Title
SIMULTANEOUS EQUATION SYSTEMS: A CONSISTENT ESTIMATOR FOR UNKNOWN PARAMETERS IN CONFINED AQUIFERS1

Journal
JAWRA Journal of the American Water Resources Association, 23(4)

ISSN
1093-474X

Authors
Loaiciga, Hugo A
Mariño, Miguel A

Publication Date
1987-08-01

DOI
10.1111/j.1752-1688.1987.tb00831.x

Peer reviewed

WATER RESOURCES BULLETIN
VOL. 23, NO. 4 AMERICAN WATER RESOURCES ASSOCIATION AUGUST 1987

SIMULTANEOUS EQUATION SYSTEMS: A CONSISTENT ESTIMATOR FOR UNKNOWN PARAMETERS IN CONFINED AQUIFERS[1]

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

ABSTRACT: This paper presents criteria for establishing the identification status of the inverse problem for confined aquifer flow. Three linear estimation methods (ordinary least squares, two-stage least squares, and three-stage least squares) and one nonlinear method (maximum likelihood) are used to estimate the matrices of parameters embedded in the partial differential equation characterizing confined flow. Computational experience indicates several advantages of maximum likelihood over the linear methods.
(KEY TERMS: parameter estimation; ordinary least squares; two- and three-stage least squares; maximum likelihood; inverse problem; finite-element method.)

INTRODUCTION

The estimation of transmissivities and storativities by statistical and other numerical techniques has received substantial attention in recent years. The estimation of ground-water parameters has been approached in a variety of ways, ranging from generalized least squares to nonlinear programming (see, e.g., Yeh, 1975; Cooley, 1977, 1979, 1982; Neuman and Yakowitz, 1979; Yakowitz and Duckstein, 1980; Neuman, 1980; Yeh and Yoon, 1981; Yeh, *et al.,* 1983; Kitanidis and Vomvoris, 1983; Aboufirassi and Mariño, 1984; and Sadeghipour and Yeh, 1984). Previous studies on the inverse problem in ground-water flow point out the difficulties stemming from the nonuniqueness and instability of parameter estimates (see, e.g., Yakowitz and Duckstein, 1980).

The objectives of this research are: (1) to develop criteria to establish the identifiability status of the inverse problem for confined ground-water flow; (2) to present three linear methods to estimate the coefficient matrices that govern the (discretized) ground-water flow equation; (3) to develop a nonlinear (i.e., maximum likelihood, ML) method to estimate directly transmissivities and storativities, as well as the coefficient matrices that govern the flow equation, and compare the nonlinear method performance with those of the linear estimation techniques; and (4) to establish the statistical properties of the linear and nonlinear estimation methods.

PROBLEM STATEMENT

Discretization of the Confined Aquifer Flow Equation

The equation of flow in a heterogeneous and isotropic confined aquifer is

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

in which ϕ denotes the piezometric head (units, L); $T = T(x,y)$ denotes transmissivity (units, L^2T^{-1}); $S = S(x,y)$ represents storativity (dimensionless); and F denotes either a distributed (units, LT^{-1}) or point (units, L^3T^{-1}) sink/source. Equation (1) can be discretized by the finite element method and expressed as a linear first-order system of differential equations

$$\underset{G\times G}{A}\;\underset{G\times 1}{\underline{\phi}} + \underset{G\times G}{B}\;\underset{G\times 1}{\dot{\underline{\phi}}} + \underset{G\times 1}{\underline{F}} = \underset{G\times 1}{\underline{0}} \tag{2}$$

in which A is the conductance matrix; $\underline{\phi}$ is the vector of unknown heads in the flow domain (the dot indicates time derivative); B is the capacity matrix; and $\underline{F}$ is a vector containing boundary condition values as well as terms related to the sink/source distribution. In Equation (2), only the subset of unknown nodal heads is represented: known (i.e., prescribed) nodal heads in the flow domain have been condensed in typical fashion by the finite-element procedure. Equation (2) can be rewritten as

$$\begin{aligned}\dot{\underline{\phi}} &= -B^{-1}A\underline{\phi} - B^{-1}\underline{F}\\ &= C\underline{\phi} + D\underline{F}\end{aligned} \tag{3}$$

By letting the initial condition $\underline{\phi}(0) = \underline{\phi}_0$, the system of differential equations in Equation (3) has an analytical solution for $\phi(t)$, which is given by

[1]Paper No. 86045 of the *Water Resources Bulletin.* **Discussions are open until April 1, 1988.**

[2]Respectively, Assistant Professor of Geological Sciences, Department of Geological Sciences, Wright State University, Dayton, Ohio 45435; and Professor of Civil Engineering and Water Science, University of California, 119 Veihmeyer Hall (LAWR), Davis, California 95616.

$$\underline{\phi}(t) = e^{tC}[\underline{\phi}_0 + \int_0^t e^{-\tau C}\, D\underline{F}(\tau)d\tau] \tag{4}$$

(see, e.g., Polak and Wong, 1970, pp. 9-17). The matrix e^{tC} in Equation (4) can be expressed as a function of t by the method of interpolating polynomials (see, e.g., Gantmacher, 1959).

The continuous system of Equation (4) is discretized as follows,

$$(\frac{B}{\Delta t} + \eta A)\,\underline{\phi}_t + [A(1-\eta) - \frac{B}{\Delta t}]\,\underline{\phi}_{t-1} + \underline{F}_t = \underline{u}_t \tag{5}$$

$$t = 1, 2, \ldots, n$$

in which $0 \leqslant \eta \leqslant 1$; Δt is a time step; n is the number of simulation periods; $\underline{F}_t$ includes the effect of boundary conditions and sink/sources; and $\underline{u}_t$ is a white noise vector that accounts for errors in measurements and modeling by Equation (1). It is assumed that

$$E(\underline{u}_t) = \underline{0} \tag{6}$$

$$E(\underline{u}_t\underline{u}_s^T) = \begin{cases} [\sigma_{ij}]_{G\times G} = \Sigma \text{ if } t = s \\ 0 \qquad\qquad \text{otherwise} \end{cases} \tag{7}$$

System-Equation Notation

Equation (5) can be rewritten as

$$\underset{G\times G}{\psi}\ \underset{G\times 1}{\underline{\phi}_t} + \underset{G\times K}{\Gamma}\ \underset{K\times 1}{\underline{x}_t} = \underset{G\times 1}{\underline{u}_t},\ \forall t \tag{8}$$

in which

$$\psi = \frac{B}{\Delta t} + \eta A \tag{9}$$

$$\Gamma = [A - \psi \ \vdots\ M] \tag{10}$$

$$\underline{x}_t^T = (\phi_{t-1}^T, \underline{p}_t^T) \tag{11}$$

where the G × (K–G) matrix M and the K–G) × 1 vector $\underline{p}_t$ are determined by the nature of the boundary conditions and the sink/source distribution over the flow domain.

