\varphi}_t + \Gamma \mathbf{x}_t)(\psi \boldsymbol{\varphi}_t + \Gamma \mathbf{x}_t)^T \right] \tag{12}$$

By substituting eq. 12 into eq. 7 and multiplying the resulting expression by $-1$, one obtains the following expression for negative log-likelihood function $f$

$$f = \frac{nG}{2} \ln (2\pi) + \frac{n}{2} \ln \left| \frac{1}{n} \sum_{t=1}^{n} (\psi \boldsymbol{\varphi}_t + \Gamma \mathbf{x}_t)(\psi \boldsymbol{\varphi}_t + \Gamma \mathbf{x}_t)^T \right| - n \ln |\psi| + \frac{n^2}{2}$$

$$= C + \frac{n}{2} \ln |\hat{\Sigma}| - n \ln |\psi| \tag{13}$$

in which

$$C = \frac{nG}{2} \ln (2\pi) + \frac{n^2}{2} = \text{constant} \tag{14}$$

In order to compute $\nabla f$ and $G$ at $\boldsymbol{\theta}_k$, matrix derivative results are useful (Graybill, 1983)

$$\frac{\partial \ln |A|}{\partial \theta_i} = \text{tr} \left[ A^{-1} \frac{\partial A}{\partial \theta_i} \right] \qquad |A| > 0 \tag{15}$$

$$\frac{\partial A^{-1}}{\partial \theta_i} = -A^{-1} \frac{\partial A}{\partial \theta_i} A^{-1} \tag{16}$$

$$\frac{\partial^2 \ln |A|}{\partial \theta_i^2} = \text{tr} \left[ -A^{-1} \frac{\partial A}{\partial \theta_i} A^{-1} \frac{\partial A}{\partial \theta_i} + A^{-1} \frac{\partial^2 A}{\partial \theta_i^2} \right] \tag{17}$$

in which $A$ should be replaced by $\hat{\Sigma}$ or $\psi$ in actual computations.

Equations 15–17 provide elements of $\nabla f$ and $G$ that are evaluated at the current point in iterations of the Newton method. Optimum value $\boldsymbol{\theta}^*$ that minimizes eq. 13 is the maximum likelihood (ML) estimator of unknown parameter vector $\boldsymbol{\theta}$.

## PROPERTIES OF ESTIMATORS

Suppose $\tilde{\boldsymbol{\theta}}$ denotes the true but unknown vector of parameters. Let

$$I(\boldsymbol{\theta}) = E(\partial^2 f / \partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T) \tag{18}$$

$$i(\boldsymbol{\theta}) = \partial^2 f / \partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T \tag{19}$$

in which the expectation in eq. 18 is with respect to $\boldsymbol{\varphi}_t$. Matrices given in eqs. 18 and 19 are Fisher information and sample information matrices, respectively. For a sample size $n$ sufficiently large, distribution of ML estimator $\boldsymbol{\theta}^*$ is approximately

$$\boldsymbol{\theta}^* \sim N[\tilde{\boldsymbol{\theta}}, I^{-1}(\tilde{\boldsymbol{\theta}})] \tag{20}$$

in which $N$ denotes the multivariate normal distribution. From eq. 20, it follows that expression

$$(\boldsymbol{\theta} - \boldsymbol{\theta}^*)^T I^{-1}(\boldsymbol{\theta}^*)(\boldsymbol{\theta} - \boldsymbol{\theta}^*) \leq \chi^2_\alpha(q) \tag{21}$$

in which $\chi^2_\alpha(q)$ is the $(1 - \alpha)$th percentile of a $\chi^2$ variable with $q$ degrees of freedom, represents an ellipsoid in $q$-dimensional $\boldsymbol{\theta}$ space centered at $\boldsymbol{\theta}^*$; probability that this random ellipsoid covers the true parameter $\tilde{\boldsymbol{\theta}}$ is $1 - \alpha$. Equation 20 allows construction of a hypothesis test. Let

$$H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}^0$$

$$H_1 : \boldsymbol{\theta} \neq \boldsymbol{\theta}^0 \tag{22}$$

in which $\boldsymbol{\theta}^0$ is the specified value in the null hypothesis $H_0$. The null hypothesis $H_0$ is rejected at a significance level $\alpha$ if

$$(\boldsymbol{\theta}^0 - \boldsymbol{\theta}^*)^T I^{-1}(\boldsymbol{\theta}^0)(\boldsymbol{\theta}^0 - \boldsymbol{\theta}^*) > \chi^2_\alpha(q) \tag{23}$$

In practice, $I(\cdot)$ (see eq. 18) may be difficult to obtain, so sample information matrix $i(\cdot)$ (see eq. 19) can replace $I(\cdot)$ in eqs. 20, 21, and 23.