The aquifer shown in Figure 1 is used as an illustrative case for the applicability of Equation (8). By using linear basis interpolation functions in the finite element discretization of Equation (1), one obtains

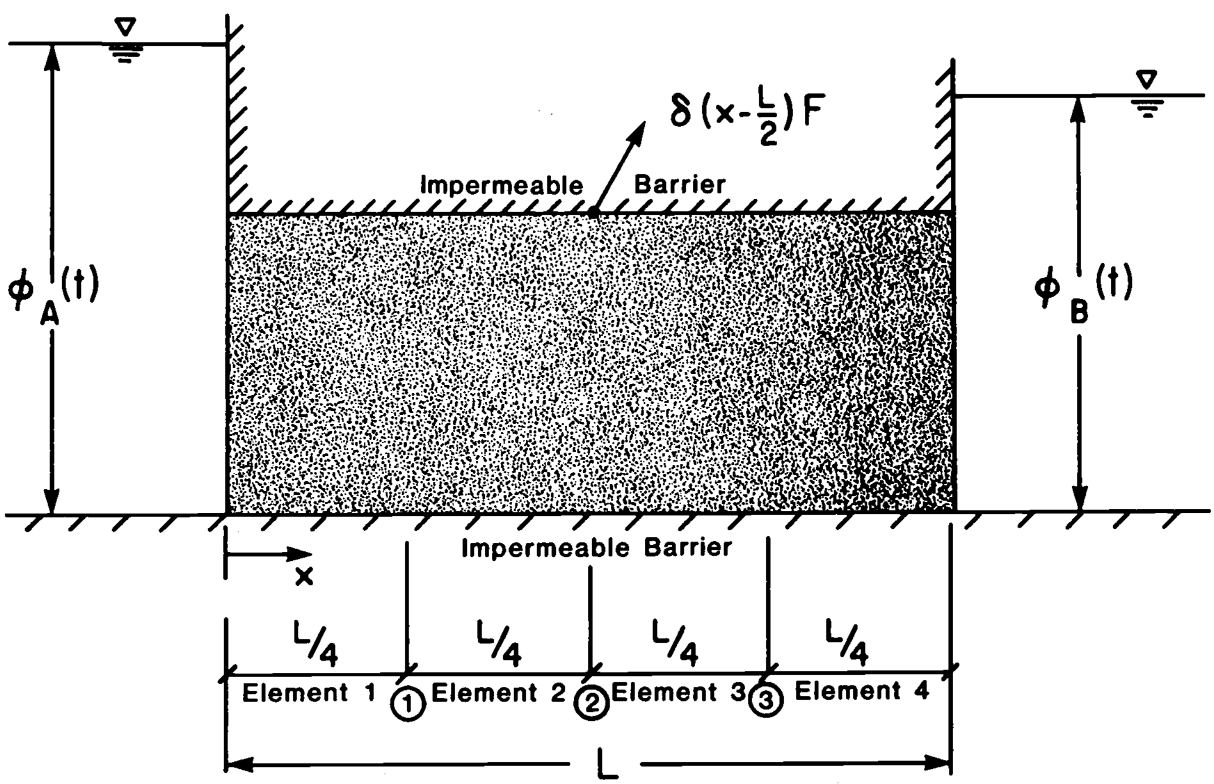

Figure 1. Confined Aquifer Subject to Time-Dependent Boundary Conditions and a Discharge (of units $L^3T^{-1}L^{-1}$) at x = L/2.

$$\psi = [\psi_{ij}] = \begin{bmatrix} \eta\left(\frac{T^{(1)}+T^{(2)}}{\ell}\right)+\frac{1}{3}\left(\frac{S^{(1)}+S^{(2)}}{\Delta t}\right)\ell & -\eta\frac{T^{(2)}}{\ell}+\frac{S^{(2)}}{6\Delta t}\ell & 0 \\ & \eta\left(\frac{T^{(2)}+T^{(3)}}{\ell}\right)+\frac{1}{3}\left(\frac{S^{(2)}+S^{(3)}}{\Delta t}\right)\ell & -\frac{\eta}{\ell}T^{(3)}+\frac{S^{(3)}\ell}{6\Delta t} \\ \text{Symmetric} & & \eta\left(\frac{T^{(3)}+T^{(4)}}{\ell}\right)+\frac{1}{3}\left(\frac{S^{(3)}+S^{(4)}}{\Delta t}\right)\ell \end{bmatrix} \tag{12}$$

$$\Gamma = [\gamma_{ij}] = \left[\begin{array}{ccc:ccccc} \eta'\left(\frac{T^{(1)}+T^{(2)}}{\ell}\right)-\frac{\ell}{3\Delta t}(S^{(1)}+S^{(2)}) & -\eta'\frac{T^{(2)}}{\ell}-\frac{\ell}{6\Delta t}S^{(2)} & 0 & -\frac{T^{(1)}}{\ell} & 0 & \frac{\ell S^{(1)}}{6} & 0 & 0 \\ & \eta'\left(\frac{T^{(2)}+T^{(3)}}{\ell}\right)-\frac{\ell}{3\Delta t}(S^{(2)}+S^{(3)}) & -\frac{\eta'}{\ell}T^{(3)}-\frac{\ell}{6\Delta t}S^{(3)} & 0 & 0 & 0 & 0 & 1 \\ \text{Symmetric} & & \eta'\left(\frac{T^{(3)}+T^{(4)}}{\ell}\right)-\frac{\ell}{3\Delta t}(S^{(3)}+S^{(4)}) & 0 & -\frac{T^{(4)}}{\ell} & 0 & \frac{\ell S^{(4)}}{6} & 0 \end{array}\right] \tag{13}$$

in which $\eta' = 1 - \eta$; $T^{(i)}$ and $S^{(i)}$ denote transmissivity and storativity of the ith element, respectively; and $\ell = L/4$, where L is the total aquifer length.

$$\underline{x}_t^T = [\phi_1(t-1), \phi_2(t-1), \phi_3(t-1); \overline{\phi}_A, \overline{\phi}_B, \overline{\dot{\phi}}_A, \overline{\dot{\phi}}_B, \overline{F}] \tag{14}$$

in which $\overline{F} = \eta F(t) + (1-\eta)F(t-1)$ is the averaged sink strength (see Figure 1); $\overline{\phi}_A = \eta\phi_A(t) + (1-\eta)\phi_A(t-1)$; $\overline{\dot{\phi}}_A = [\phi_A(t) - \phi_A(t-1)]\ /\Delta t$; $\overline{\phi}_B$ and $\overline{\dot{\phi}}_B$ are defined similarly. From Equations (13) and (14) it can be seen that K = 8 and G = 3. The dependence of the elements of ψ and Γ (i.e., ψ_{ij} and γ_{ij}, respectively) on $T^{(i)}$ and $S^{(i)}$ (i = 1, 2, 3, 4) is evident from Equations (12) and (13).

The purpose of this study is to obtain statistically consistent (see, e.g., Rao, 1965, pp. 344-345) estimators of the elements of the matrices ψ and Γ which govern the equation of flow and of the covariance matrix B (see Equation 7). Such task is to be done via linear and nonlinear estimation. The identifiability status of the inverse problem, as well as the statistical properties of estimators are to be addressed also, and Equation (8) constitutes the basic relationship for the analysis.

IDENTIFICATION STATUS OF THE INVERSE PROBLEM

Identification Criteria

The problem of identification is one of being able to determine all the (nonzero) elements of the matrices ψ and Γ, as well as the covariance matrix Σ. Notice that we look for estimators of ψ_{ij}, γ_{ij}, and σ_{ij} (i.e., the elements of the matrices ψ, Γ, and Σ, respectively).