Nonlinear least-squares methods [see, e.g., Neuman (1980) and Cooley (1982)] are related to the ML approach presented in this work, in the sense that they are based on the discretized groundwater flow equation that leads to minimization of quadratic or nonlinear functions and impose similar assumptions on the noise term. In addition, such nonlinear or generalized regression models are solved commonly by means of iterative search techniques based on Newton-type algorithms. Prior information on ML estimators may be incorporated through selection of adequate initial parameter estimates, whereas Cooley's (1982) nonlinear regression approach incorporates such priors by expanding the set of regression equations. Maximum likelihood estimates have desirable asymptotic properties as discussed above and, in addition, small-sample ML estimators usually are more efficient (i.e., have smaller variance) than competing nonlinear least-squares estimators.

## APPLICATION OF THE ESTIMATION PROCEDURE

A one-dimensional confined aquifer (Fig. 1) is used as a test case for the ML estimation approach discussed above. Unknown transmissivity $T$ and storativity $S$ are estimated via ML, and subsequently, matrices $\psi$ and $\Gamma$ are formed by using functional relationships between their elements and $T$ and $S$.

### Generation of Head Values

Data upon which the ML method is implemented are obtained by deriving an analytical solution to the confined aquifer flow problem shown (Fig. 1). The closed-form (exact) solution is utilized to generate nodal values for times $t = 1, 2, \ldots, n$, which are used to substitute for $\boldsymbol{\varphi}_t$, $\forall t$, in eq. 13.

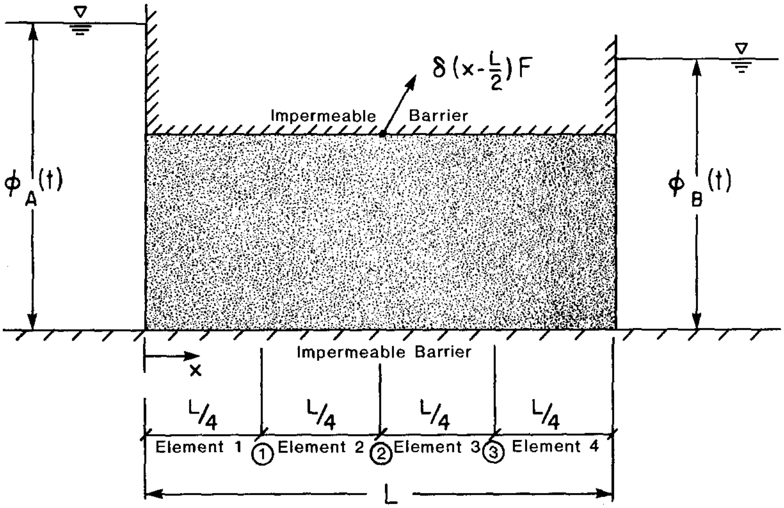


**Fig. 1.** Confined aquifer subject to time-dependent conditions and a discharge (of units $L^3T^{-1}L^{-1}$) at $x = L/2$.

The equation characterizing flow in the aquifer (Fig. 1.) is

$$T\frac{\partial^2\phi}{\partial x^2} - S\frac{\partial\phi}{\partial t} = F\delta\left(x - \frac{L}{2}\right) \tag{24}$$

and boundary and initial conditions are

$$\phi_A(t) = H_A(t) \qquad x = 0 \qquad t > 0 \tag{25}$$

$$\phi_B(t) = H_B(t) \qquad x = L \qquad t > 0 \tag{26}$$

$$\phi(0) = g(x) \qquad 0 \le x \le L \qquad t = 0 \tag{27}$$

By letting

$$X = \frac{x\pi}{L} \tag{28}$$

$$F^* = \frac{F}{S}\delta\left[\frac{L}{\pi}\left(X - \frac{\pi}{2}\right)\right] \quad \text{and} \tag{29}$$

$$c = \frac{T}{S}\left(\frac{\pi}{L}\right)^2 \tag{30}$$

Equations 18–21 become

$$-c\frac{\partial^2\phi}{\partial X^2} + \frac{\partial\phi}{\partial t} = -F^* \tag{31}$$

$$\phi_A(t) = H_A(t) \qquad X = 0, \qquad t > 0 \tag{32}$$

$$\phi_B(t) = H_B(t) \qquad X = \pi \qquad t > 0 \tag{33}$$

$$\phi(0) = g(X) \qquad 0 \leq X \leq \pi \qquad t = 0 \tag{34}$$

Solution to the problem defined by eqs. 31–34 is

$$\phi(X, t) = H_A(t)\left[1 - \frac{X}{\pi}\right] + H_B(t)\frac{X}{\pi} + \sum_{n=1}^{\infty} \cdot \sin(nX)\left[\int_0^t e^{-cn^2(t-u)}F_n(u)\,du + b_n e^{-cn^2 t}\right] \tag{35}$$

in which

$$F_n(t) = \frac{2}{\pi}\int_0^{\pi} \sin(nX)[P(X, t) - F^*(X, t)]\,dX \tag{36}$$

$$b_n = \frac{2}{\pi}\int_0^{\pi} \sin(nX)[g(X) + f_1(X)]\,dX \tag{37}$$

$$P(X, t) = -H_A'(t)\left[1 - \frac{X}{\pi}\right] - H_B'(t)\frac{X}{\pi} \tag{38}$$

$$f_1(X) = -H_A(0)\left[1 - \frac{X}{\pi}\right] - H_B(0)\frac{X}{\pi} \tag{39}$$

Equations 35–39 were utilized to generate head values for time periods $t = 1, 2, \ldots, 20$ (upon which Gaussian white noise was added), corresponding to data $H_A$, $H_B$, $g(x)$, and $F$ (Table 1).

**Table 1.** Basic Data for the Estimation Problem[a]

| | Nodal head ($M$) | | | | Nodal head ($M$) | | |
|---|---|---|---|---|---|---|---|
| Period, $t$ | 1 | 2 | 3 | Period, $t$ | 1 | 2 | 3 |
| 1 | 84.93 | 87.70 | 94.88 | 11 | 85.87 | 82.62 | 87.83 |
| 2 | 84.75 | 86.75 | 94.38 | 12 | 86.18 | 82.36 | 87.16 |
| 3 | 84.62 | 86.01 | 93.72 | 13 | 86.51 | 82.13 | 86.50 |
| 4 | 84.58 | 85.40 | 92.98 | 14 | 86.86 | 81.93 | 85.85 |
| 5 | 84.62 | 84.86 | 92.22 | 15 | 87.23 | 81.74 | 85.22 |
| 6 | 84.72 | 84.38 | 91.46 | 16 | 87.61 | 81.57 | 84.60 |
| 7 | 84.88 | 83.95 | 90.70 | 17 | 88.00 | 81.42 | 84.00 |
| 8 | 85.08 | 83.57 | 89.96 | 18 | 88.40 | 81.28 | 83.40 |
| 9 | 85.31 | 83.22 | 89.23 | 19 | 88.81 | 81.15 | 82.81 |
| 10 | 85.58 | 82.90 | 88.52 | 20 | 89.23 | 81.04 | 82.23 |

[a] Nodal heads were generated with Eq. 35 using $T = 500$ m$^2$/day, $S = 12 \times 10^{-3}$, $l = 500$ m, $\omega = 0.5$, and $\Delta t = 1$ day; $F = 10$ m$^2$/day; $g(x) = H_A(t) + [H_B(t) - H_A(t)]x/L$; $H_A(t) = 80 + t$; $H_B(t) = 100 - t$.

## Computation of (Negative) Log-Likelihood Function

For the aquifer (Fig. 1) with constant values for $T$ and $S$ (the methodology does not require constant parameters, but in order to be able to use the exact solution eq. 35 to generate head values such simplification is needed), the $\psi$ and $\Gamma$ matrices (see eq. 2) are

$$\psi = \begin{bmatrix} \frac{2\omega T}{l} + \frac{2lS}{3\Delta t} & -\frac{\omega T}{l} + \frac{lS}{6\Delta t} & 0 \\ -\frac{\omega T}{l} + \frac{lS}{6\Delta t} & \frac{2\omega T}{l} + \frac{2lS}{3\Delta t} & -\frac{\omega T}{l} + \frac{lS}{6\Delta t} \\ 0 & -\frac{\omega T}{l} + \frac{lS}{6\Delta t} & \frac{2\omega T}{l} + \frac{2lS}{3\Delta t} \end{bmatrix} \quad (40)$$

$$\Gamma = \begin{bmatrix} \frac{2\omega' T}{l} - \frac{2lS}{3\Delta t} & -\frac{\omega' T}{l} - \frac{lS}{6\Delta t} & 0 & -\frac{T}{l} & 0 & \frac{lS}{6} & 0 & 0 \\ -\frac{\omega' T}{l} - \frac{lS}{6\Delta t} & \frac{2\omega' T}{l} - \frac{2lS}{3\Delta t} & -\frac{\omega' T}{l} - \frac{lS}{6\Delta t} & 0 & 0 & 0 & 0 & 1 \\ 0 & -\frac{\omega' T}{l} - \frac{lS}{6\Delta t} & \frac{2\omega' T}{l} - \frac{2lS}{3\Delta t} & 0 & -\frac{T}{l} & 0 & \frac{lS}{6} & 0 \end{bmatrix} \quad (41)$$

in which $\frac{1}{2} \leq \omega \leq 1$ is a weighting factor, $\omega' = 1 - \omega$, $l = L/4$, and $\Delta t$ is the simulation time period. Linear basis functions were used in finite-element discretization of eq. 24. Vector $\mathbf{x}_t$ in eq. 2 is

$$\mathbf{x}_t^T = [\phi_1(t-1), \phi_2(t-1), \phi_3(t-1), \bar{\phi}_A, \bar{\phi}_B, \bar{\dot{\phi}}_A, \bar{\dot{\phi}}_B, \bar{F}] \quad (42)$$

in which $\bar{F} = \omega F(t) + (1 - \omega)\, F(t-1)$ is average discharge (see Fig. 1) at $x = L/2$; $\bar{\phi}_A = \omega\phi_A(t) + (1-\omega)\,\phi_A(t-1)$; $\bar{\dot{\phi}}_A = \{\omega[\phi_A(t) - \phi_A(t-1)] + (1-\omega)[\phi_A(t) - \phi_A(t-1)]\}/\Delta t$; and similar definitions hold for $\bar{\phi}_B$ and $\bar{\dot{\phi}}_B$.