By premultiplying Equation (8) by ψ^{-1} and solving for $\underline{\phi}_t$ one obtains

$$\underline{\phi}_t = (-\psi^{-1}\Gamma)\underline{x}_t + (\psi^{-1}\underline{u}_t)$$

$$= \Pi\,\underline{x}_t + \underline{e}_t,\ t = 1, 2, \ldots, n \tag{15}$$

in which $E(\underline{e}_t\underline{e}_t^T) = \psi^{-1}\Sigma(\psi^{-1})^T$. As is shown subsequently, the G x K (full and unknown) matrix Π can be estimated consistently by ordinary least squares (OLS). The identification problem can then be stated as follows: "given a consistent estimator of Π (= $-\psi^{-1}\Gamma$), is it possible to estimate (consistently) ψ and Γ, and if so, what are the properties of estimators?" It is important to point out that if ψ and Γ can be estimated, then immediately one has all the information required to simulate the discretized ground-water flow Equation (8). The estimated covariance Σ is useful to determine the properties of the estimators for ψ and Γ.

From Equation (15)

$$\underset{(G\times K)}{\Pi} = -\underset{(G\times G)}{\psi^{-1}} \underset{(G\times K)}{\Gamma} \qquad (16)$$

Thus,

$$\psi\Pi = -\Gamma \qquad (17)$$

The jth row (j=1, 2, . . . ,G) of Equation (17) can be written as

$$(\psi_{j1}\ \psi_{j2} \cdots \psi_{jG})\Pi = -(\gamma_{j1}\ \gamma_{j2} \cdots \gamma_{jK}) \qquad (18)$$

The elements of the right- and left-hand side vectors of Equation (18) can be rearranged so that their nonzero elements lead those that are equal to zero. The matrix Π can be conformally rearranged so that Equation (18) can be rewritten as

$$(\underline{\psi}_\Delta^T\ \underline{0}_{\Delta\Delta}^T) \begin{bmatrix} \Pi_{\Delta *} & \Pi_{\Delta **} \\ \Pi_{\Delta\Delta *} & \Pi_{\Delta\Delta **} \end{bmatrix} = -(\underline{\gamma}_*^T\ \underline{0}_{**}^T) \qquad (19)$$

in which $\underline{\psi}_\Delta$ and $\underline{0}_{\Delta\Delta}$ are the $G^\Delta \times 1$ and $(G-G^\Delta) \times 1$ subvectors of nonzero and zero elements in the left-hand side vector of Equation (19); and $\underline{\gamma}_*$ and $\underline{0}_{**}$ are the $K^* \times 1$ and $(K-K^*) \times 1$ subvectors of nonzero and zero elements of the right-hand side vector of Equation (19). The submatrices $\Pi_{\Delta *}$, $\Pi_{\Delta **}$, $\Pi_{\Delta\Delta *}$, and $\Pi_{\Delta\Delta **}$ are of dimensions $G^\Delta \times K^*$, $G^\Delta \times (K-K^*)$, $(G-G^\Delta) \times K^*$, and $(G-G^\Delta) \times (K-K^*)$, respectively, and correspond to a conformally rearranged matrix Π as required by the vector partition in Equation (19). Equation (19) leads to the following expressions,

$$\underset{(1\times G^\Delta)}{\underline{\psi}_\Delta^T} \underset{(G^\Delta\times K^*)}{\Pi_{\Delta *}} = \underset{(1\times K^*)}{-\underline{\gamma}_*^T} \qquad (20)$$

$$\underset{(1\times G^\Delta)}{\underline{\psi}_\Delta^T} \underset{G^\Delta\times(K-K^*)}{\Pi_{\Delta **}} = \underset{1\times(K-K^*)}{\underline{0}_{**}^T} \qquad (21)$$

In Equations (20) and (21), it is possible to divide both sides by any of the (nonzero) elements of $\underline{\psi}_\Delta$, so that one of the elements of $\underline{\psi}_\Delta$ can be normalized to unity. Then, there are $(G^\Delta-1) + K^*$ unknown variables in the jth equation. If Equation (21) could be solved for $\underline{\psi}_\Delta$, then $\underline{\gamma}_*$ would be immediately determined from Equation (20). From basic matrix theory (Graybill, 1983, pp. 149-178), it is known that at least $G^\Delta -1$ equations are needed to solve for $\underline{\psi}_\Delta$ in Equation (21). The vector $\underline{\psi}_\Delta$ has $G^\Delta - 1$ unknown elements since one of its elements can be normalized to unity, as stated above. Therefore, it is required that

$$K - K^* \geqslant G^\Delta - 1 \qquad (22)$$

since there are $K - K^*$ columns in $\Pi_{\Delta **}$.

Equation (22) is only a necessary condition for identifiability, because even if it is satisfied, the columns of $\Pi_{\Delta **}$ may not be linearly independent. A necessary and sufficient condition for the identification of $\underline{\psi}_\Delta$ and $\underline{\gamma}_*$ in the jth equation is that the number of linearly independent columns of $\Pi_{\Delta **}$ be equal to $G^\Delta - 1$, i.e.,

$$\mathrm{rank}(\Pi_{\Delta **}) = G^\Delta - 1 \qquad (23)$$

To summarize, the identification status of the jth equation (see Equation 18), j = 1, 2, . . . , G, must belong to one of the possible cases:

(1) If $K - K^* > G^\Delta - 1$ and rank $(\Pi_{\Delta **}) = G^\Delta - 1$, the jth equation is overidentified. There are multiple ways of estimating consistently $\underline{\psi}_\Delta$ and $\underline{\gamma}_*$.

(2) If $K - K^* = G^\Delta -1$ and $\mathrm{rank}(\Pi_{\Delta **}) = G^\Delta - 1$, there exists exact identification. It is possible to solve uniquely and consistently for $\underline{\psi}_\Delta$ and $\underline{\gamma}_*$.

(3) If $K - K^* > G^\Delta - 1$ and $\mathrm{rank}(\Pi_{\Delta **}) < G^\Delta -1$, or if $K - K^* < G^\Delta -1$, the equation is underidentified. In this case, it is not possible to estimate consistently the parameters in $\underline{\psi}_\Delta$ and $\underline{\gamma}_*$. This is equivalent to saying that there are more unknowns than there are (independent) equations to estimate them.

It is known that for well-posed problems the finite element method yields a matrix ψ (see Equation 17) that is at least tridiagonal (i.e., $G^\Delta \geqslant 2$). Notice also that in Equation (21) the rank of $\Pi_{\Delta **}$ must be equal to or less than $G^\Delta - 1$. Otherwise, i.e., if rank $(\Pi_{\Delta **}) = G^\Delta$, $\underline{\psi}_\Delta$ would be a null vector which contradicts the known fact that $\underline{\psi}_\Delta$ has at least two nonzero elements. Therefore, the condition given by Equation (22) together with possibilities (1)–(3) cited above cover all the feasible cases that can be encountered in establishing the identification status for any of the G structural equations (18).

It is shown next that the problem of estimating ψ and Γ in confined aquifer problems is most likely to be overidentified, so that there are different, but all consistent, methods of estimating ψ and Γ.

An Application of the Identifiability Criterion

The rank condition Equation (23) can be more easily tested by using an equivalent expression, i.e.,

$$\operatorname{rank}(\Pi_{\Delta**}) = \operatorname{rank}\Big(\underset{G-1\times G-G^{\Delta}}{\psi_{\Delta\Delta}} \;\vdots\; \underset{G-1\times K-K^{*}}{\Gamma_{**}} \Big) - (G - G^{\Delta})$$

$$= G^{\Delta} - 1 \tag{24}$$

in which $\psi_{\Delta\Delta}$ and Γ_{**} are submatrices of ψ and Γ, respectively, corresponding to the variables omitted from the jth equation but included in all other rows of Equation (17).