Matrices given in eqs. 40–41 show explicitly the functional relationships between their elements and unknown parameters $T$ and $S$. By having $\psi$, $\Gamma$, $\varphi_t$, and $\mathbf{x}_t, \forall t$, the negative log-likelihood function $f$ of eq. 13, is completely defined, and the expressions given by eqs. 15–17 readily are evaluated when implementing Newton's method.

## Estimation and Analysis of Results

The sample used to estimate $T$ and $S$ is tabulated (Table 1). Piezometric heads were generated by eq. 35 and are accurate to $\pm 0.001$ m (in approximation of the infinite series). Upon differentiation, the negative log-likelihood function (eq. 13) yielded

$$\frac{\partial f}{\partial \theta_i} = \frac{n}{2} \operatorname{tr} \left( \hat{\Sigma}^{-1} \frac{\partial \hat{\Sigma}}{\partial \theta_i} \right) - n \operatorname{tr} \left( \psi^{-1} \frac{\partial \psi}{\partial \theta_i} \right) \tag{43}$$

$$\frac{\partial^2 f}{\partial \theta_i^2} = \frac{n}{2} \operatorname{tr} \left[ -\hat{\Sigma}^{-1} \frac{\partial \hat{\Sigma}}{\partial \theta_i} \hat{\Sigma}^{-1} \frac{\partial \hat{\Sigma}}{\partial \theta_i} + \hat{\Sigma}^{-1} \frac{\partial^2 \hat{\Sigma}}{\partial \theta_i^2} \right] - n \operatorname{tr} \left[ -\psi^{-1} \frac{\partial \psi}{\partial \theta_i} \psi^{-1} \frac{\partial \psi}{\partial \theta_i} \right] \tag{44}$$

$$\frac{\partial^2 f}{\partial \theta_2 \partial \theta_1} = \frac{n}{2} \operatorname{tr} \left[ -\hat{\Sigma}^{-1} \frac{\partial \hat{\Sigma}}{\partial \theta_2} \hat{\Sigma}^{-1} \frac{\partial \hat{\Sigma}}{\partial \theta_1} + \hat{\Sigma}^{-1} \frac{\partial^2 \hat{\Sigma}}{\partial \theta_2 \partial \theta_1} \right] - n \operatorname{tr} \left[ -\psi^{-1} \frac{\partial \psi}{\partial \theta_2} \psi^{-1} \frac{\partial \psi}{\partial \theta_1} \right] \tag{45}$$

Equations 43–45 (with $\theta_1 = T$ and $\theta_2 = S$) were used to evaluate $\nabla f$ and $G$ at the current iteration point $\boldsymbol{\theta}_k$ during the $k$th iteration of Newton's method (see eqs. 9–10). Second derivatives with respect to $\psi$ vanish due to linear dependence of its elements on $T$ and $S$ (see eq. 40). Expanded details on expressions $\partial\hat{\Sigma}/\partial\theta_i$, $\partial^2\hat{\Sigma}/\partial\theta_i^2$ ($i = 1, 2$), and $\partial^2\hat{\Sigma}/\partial\theta_2\partial\theta_1$ are given in Appendix A.

Several initial estimates $\boldsymbol{\theta}_0$ were tried to test if convergence occurred to the same local optimum. Convergence to a unique point occurred for all initial estimates tried (which were within 50% of true values). In all cases, convergence occurred within five iterations of Newton's method, and a quadratic convergence rate was observed. Contour plots of the negative log-likelihood function indicated that it is convex, with a flat surface around the unique local optimum, as shown (Fig. 2) Optimal ML estimator point was $T^* = 456$ m$^2$/day and $S^* = 0.0108$, whereas true values are 500 m$^2$/day and 0.012, respectively. The convergence path for initial estimators $T^{(0)} = 350$, $S^{(0)} = 0.006$, and standard errors of optimal estimators are shown (Table 2). The covariance matrix of ML estimators is approximated by the inverse of the sample information matrix (see eq. 19) evaluated at convergence values $T^* = 456$ m$^2$/day and $S^* = 0.0108$ (Rao, 1965) and is equal to

$$\operatorname{cov}(T^*, S^*) = \begin{bmatrix} 2273.5 & 0.12285 \\ 0.12285 & 0.0000137 \end{bmatrix}$$

Thus, standard errors of $T^*$ and $S^*$ are 47.7 m$^2$/day and 0.00370, respectively. Matrices $\psi$ and $\Gamma$ of the governing flow equation (see eqs. 40–41) were estimated by replacing $T$ and $S$ in those equations by their estimators $T^*$ and $S^*$, respectively. True values of $\psi$ and $\Gamma$, as well as those of their estimators $\hat{\psi}$ and

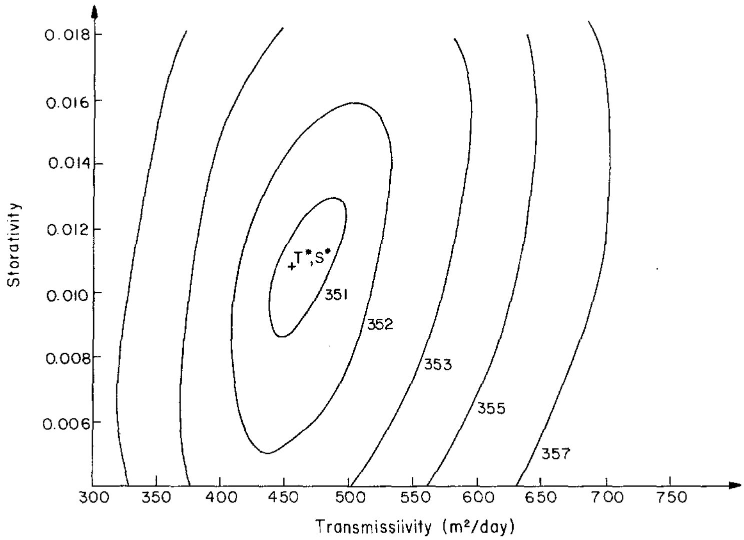


**Fig. 2.** Contour plot of negative log-likelihood function.

$\hat{\Gamma}$, are as follows

$$\psi = \begin{bmatrix} 5 & \frac{1}{2} & 0 \\ \frac{1}{2} & 5 & \frac{1}{2} \\ 0 & \frac{1}{2} & 5 \end{bmatrix} \qquad \hat{\psi} = \begin{bmatrix} 4.337 & 0.282 & 0 \\ 0.282 & 4.337 & 0.282 \\ 0 & 0.282 & 4.337 \end{bmatrix}$$

$$\Gamma = \begin{bmatrix} -3 & -1.5 & 0 & -1 & 0 & 1 & 0 & 0 \\ -1.5 & -3 & -1.5 & 0 & 0 & 0 & 0 & 1 \\ 0 & -1.5 & -3 & 0 & -1 & 0 & 1 & 0 \end{bmatrix}$$

$$\hat{\Gamma} = \begin{bmatrix} -2.197 & -1.352 & 0 & -1.070 & 0 & 0.817 & 0 & 0 \\ -1.352 & -2.197 & -1.352 & 0 & 0 & 0 & 0 & 1 \\ 0 & -1.352 & -2.197 & 0 & -1.070 & 0 & 0.817 & 0 \end{bmatrix}$$

From asymptotic properties of ML estimators (Lehmann, 1983), $\hat{\psi}$ and $\hat{\Gamma}$ are consistent estimators of $\psi$ and $\Gamma$, respectively, i.e.

$$p \lim_{n \to \infty} \hat{\psi} = \psi \qquad \text{and}$$

$$p \lim_{n \to \infty} \hat{\Gamma} = \Gamma$$

Therefore from eq. 2

$$\varphi_t = (-\psi^{-1}\Gamma)\mathbf{x}_t + \psi^{-1}\mathbf{u}_t = \Pi\mathbf{x}_t + \mathbf{e}_t, \ \forall t \tag{46}$$

**Table 2.** Synopsis of Newton's Search

| Iteration | $T^a$ | $S^b$ | $f$ |
|---|---|---|---|
| 0 | 350 | 0.0060 | 356 |
| 1 | 446 | 0.0080 | 352 |
| 2 | 445 | 0.0095 | 351.2 |
| 3 | 453 | 0.0105 | 350.8 |
| 4 | 455 | 0.0107 | 350.5 |
| 5 | 456 | 0.0108 | 350 |
| | (47.7)[c] | (0.00370)[c] | |

[a] $T$ = transmissivity (m$^2$/day).
[b] $S$ = storativity (dimensionless).
[c] Standard errors of optimal estimators.

and expected value of $\varphi_t$, given $\varphi_{t-1}$, is

$$E(\varphi_t|\varphi_{t-1}) = \Pi\mathbf{x}_t, \ \forall t \tag{47}$$

which can be used to simulate expected values of heads for any time $t$, given heads at time $t - 1$. From the invariance property of ML estimators (see, e.g., Bickel and Doksum, 1977, p. 99), $\hat{\Pi}$ and $\hat{\Pi}\mathbf{x}_t = \hat{E}(\varphi_t|\varphi_{t-1})$ are also ML estimators because they are functions of estimators $\hat{\psi}$ and $\hat{\Gamma}$ (see eqs. 46–47).