The identifiability criteria implied by Equations (22) and (24) is applied to the example aquifer of Figure 1, with ψ and Γ given by Equations (12) and (13). Choosing the first row equation (j = 1), the following expressions are obtained,

$$(\underline{\psi}_{\Delta}^{T}\ \underline{0}_{\Delta\Delta}^{T}) = (\psi_{11}\ \psi_{12}\ 0) \tag{25}$$

$$-(\gamma_{*}^{T}\ \underline{0}_{**}^{T}) = -(\gamma_{11}\ \gamma_{12}\ \gamma_{14}\ \gamma_{16}\ 0\ 0\ 0\ 0) \tag{26}$$

$$\psi = \begin{bmatrix} \psi_{11} & \psi_{12} & \vdots & 0 \\ \psi_{21} & \psi_{22} & \vdots & \psi_{23} \\ 0 & \psi_{32} & \vdots & \psi_{33} \end{bmatrix} \tag{27}$$

(the reordered Γ matrix)

$$\Gamma = \begin{bmatrix} \gamma_{11} & \gamma_{12} & \gamma_{14} & \gamma_{16} & \vdots & 0 & 0 & 0 & 0 \\ \gamma_{21} & \gamma_{22} & 0 & 0 & \vdots & \gamma_{23} & 0 & 0 & 1 \\ 0 & \gamma_{32} & 0 & 0 & \vdots & \gamma_{33} & \gamma_{35} & \gamma_{37} & 0 \end{bmatrix} \tag{28}$$

and thus,

$$(\psi_{\Delta\Delta}\quad \Gamma_{**}) = \begin{bmatrix} \psi_{23} & \vdots & \gamma_{23} & 0 & 0 & 1 \\ \psi_{33} & \vdots & \gamma_{33} & \gamma_{35} & \gamma_{37} & 0 \end{bmatrix} \tag{29}$$

Equations (25)–(29) correspond to the first row (j = 1), and similar expressions for the second (j = 2) and third (j = 3) rows of Equation (17) (for the test aquifer of Figure 1) are obtainable, leading to the identifiability status as summarized in Table 1. The three equations are overidentified, meaning that one can solve for the unknown elements ψ_{ij} and γ_{ij} by different, consistent, methods. For example, for the first row (j = 1), Equation (21) implies that

$$(\psi_{11}\ \psi_{12}) \begin{bmatrix} \Pi_{11} & \Pi_{12} & \Pi_{14} & \Pi_{16} \\ \Pi_{21} & \Pi_{22} & \Pi_{24} & \Pi_{27} \end{bmatrix} = (0\ 0\ 0\ 0) \tag{30}$$

from which it is apparent that there are four equations and two unknowns, ψ_{11} and ψ_{12}. The elements $\Pi_{11}, \ldots, \Pi_{27}$ are estimated consistently by ordinary least-squares (see Equation 38 below); therefore, they are known quantities. By dividing Equation (30) by ψ_{11}, so that the first element is normalized to unity, and solving for $\psi_{12}^{*} = \psi_{12}/\psi_{11}$, one obtains

$$\psi_{12}^{*} = -\frac{\Pi_{21}\,\Pi_{11} + \Pi_{22}\,\Pi_{12} + \Pi_{25}\,\Pi_{15} + \Pi_{27}\,\Pi_{17}}{(\Pi_{21}^{2} + \Pi_{22}^{2} + \Pi_{25}^{2} + \Pi_{27}^{2})} \tag{31}$$

Equation (31) is a consistent estimator for ψ_{12}^{*} (this is a particular case of the general solution $\underline{x}^{*} = (A^{T}A)^{-1}A^{T}\underline{y}$ for overidentified equations $A\underline{x} = \underline{y}$ when $A^{T}A$ is full rank). As will be shown subsequently, there are alternative methods of estimating ψ_{12}^{*} consistently, that are computationally more expedient than using Equation (31), and whose asymptotic properties are easily established. In the sequel, the normalized coefficients $\psi_{ij}^{*} = \psi_{ij}/\psi_{ii}$ and $\gamma_{ij}^{*} = \gamma_{ij}/\psi_{ii}$ will be respectively represented by ψ_{ij} and γ_{ij} to simplify the notation, and from the context it should be obvious whether the raw or normalized coefficients are being used.

For the case of exact identification, i.e., $K - K^{*} = G^{\Delta} - 1$ and $\operatorname{rank}(\Pi_{\Delta**}) = G^{\Delta} - 1$, the parameter estimates are the same (i.e., have the same numerical value) regardless of whether single-equation or system-equation methods (described below) are used in the estimation. This property is of little practical relevance, because, as previously seen in the identification sample, the inverse problem in ground water is overidentified.

When there is a condition of underidentication, one can set $(G^{\Delta} - 1) - r$ ($r = \operatorname{rank}(\Pi_{\Delta**})$) coefficients artibrarily and solve for the remaining ones, which clearly leads to an infinite number of estimates, all of them inconsistent. Underidentification is most likely nonexistent in the inverse problem for ground-water flow.

TABLE 1. Identifiability Status for Test Case.

Equation j	K	K*	G	G^{Δ}	$\mathrm{rank}(\psi_{\Delta\Delta} \vdots \Gamma_{**})$	Condition Order	Condition Rank	Status
1	8	4	3	2	2	$4>1$	1=1	Overidentified
2	8	4	3	3	2	$4>2$	2=2	Overidentified
3	8	4	3	2	2	$4>1$	1=1	Overidentified

NOTES: (1) The order condition checks the inequality Equation (22).
(2) The rank condition checks the equality Equation (24).

THE TWO-STAGE LEAST-SQUARES METHOD (2SLS)

Development of the Method

In this section, the estimation of the elements of the matrices ψ and Γ is done by operating on each of the row (or structural) equations of the discretized flow Equation (8). From the identification example of the previous section it is understood that the row equations are overidentified (the most likely scenario as argued above). It is shown above that the 2SLS method is a single-equation technique, i.e., it operates on a single structural equation (see, e.g., Equation 18), one-at-a-time. Equation (8) may be written for all time periods at once, i.e.,

$$\psi[\underline{\phi}_1 \cdots \underline{\phi}_n] + \Gamma[\underline{x}_1 \cdots \underline{x}_n] = [\underline{u}_1 \cdots \underline{u}_n] \tag{32}$$

The system of Equation (32) contains G equations, each equation corresponding to one of the rows of the matrix ψ (say the jth) times the matrix of the $\underline{\phi}$'s, plus the jth row of Γ times the matrix of the $\underline{x}$'s being equal to the jth row of the right-hand side of Equation (32). By choosing the jth row equation, normalizing the ψ_{jj} parameter to unity (by dividing the entire jth equation by ψ_{jj}, an arbitrary choice), and taking the transpose of the jth equation (so that the parameters are ordered columnwise), one obtains

$$\underline{\phi}_j = \Phi_j \underline{\psi}_j + X_j \underline{\gamma}_j + \underline{u}_j \tag{33}$$

in which

$$\underline{\phi}_j^T = [\phi_j(1), \ldots, \phi_j(n)]_{1 \times n}$$

$$\underline{\psi}_j^T = [-\psi_{j1}, \ldots, -\psi_{j,j-1}, -\psi_{j,j+1}, \ldots, -\psi_{jG^{\Delta}}]_{1 \times (G^{\Delta}-1)}$$

$$\underline{\gamma}_j^T = [-\gamma_{j1}, -\gamma_{j2}, \ldots, \gamma_{jK^*}]_{1 \times K^*}$$

$$\underline{u}_j^T = [u_j(1), \ldots, u_j(n)]_{1 \times n}$$

$$\Phi_j = \begin{bmatrix} \phi_1(1) \ldots \phi_{j-1}(1)\, \phi_{j+1}(1) \ldots \phi_{G^{\Delta}}(1) \\ \phi_1(2) \ldots \phi_{j-1}(2)\, \phi_{j+1}(2) \ldots \phi_{G^{\Delta}}(2) \\ \vdots \quad \vdots \quad \vdots \quad \vdots \\ \phi_1(n) \ldots \phi_{j-1}(n)\, \phi_{j+1}(n) \ldots \phi_{G^{\Delta}}(n) \end{bmatrix}_{n \times (G^{\Delta}-1)}$$

$$X_j = \begin{bmatrix} x_1(1) & x_2(1) \ldots x_{K^*}(1) \\ x_1(2) & x_2(2) \ldots x_{K^*}(2) \\ \vdots & \vdots \quad \vdots \\ x_1(n) & x_2(n) \ldots x_{K^*}(n) \end{bmatrix}_{n \times K^*}$$

In Equation (33), $E(\underline{u}_j\underline{u}_j^T) = \sigma_{jj}I_{nn}$, according to the assumptions given in Equations (6)–(7). Also, notice that matrices Φ_j and X_j contain the variables associated with nonzero coefficients, thus, their respective column dimensions are $G^{\Delta} - 1$ and K* (see Equation 21).

Equation (33) can be written in the usual linear-model form,

$$\underline{\phi}_j = Z_j \underline{\beta}_j + \underline{u}_j \tag{34}$$

where $Z_j = [\Phi_j\ X_j]$ and $\underline{\beta}_j^T = [\underline{\psi}_j^T\ \underline{\gamma}_j^T]$. The regression model of Equation (34) may be solved by the standard ordinary least-squares (OLS) estimator, i.e.,

$$\hat{\underline{\beta}}_j = (Z_j^T Z_j)^{-1} Z_j^T \underline{\phi}_j \tag{35}$$

One inconvenience, however, is that the columns of Z_j in Equation (34) are correlated with the error term $\underline{u}_j$ in the same equation, and such correlation follows from the definition of the matrix Φ_j above. Therefore, the OLS estimator of Equation (35) is inconsistent due to such correlation. It is possible to transform Equation (34) to make the (transformed) matrix Z_j asymptotically uncorrelated with the error term $\underline{u}_j$. By taking the transpose of Equation (32),

$$\begin{bmatrix} \underline{\phi}_1^T \\ \vdots \\ \underline{\phi}_n^T \end{bmatrix} \psi^T = - \begin{bmatrix} \underline{x}_1^T \\ \vdots \\ \underline{x}_n^T \end{bmatrix} \Gamma^T + \begin{bmatrix} \underline{u}_1^T \\ \vdots \\ \underline{u}_n^T \end{bmatrix} \tag{36}$$

or in compact form, after postmultiplying by $(\psi^T)^{-1}$,

$$\underset{(n\times G)}{\Phi} = X[-\Gamma^T(\psi^T)^{-1}] + V = \underset{(n\times K)}{X}\ \underset{(K\times G)}{R} + \underset{(n\times G)}{V} \tag{37}$$

The matrix R in Equation (37) is estimated by the following multivariate regression,

$$\hat{R} = (X^TX)^{-1}X^T \Phi \tag{38}$$

in which $\hat{R}$ denotes an estimator for R. Let R_j be the following submatrix of $\hat{R}$,

$$\hat{R}_j \doteq [\underline{r}_j, \ldots, \underline{r}_{j-1}, \underline{r}_{j+1}, \ldots, \underline{r}_{G^\Delta}] \tag{39}$$

$\hat{R}_j$ equals the matrix $\hat{R}$ with its jth column suppressed. It follows from Equation (37) that

$$\Phi - V = XR \tag{40}$$

Since the matrix of disturbances V is unobservable, one can approximate the left-hand side of Equation (40) by

$$\Phi - \hat{V} = X\hat{R} \tag{41}$$

in which $\hat{V}$ is the matrix of residuals obtained from the multivariate regression in Equation (38), i.e.,

$$\hat{V} = \Phi - X\hat{R} \tag{42}$$

From Equation (41) it is clear that by deleting the jth column,

$$\phi_j - \hat{V}_j = X\hat{R}_j \tag{43}$$

Finally, Equation (33) can be transformed to

$$\begin{aligned} \underline{\phi}_j &= [\Phi_j - \hat{V}_j]\ \underline{\psi}_j + X_j\underline{\gamma}_j + (\underline{u}_j + \hat{V}_j\underline{\psi}_j) \\ &= [X\hat{R}_j \vdots X_j] \begin{bmatrix} \underline{\psi}_j \\ \underline{\gamma}_j \end{bmatrix} + \underline{w}_j \\ &= \hat{Z}_j\underline{\beta}_j + \underline{w}_j \end{aligned} \tag{44}$$

Since $\Phi_j - \hat{V}_j$ and $\underline{u}_j + \hat{V}_j\ \underline{\psi}_j$ (= $\underline{w}_j$) are asymptotically uncorrelated (i.e., the probability limit of $\Phi_j - \hat{V}_j$ converges to $X\ R_j$, which is uncorrelated with $\underline{w}_j$), $\underline{\beta}_j$ can be estimated consistently by OLS in Equation (44). The OLS method applied to Equation (44) yields

$$\underline{\beta}_j = (\hat{Z}_j^T\hat{Z}_j)^{-1}\hat{Z}_j^T\underline{\phi}_j \tag{45}$$

It is shown in the next subsection that $\underline{\tilde{\beta}}_j$ is a consistent estimator of $\underline{\beta}_j$. Notice that the computation of $\tilde{\beta}_j$ involves first the construction of $\hat{Z}_j$ through the regression estimator $X\ \hat{R}_j$ (see Equation 44) and as a second step, the regression of Equation (45) is carried on, hence its name 2SLS.

Asymptotic Properties of the 2SLS Estimator

From classical multivariate theory, the asymptotic covariance of $\tilde{\beta}_j$ is approximated by

$$\tilde{\Omega} = \hat{s}_{jj}(\hat{Z}_j^T\hat{Z}_j)^{-1} \tag{46}$$

in which $\hat{s}_{jj}$ is the estimator of σ_{jj}, i.e.,

$$s_{jj} = \frac{(\underline{\phi}_j - \hat{Z}_j\underline{\tilde{\beta}}_j)^T(\underline{\phi}_j - \hat{Z}_j\tilde{\beta}_j)}{n-(G^\Delta - 1 + K^*)} \tag{47}$$

From the linear dependence of $\underline{\tilde{\beta}}_j$ on $\underline{\phi}_j$, see Equation (45), large sample theory of OLS (see, e.g., Rao, 1965), the asymptotic distribution of $\sqrt{n}\,(\underline{\tilde{\beta}}_j - \underline{\beta}_j)$ is

$$\sqrt{n}\,(\underline{\tilde{\beta}}_j - \underline{\beta}_j) \to N(\underline{0}, \sigma_{jj} \operatorname*{plim}_{n\to\infty}[n^{-1}\hat{Z}_j^T\hat{Z}_j]^{-1}) \tag{48}$$

i.e., it is multivariate normal with zero mean and a limiting (asymptotic) covariance by the covariance in Equation (48). The consistency of $\underline{\tilde{\beta}}$ and of s_{jj} has been established by Theil (1971).