Table 3 shows the estimated $\hat{\Pi}$ matrix by ML method and by a linear regression (LR) technique (Loaiciga and Mariño, 1986). A derivation of $\hat{\Pi}$ by LR is given in Appendix B. Simulated head values using ML and LR estimators, as well as exact head values obtained from eq. 35 for $t = 1, 2, \ldots, 20$, are shown (Table 4). Observe that LR estimates tend to slightly overestimate head values at early time steps (i.e., $t = 1, 2, \ldots, 10$). For time steps larger than $t = 15$, a slight underestimation of head values by LR technique is apparent. Overall departures from exact values are within $-0.10$ and 0.50 m. ML estimated heads show a tendency to slightly underestimate heads at nodes 1 and 2 for early time steps, but the tendency reverses to a moderate overestimation for time steps after $t = 15$, at nodes 1 and 2. At node 3, estimated values exceed actual ones at early time steps ($t \leq 2$) and subsequently tend to fall below actual head values. Accuracy of estimated heads by ML method is within 0.5 m for the entire simulation period.

In summary, performance of ML and LR estimators based on the predicted response of an aquifer to pumping and time-varying boundary conditions is roughly equivalent. The ML method yields unique and reasonable estimates of $T$ and $S$ (deviations of 9 and 10% were observed about true values, respectively). A disadvantage of the ML technique is the need for prior information on $T$ and $S$ to initialize Newton's method. The LR estimation technique does not require such prior information, is more easily programmable, and, as a subproduct, directly provides estimates of the covariance structure of esti-

**Table 3.** Estimated Parameter Matrix $\hat{\Pi}$[a]

| Column | Row 1 | | Row 2 | | Row 3 | |
|---|---|---|---|---|---|---|
| 1 | 0.48829 | (0.38250) | 0.28118 | (0.30894) | −0.01828 | (0.10821) |
| 2 | 0.28118 | (0.33058) | 0.47001 | (0.48875) | 0.28118 | (0.12666) |
| 3 | −0.01828 | (0.05838) | 0.28118 | (0.12010) | 0.48829 | (0.51841) |
| 4 | 0.24777 | (0.27996) | −0.01618 | (−0.03454) | 0.001052 | (−0.03650) |
| 5 | 0.00105 | (−0.05341) | −0.01618 | (0.09242) | 0.24777 | (0.28496) |
| 6 | −0.18911 | (−0.01873) | 0.01235 | (−0.02350) | $-8.0293 \times 10^{-4}$ | (0.06790) |
| 7 | $-8.0293 \times 10^{-4}$ | (−0.02124) | 0.01235 | (−0.02415) | −0.18911 | (0.06514) |
| 8 | 0.01512 | ($7.6875 \times 10^{-4}$) | −0.23254 | (−0.00874) | 0.01512 | (0.00380) |

[a] Entries show ML and LR (within parentheses) estimates.

**Table 4.** Actual and Forecasted Piezometric Heads[a]

| Time | Node 1 | | | Node 2 | | | Node 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| | A | ML | LR | A | ML | LR | A | ML | LR |
| 0 | 85.00 | — | — | 90.00 | — | — | 95.00 | — | — |
| 1 | 84.93 | 85.09 | 85.05 | 87.70 | 87.68 | 87.99 | 94.88 | 95.21 | 95.29 |
| 2 | 84.75 | 84.72 | 84.77 | 86.75 | 86.67 | 86.95 | 94.38 | 94.42 | 94.85 |
| 3 | 84.62 | 84.52 | 84.64 | 86.01 | 85.87 | 86.18 | 93.72 | 93.51 | 94.11 |
| 4 | 84.58 | 84.45 | 84.60 | 85.40 | 85.18 | 85.53 | 92.98 | 92.60 | 93.32 |
| 5 | 84.62 | 84.49 | 84.65 | 84.86 | 84.58 | 84.96 | 92.22 | 91.71 | 92.50 |
| 6 | 84.72 | 84.61 | 84.75 | 84.38 | 84.06 | 84.46 | 91.46 | 90.86 | 91.68 |
| 7 | 84.88 | 84.78 | 84.91 | 83.95 | 83.61 | 84.03 | 90.70 | 90.05 | 90.87 |
| 8 | 85.08 | 85.00 | 85.10 | 83.57 | 83.22 | 83.63 | 89.96 | 89.28 | 90.09 |
| 9 | 85.31 | 85.26 | 85.33 | 83.22 | 82.88 | 83.27 | 89.23 | 88.54 | 89.34 |
| 10 | 85.58 | 85.55 | 85.59 | 82.90 | 82.58 | 82.93 | 88.52 | 87.84 | 88.62 |
| 11 | 85.87 | 85.87 | 85.87 | 82.62 | 82.33 | 82.64 | 87.83 | 87.16 | 87.90 |
| 12 | 86.18 | 86.21 | 86.17 | 82.36 | 82.11 | 82.36 | 87.16 | 86.50 | 87.21 |
| 13 | 86.51 | 86.57 | 86.49 | 82.13 | 81.91 | 82.12 | 86.50 | 85.86 | 86.54 |
| 14 | 86.86 | 86.96 | 86.83 | 81.93 | 81.75 | 81.90 | 85.85 | 85.25 | 85.88 |
| 15 | 87.23 | 87.35 | 87.19 | 81.74 | 81.60 | 81.70 | 85.22 | 84.64 | 85.23 |
| 16 | 87.61 | 87.76 | 87.56 | 81.57 | 81.48 | 81.53 | 84.60 | 84.05 | 84.58 |
| 17 | 88.00 | 88.19 | 87.95 | 81.42 | 81.37 | 81.36 | 84.00 | 83.48 | 83.95 |
| 18 | 88.40 | 88.62 | 88.34 | 81.28 | 81.27 | 81.20 | 83.40 | 82.91 | 83.34 |
| 19 | 88.81 | 89.06 | 88.75 | 81.15 | 81.19 | 81.05 | 82.81 | 82.35 | 82.72 |
| 20 | 89.23 | 89.51 | 89.17 | 81.04 | 81.12 | 80.93 | 82.23 | 81.80 | 82.13 |