The 2SLS method is applied to each of the j (j=1,2,. . . ,G) structural equations (see, e.g., Equation 33) one-at-a-time, to estimate the entire set of parameters ψ_{ij} and γ_{ij}, $\forall i,j$. Notice that in deriving the 2SLS estimator the coefficient ψ_{jj} was normalized to 1. Clearly, any of the ψ_{ij}'s in the jth structural equation could have been chosen for normalization. In general, the 2SLS estimator $\tilde{\underline{\beta}}_j$ changes as one changes the normalized parameter and, in this sense, the method is not invariant with respect to normalization. However, the statistical properties of the estimators remain the same regardless of the normalization choice.

THE THREE-STATE LEAST SQUARES METHOD (3SLS)

Development of the Method

It is possible to write the G structural equations (see Equation 44) in a single expression, i.e.,

$$\begin{bmatrix} \underline{\phi}_1 \\ \underline{\phi}_2 \\ \vdots \\ \underline{\phi}_G \end{bmatrix} = \begin{bmatrix} \hat{Z}_1 & 0 & \dots 0 \\ 0 & \hat{Z}_2 & \dots 0 \\ \vdots & & \\ 0 & 0 & \dots \hat{Z}_G \end{bmatrix} \begin{bmatrix} \underline{\beta}_1 \\ \underline{\beta}_2 \\ \vdots \\ \underline{\beta}_G \end{bmatrix} + \begin{bmatrix} \underline{w}_1 \\ \underline{w}_2 \\ \\ \underline{w}_G \end{bmatrix}$$

or in compact notation,

$$\underline{\phi} = \hat{Z}\underline{\beta} + \underline{w} \tag{50}$$

in which $\underline{\phi}$ and $\underline{w}$ are both of dimension nG × 1; $\underline{\beta}$ is of dimension $[\sum_{j=1}^{G} (G_j^{\Delta} - 1 + K_j^{*})] \times 1$, in which G_j^{Δ} and K_j^{*} are the number of nonzero parameters in the jth row of the ψ and Γ matrices, respectively (see Equation 19); and $\hat{Z}$ is dimensioned conformally

The vector of coefficients $\underline{\beta}$ in Equation (50) is estimated by the generalized least-squares regression,

$$\bar{\underline{\beta}} = [\hat{Z}^T (\hat{\Sigma}^{-1} \otimes I_{nn})\hat{Z}]^{-1} [\hat{Z}^T (\hat{\Sigma}^{-1} \otimes I_{nn})]\underline{\phi} \tag{51}$$

in which $\hat{\Sigma}^{-1} \otimes I_{nn} = [s^{ij} I_{nn}]_{nG \times nG}$ is the Kronecker product of the inverse of the covariance matrix $\hat{\Sigma}$ and the identity matrix I_{nn}. $\Sigma \otimes I_{nn}$ is the covariance matrix of $\underline{w}$ in Equation (50). The elements of $\hat{\Sigma}^{-1}$ (denoted as s^{ij}) are obtained by inverting $\hat{\Sigma} = [s_{ij}]_{G \times G}$, in which s_{ij} is obtained from the 2SLS method, i.e.,

$$s_{ij} = \frac{(\underline{\phi}_i - \hat{Z}_i \tilde{\underline{\beta}}_i)^T (\underline{\phi}_j - \hat{Z}_j \tilde{\underline{\beta}}_j)}{n - \max[G_i^{\Delta} - 1 + K_i^{*}, G_j^{\Delta - 1 +} K_j]} \tag{52}$$

The main idea behind the point extimation of the 3SLS is a gain in asymptotic efficiency relative to single-equation methods of estimation such as the 2SLS. The implementation of the 3SLS method requires that the number of observations, n, be larger than the number of equations, G, to avoid the singularity of the covariance matrix $\Sigma \otimes I_{nn}$. Since $\hat{\Sigma}$ in Equation (51) is computed based on the 2SLS estimators s_{ij} (see Equation 52), it is required to first obtain 2SLS estimators (a two-step process), and subsequently the generalized least-squares (GLS) estimator of Equation (51). The 3SLS method derives its name from this sequence of steps (2SLS first, GLS second).

Asymptotic Properties of the 3SLS Estimator

From the expression for $\bar{\underline{\beta}}$ in Equation (51), it follows that the asymptotic covariance of $\bar{\underline{\beta}}$ is approximated by

$$\bar{\Omega} = [\hat{Z}^T (\Sigma^{-1} \otimes I_{nn})\hat{Z}]^{-1} \tag{53}$$

From large sample theory for linear least-squares estimators, the asymptotic distribution for

$$\sqrt{n}\ (\bar{\underline{\beta}} - \underline{\beta}) \rightarrow N(\underline{0}, \text{plim}[n^{-1}\hat{Z}^T (\hat{\Sigma}^{-1} \otimes I_{nn})\hat{Z}]^{-1} \tag{54}$$

from which the consistency of $\bar{\underline{\beta}}$ is readily established (Theil, 1971).

MAXIMUM LIKELIHOOD ESTIMATION (MLE)

The Negative Log-Likelihood Function

In this section, a nonlinear method for estimating directly transmissivities and storavities is developed. The unknown parameters $T^{(i)}$'s and $S^{(i)}$'s are stored in a q × 1 vector of parameters $\underline{\theta}$, whose component, for the sake of simplicity, are denoted as θ_i, i=1,2,. . ., q. The maximum likelihood (ML) method is based on the likelihood function associated with Equation (8). Assuming that $\underline{u}_t$, $\forall t$, is a normal white-noise sequence, and that $\underline{x}_t$ is a fixed vector of exogenous variables in Equation (8), the likelihood function for the left-hand side of Equation (8) is given by

$$L = \frac{|\psi|^n}{(2\pi)^{nG/2}} |\Sigma|^{-n/2} \exp\left\{-\frac{1}{2} \sum_{t=1}^{n} (\psi\,\underline{\phi}_t + \Gamma \underline{x}_t)^T \Sigma^{-1} (\psi\,\underline{\phi}_t + \Gamma\,\underline{x}_t)\right\} \tag{55}$$