[a]Measured in meters. A = actual heads, ML = maximum likelihood head estimators, and LR = linear regression head estimators.

mators. However, unlike the ML method, LR estimation does not yield estimates directly of $T$ and $S$ but of elements of matrices $\psi$ and $\Gamma$ only. The ML estimation algorithm was implemented in a DEC-VAX 11/780 minicomputer with a CPU time of approximately 30 s.

## 5. SUMMARY AND CONCLUSIONS

The equation of flow for confined aquifers has been discretized and expressed in linear form. The corresponding negative log-likelihood function of the linearized flow equation has been derived and expressed as a nonlinear function of unknown transmissivities and storativities. A global optimum of the negative log-likelihood function has been obtained by Newton's method. A sensitivity analysis with respect to the initial starting search point was conducted, leading in all cases to the same estimates. Upon computation of transmissivity and storativity estimates, matrices governing the flow equation were constructed and head values simulated. Head values used for implementing the maximum likelihood approach were generated by exact solution to a one-dimensional flow problem (and noise-corrupted with Gaussian white noise), in which the confined aquifer is subject to time-varying boundary head values and pumping. Experiences of this study point out some interesting conclusions: (1) Maximum likelihood estimation leads to a nonlinear estimation problem. Unless good prior information on parameter values (to start Newton's method) exist, convergence may be reached at undesirable estimate values. (2) If initial estimators are adequate, convergence is fast (in fact, quadratic in Newton's method) to a reasonably accurate global optimum, when the negative log-likelihood function is convex. (3) When ML and LR estimates of aquifer response (i.e., piezometric heads) were compared, both methods produced predicted values of similar accuracy (within 0.5 m from exact heads). (4) The choice between linear (e.g., LR) and nonlinear (e.g., ML) estimation presents analysts with a variety of trade-offs. The main advantage of LR estimators is easy implementation. In contrast, ML yields directly estimates for $T$ and $S$, which also have desirable asymptotic properties such as consistency and efficiency.

## ACKNOWLEDGMENT

Research loading to this report was supported by the University of California, Water Resources Center, as part of Water Resources Center Project UCAL-WRC-W634. We thank journal reviewers for their valuable suggestions.

## APPENDIX A: MATRIX DERIVATIVE EXPRESSIONS

In implementation of Newton's method, expressions $\partial\hat{\Sigma}/\partial\theta_i$, $\partial^2\hat{\Sigma}/\partial\theta_i^2$, and $\partial^2\hat{\Sigma}/\partial\theta_2\partial\theta_1$ (in which $\theta_1 = T$ and $\theta_2 = S$) are required. Let

$$A_1 = \sum_{t=1}^{n} \varphi_t \varphi_t^T \tag{A1}$$

$$A_2 = \sum_{t=1}^{n} \varphi_t \mathbf{x}_t^T \tag{A2}$$

$$A_3 = \sum_{t=1}^{n} \mathbf{x}_t \mathbf{x}_t^T \tag{A3}$$

By using the expression for $\hat{\Sigma}$ (see eq. 12) and the product rule for matrix derivatives (i.e., for arbitrary matrices $A(\theta_i)$ and $B(\theta_i)$, $\partial AB/\partial\theta_i = A(\partial B/\partial\theta_i) + (\partial A/\partial\theta_i)B$), the following expressions are derived readily

$$\frac{\partial\hat{\Sigma}}{\partial\theta_i} = \frac{1}{n}\left[\sum_{t=1}^{n} (\psi_{\theta_i}A_1\psi^T + \psi A_1\psi_{\theta_i}^T + \psi_{\theta_i}A_2\Gamma^T + \psi A_2\Gamma_{\theta_i}^T + \Gamma_{\theta_i}A_2^T\psi^T + \Gamma A_2^T\psi_{\theta_i}^T + \Gamma_{\theta_i}A_3\Gamma^T + \Gamma A_3\Gamma_{\theta_i}^T)\right] \tag{A4}$$

$$\frac{\partial^2\hat{\Sigma}}{\partial\theta_i^2} = \frac{2}{n}\left[\sum_{t=1}^{n} (\psi_{\theta_i}A_1\psi_{\theta_i}^T + \psi_{\theta_i}A_2\Gamma_{\theta_i}^T + \Gamma_{\theta_i}A_2^T\psi_{\theta_i}^T + \Gamma_{\theta_i}A_3\Gamma_{\theta_i}^T)\right] \tag{A5}$$

$$\frac{\partial^2\Sigma}{\partial\theta_2\partial\theta_1} = \frac{1}{n}\left[\sum_{t=1}^{n} (\psi_{\theta_1}A_1\psi_{\theta_2}^T + \psi_{\theta_2}A_1\psi_{\theta_1}^T + \psi_{\theta_1}A_2\Gamma_{\theta_2}^T + \psi_{\theta_2}A_2\Gamma_{\theta_1}^T + \Gamma_{\theta_1}A_2^T\psi_{\theta_2}^T + \Gamma_{\theta_2}A_2^T\psi_{\theta_1}^T + \Gamma_{\theta_1}A_3\Gamma_{\theta_2}^T + \Gamma_{\theta_2}A_3\Gamma_{\theta_1}^T)\right] \tag{A6}$$

in which $\psi_{\theta_i} = \partial\psi/\partial\theta_i$, $i = 1, 2$, and similarly for the $\Gamma$ matrix. Equations A4–A6 are used when evaluating eqs. 43–45.

## APPENDIX B: DERIVATION OF LINEAR REGRESSION ESTIMATOR $\hat{\Pi}$

Forecasting heads via eq. 47 using linear estimation requires estimate $\Pi$ by $\hat{\Pi}$ using multivariate linear regression. Equation 2 can be written for all time indexes $t = 1, 2, \ldots, n$ at once, as

$$\psi[\varphi_1, \varphi_2, \ldots, \varphi_n] + \Gamma[\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n] = [\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n] \tag{B1}$$

or in compact form

$$\psi\Phi + \Gamma X = U \tag{B2}$$

Solving for $\Phi$ in eq. B2 yields

$$\begin{aligned}\Phi &= -\psi^{-1}\Gamma X + \psi^{-1}U \\ &= \Pi X + V\end{aligned} \tag{B3}$$

From eq. B3 it follows immediately that the multivariate least-squares (linear) estimator for $\Pi$ is

$$\hat{\Pi} = (X^T X)^{-1} X^T \Phi \tag{B4}$$

Heads forecasts (eq. 47) are straightforwardly computed by using $\hat{\Pi}$ of eq. B4 in eq. 47.

## REFERENCES

Aboufirassi, M. and Mariño, M. A., 1984, A geostatically based approach to the identification of aquifer transmissivities in Yolo Basin, California: Math. Geol., v. 16, p. 125–137.

Bickel, P. J. and K. A. Doksum, 1977, Mathematical statistics: Holden-Day, San Francisco, 492 p.

Cooley, R. L., 1982, Incorporation of prior information on parameters into nonlinear regression groundwater flow models, 1, Theory: Water Resour. Res., v. 18, p. 965–976.

Graybill, F. A., 1983, Matrices with applications in statistics: 2nd ed., Wadsworth International Group, Belmont, Calif., 460 p.

Kitanidis, P. K. and E. G. Vomvoris, 1983, A geostatistical approach to the inverse problem in groundwater modeling (steady state) and one dimensional simulations: Water Resour. Res., v. 19, p. 677–690.

Lehmann, E. L., 1983, Theory of point estimation: Wiley, New York, 506 p.

Loaiciga, H. A. and M. A. Mariño, 1986, The inverse problem in confined aquifer flow: Identification and estimation with extensions: Water Resour. Res., submitted.

Mariño, M. A. and J. N. Luthin, 1982, Seepage and groundwater: Developments in Water Science, 13: Elsevier, Amsterdam/Oxford/New York, 499 p.

McLaughlin, D., 1975, Investigation of alternative procedures for estimating groundwater basin parameters: Report prepared for the Office of Water Research and Technology, Water Resources Engineers, Walnut Creek, California, 130 p.

Neuman, S. P., 1980, A statistical approach to the inverse problem of aquifer hydrology, 3, Improved solution method and added perspective: Water Resour. Res., v. 16, p. 331–346.

Rao, C. R., 1965, Linear statistical inference and its applications: Wiley, New York, 625 p.

Yeh, W. W-G., Yoon, Y. S., and Lee, K. S., 1983, Aquifer parameter identification with kriging and optimum parametrization: Water Resour. Res., v. 19, p. 225–233.