For estimation purposes, it is convenient to take the natural logarithm of Equation (55), multiply the resulting expression by −1, to obtain the following function

$$f = k + \frac{n}{2} \ln|\Sigma| - n \ln|\psi| + \frac{1}{2} \operatorname{tr}[\Sigma^{-1} (\psi A_1 \psi^T + \psi A_2 \Gamma^T + \Gamma A_2^T \psi^T + \Gamma A_3 \Gamma^T)] \tag{56}$$

in which

$$k = \frac{nG}{2} \ln(2\pi)$$

$$A_1 = \sum_{t=1}^{n} \underline{\phi}_t \underline{\phi}_t^T$$

$$A_2 = \sum_{t=1}^{n} \underline{\phi}_t \underline{x}_t^T$$

$$A_3 = \sum_{t=1}^{n} \underline{x}_t \underline{x}_t^T$$

It is convenient to simplify Equation (56) by differentiating it with respect to Σ and solving for Σ to obtain

$$\hat{\Sigma} = \frac{1}{n} [(\psi A_1 \psi^T + \psi A_2 \Gamma^T + \Gamma A_2^T \psi^T + \Gamma A_3 \Gamma^T)] \tag{57}$$

in which the following matrix derivatives were used,

$$\frac{d \ln|P|}{dP} = (P^T)^{-1} \quad (|P| > 0) \tag{58}$$

and

$$\frac{d \operatorname{tr}(QPM)}{dP} = Q^T M^T \tag{59}$$

in which Q, P, and M are conformable and square matrices. By substituting Equation (57) into Equation (56), one obtains

$$f = c + \frac{n}{2} \ln|\hat{\Sigma}| - n \ln|\psi| \tag{60}$$

in which $c = k + \frac{n^2}{2}$. The negative log-likelihood f depends on the parameter vector $\underline{\theta}$, since the elements of $\hat{\Sigma}$ and ψ are functions of $\underline{\theta}$. The objective is to minimize f with respect to $\underline{\theta}$ to find the ML estimator of transmissivities and storativities.

Minimization of the Negative Log-Likelihood Function

Newton's method has been chosen for the minimization of Equation (60). The method requires an initial estimate of $\underline{\theta}$ (i.e., $\underline{\theta}_0$). At the kth iteration, while the search is at point $\underline{\theta}_k$, Newton's method finds the step vector, $\underline{p}_k$, such that the function f in Equation (60) evaluated at the next search point $\underline{\theta}_k + \underline{p}_k$ is minimized. By taking a second-order Taylor expansion of f about $\underline{\theta}_k$ one obtains

$$f(\underline{\theta}_k + \underline{p}_k) \simeq f(\underline{\theta}_k) + \underline{p}_k^T \nabla \underline{f}_k + \frac{1}{2} \underline{p}_k^T G_k \underline{p}_k \tag{61}$$

in which $\nabla \underline{f}_k$ and G_k are the gradient and Hessian (matrix of second derivatives of f evaluated at $\underline{\theta}_k$. By applying the necessary conditions in Equation (61) to find the minimizing step vector $\underline{p}_k$, it is found that

$$\underline{p}_k = - G_k^{-1} \nabla \underline{f}_k \tag{62}$$

and the next search point is given by

$$\underline{\theta}_{k+1} = \underline{\theta}_k + \alpha_k \underline{p}_k \tag{63}$$

in which α_k is a scale factor ($0 < \alpha_k \leqslant 1$) introduced to avoid "overshooting" in the search for a minimum. In well-behaved functions, $\alpha_k = 1$; otherwise, having found $\underline{p}_k$, $f(\underline{\theta}_k + \alpha_k \underline{p}_k)$ is minimized with respect to α_k to obtain the appropriate scale factor.

The implementation of Newton's method requires the evaluation of the gradient and Hessian of f about $\underline{\theta}_k$. The elements of the gradient of f are given by $\partial f/\partial \theta_i$, i=1, 2, . . . ,q. The elements of the Hessian matrix are obtained from $\frac{\partial^2 f}{\partial \theta_i \partial \theta_j}$, $\forall$i,j. The following matrix results are useful for expressing analytically both the Hessian and the gradient.

$$\frac{\partial \ln|P|}{\partial \theta_i} = \operatorname{tr}[P^{-1} \frac{\partial P}{\partial \theta_i}] \tag{64}$$

$$\frac{\partial P^{-1}}{\partial \theta_i} = - P^{-1} \frac{\partial P}{\partial \theta_i} P^{-1} \tag{65}$$

$$\frac{\partial^2 \ln|P|}{\partial \theta_i^2} = \operatorname{tr}[-P^{-1} \frac{\partial P}{\partial \theta_i} + P^{-1} \frac{\partial^2 P}{\partial \theta_i^2}] \tag{66}$$

in which P is a square nonsingular matrix with $|P| > 0$, and whose elements are functions of the parameter vector $\underline{\theta}$. Clearly, P will be either ψ or Σ when computing $\nabla \underline{f}_k$ and/or G_k.

For well-behaved convex functions, the convergence of Newton's method is quadratic, which makes it an exceptionally attractive algorithm, and Newton's method is often regarded as the standard against which other algorithms are measured.

Properties of the ML Estimators

For reasonably large sample sizes, MLE's have all the desirable properties of estimators, i.e., consistency, asymptotic normality, and efficiency (see, e.g., Bickel and Doksum, 1977; Lehmann, 1983). The variance-covariance matrix of MLE's can be approximated by the inverse of the Hessian matrix (see, e.g., Rao, 1965), and in particular, standard errors of estimators are estimated by taking the square root of the diagonal elements of the inverted Hessian. A remarkable property of the negative log-likelihood function is its convexity for the case of exponential distribution functions (among them the normal distribution).

APPLICATION OF METHODOLOGIES

Estimation of ψ and Γ

The estimation experiment was based on piezometric heads generated by the continuous-time solution Equation (4) which were corrupted with a white noise sequence. Relevant data are given in Table 2. The heads $\underline{\phi}_1, \underline{\phi}_2, \ldots, \underline{\phi}_{20}$ are shown in Table 3. The elements of the matrices ψ and Γ that govern the flow equation (see Equation 15) were estimated by OLS, 2SLS, 3SLS, and maximum likelihood, and are shown in Table 4. Clearly, MLE's show smaller biases for all of the estimated elements than those exhibited by any of the other methods. The standard errors of MLE's are noticeably smaller also, except for the second element of $\underline{\beta}_1$, the fourth and sixth elements of $\underline{\beta}_2$, and the third element of $\underline{\beta}_3$. The relative merits of the OLS, 2SLS, and 3SLS are not as conclusive as those of MLE's. OLS appears to provide a better approximation of the parameters than do 2SLS and 3SLS. In contrast, 2SLS tends to present in general smaller standard errors than OLS and 3SLS, with the latter yielding larger biases and standard errors than the other three estimation methods. Interestingly, 3SLS should in theory provide a gain in efficiency (i.e., reduction in the standard error of estimators) due to its unique joint estimation feature. However, such efficiency gain is effective under the condition of a known covariance Σ, which is not the case in this study. As a consequence of the due estimation of Σ, as well as to the small-sample size nature of the problem, the asymptotic gain of 3SLS does not materialize, and in this study the method trails in performance the more easily implementable OLS and 2SLS, not to mention MLE's.

Figure 2 shows actual heads (see Table 3), simulated heads $\hat{\Phi} = X\hat{R}$ (see Equations 37 and 38) and $\hat{\underline{\phi}}_t = -\psi^{-1}\Gamma\underline{x}_t$, in which ψ and Γ are obtained from the MLE's estimates of ψ and Γ. The MLE of $\Pi = -\psi^{-1}\Gamma$ is shown in Table 5. Notice that X contains the actual lagged variables $\underline{\phi}_0, \underline{\phi}_1, \ldots, \underline{\phi}_{19}$ (follows from Equation 11), and this allows the estimated heads $\hat{\Phi}$ to follow the overall pattern of the actual heads with an underestimation of actual high values, and an overestimation of low heads. In contrast, the MLE's $\hat{\underline{\phi}}_t = \Pi\underline{x}_t$, $\forall t$, are generated recursively, starting with $\hat{\underline{\phi}}_1 = \Pi\underline{x}_0$, followed by $\hat{\underline{\phi}}_2 = \Pi\underline{x}_1$, and so forth. Therefore, the MLE's of heads forecast or predict the expected values ($E(\underline{\phi}_t) = \Pi\underline{x}_t$, $\forall t$) or average levels of the actual heads, as is evident from Figure 2.

TABLE 2. Data for the Example Aquifer.

Element i	Transmissivity T	Storativity S	Length ℓ
1	500 m^2/d	12×10^{-3}	500 m
2	500 m^2/d	12×10^{-3}	500 m
3	500 m^2/d	12×10^{-3}	500 m
4	500 m^2/d	12×10^{-3}	500 m

Matrix A (Equation 2) Matrix B (Equation 2)

$$A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \text{m/d} \qquad B = \begin{bmatrix} 4 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 4 \end{bmatrix} \text{m}$$

Vector $\underline{F}$ (Equation (2)

$$\begin{bmatrix} -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} \phi_1 \\ \phi_5 \\ \dot{\phi}_1 \\ \dot{\phi}_5 \\ F \end{bmatrix}$$

$\phi_A(t) = 80 + t$, m $\qquad \dot{\phi}_A = 1$, m/d

$\phi_B(t) = 100 - t$, m $\qquad \dot{\phi}_B = -1$, m/d

$F = 10$, $m^3/m/d$

Direct Estimates of Aquifer Parameters

The method of maximum likelihood yields direct estimates of transmissivity (T) and storavity (S), which without loss of generality are assumed to be constant in this study. Upon estimation of T and S, the matrices Ψ and Γ are readily computed via Equations (12) and (13), respectively. Newton's method was initialized at the values T= 250 m^2/day and S = 0.007, with the actual (true) values being 500 m^2/day and 0.012 for T and S, respectively (see Table 2). Figure 3 shows a contour plot of the negative log-likelihood function. It is a convex function with a relatively flat surface around the convergence values, T* = 452 m^2/day and S* = 0.0105. Table 6 contains a summary of the convergence path of Newton's method. Points 0-1-2-3-4-5 in Figure 3 show the search path of Newton's method. The method was tested with other initial estimates of T and S and converged (quadratically) to the same unique global optimum in all cases.

TABLE 3. Piezometric Heads (in meters).

Time t	Node 1	Node 2	Node 3
1	84.15	87.69	93.72
2	85.35	84.35	93.69
3	84.98	85.78	93.42
4	84.28	85.53	93.54
5	84.68	85.39	93.16
6	84.86	83.56	92.18
7	85.98	82.80	90.28
8	85.17	83.18	89.92
9	84.31	82.92	89.02
10	85.13	85.24	88.70
11	87.69	79.95	86.39
12	86.09	84.11	87.41
13	86.89	83.18	86.97
14	87.64	80.35	85.13
15	87.22	82.08	86.23
16	88.99	79.27	83.93
17	89.92	80.60	83.20
18	89.98	81.95	82.28
19	88.07	79.53	82.31
20	90.86	80.95	82.09

The covariance matrix of the MLE's for T and S was approximated by

$$\Omega_{MLE} = \begin{bmatrix} 2479 & 0.121 \\ 0.121 & 1.2 \times 10^{-5} \end{bmatrix}$$

which implies that the standard errors for T and S are $\sqrt{2479} \cong 50\ m^2/day$ and $\sqrt{1.2\text{x}10^{-5}} = 0.00346$, respectively. Ω_{MLE} is assumed to be well approximated by the inverse of the Hessian matrix in Newton's method evaluated at the convergence values of T and S (see Rao, 1965).

SUMMARY AND CONCLUSIONS

A methodology for establishing the identification status of the inverse problem in confined aquifer flow has been presented. It was shown that the structural equation of confined flow is overidentified with respect to the parameters forming the elements of the matrices ψ and Γ, which govern the flow equation. As a consequence of overidentification there exists several statistically consistent methods for estimating the elements of ψ and Γ. Three linear estimation techniques (OLS, 2SLS, and 3SLS) and a nonlinear method (MLE) were developed. The latter method yields directly estimates for transmissivities and storavities, whereas the former three methods estimate the elements of ψ and Γ, without solving for T and S.

The theoretical developments and applications of this paper indicate that: (1) there do not exist unique estimators for the elements of the matrices ψ and Γ, due to the overidentification condition; (2) OLS, despite having rather dismal asymptotic properties (i.e., it is inconsistent) can provide easily computable and accurate parameter estimates; (3) 3SLS, even though a system-equation (i.e., joint) method of estimation, may fail to perform better than 2SLS, and even OLS, in terms of the biases and standard error of estimates for small-sample estimation with unknown covariance matrix; and (4) MLE's have proven to have smaller biases and standard errors than OLS, 2SLS, and 3SLS, and Newton's method showed excellent (i.e., quadratic) convergence rates.

ACKNOWLEDGMENTS

The research leading to this report was supported by the University of California, Water Resources Center, as part of Water Resources Center Project UCAL-WRC-W-692.

LITERATURE CITED

Aboufirassi, M. and M. A. Mariño, 1984. A Geostatistically Based Approach to the Identification of Aquifer Transmissivities in Yolo Basin, California. Math Geology 16(2):125-137.

Bickel, P. J. and K. A. Doksum, 1977. Mathematical Statistics: Basic Ideas and Topics. Holden-Day, San Francisco, California.

Cooley, R. L., 1977. A Method of Estimating Parameters and Assessing Reliability for Models of Steady State Groundwater Flow, 1, Theory and Numerical Properties. Water Resources Research 13(2):318-324.

Cooley, R. L., 1979. A Method of Estimating Parameters and Assessing Reliability for Models of Steady State Groundwater Flow, 2, Applications of Statistical Analysis. Water Resources Research 15(3):603-617.

Cooley, R. L., 1982. Incorporation of Prior Information on Parameters Into Nonlinear Regression Groundwater Flow Models, 1, Theory. Water Resources Research 18(4):965-976.

Gantmacher, F. R., 1959. The Theory of Matrices, Vol. I. Chelsea, New York.

Graybill, F. A., 1983. Matrices With Applications in Statistics, 2nd Edition. Wadsworth Int. Group, Belmont, California.

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 Resources Research 19(3):677-690.

Lehmann, E. L., 1983. Theory of Point Estimation. Wiley, New York, New York.

Neuman, S. P., 1980. A Statistical Approach to the Inverse Problem of Aquifer Hydrology, 3, Improved Solution Method and Added Perspective. Water Resources Research 16(2):331-346.

Neuman, S. P. and S. Yakowitz, 1979. A Statistical Approach to the Inverse Problem of Aquifer Hydrology, 1, Theory. Water Resources Research 15(4):845-860.

Polak, E. and E. Wong, 1970. A First Course on Linear Systems. Van Nostrand-Reinhold Co., New York, New York.

TABLE 4. Estimated Parameters and Their Standard Errors (within parentheses).

	β_1				
Method	1	2	3	4	5
OLS	0.00602 (0.15866)	0.38127 (0.26108)	0.28066 (0.12402)	0.33055 (0.12545)	0.05062 (1.18355)
2SLS	0.20951 (0.22599)	0.17321 (0.30082)	0.20167 (0.13421)	0.41625 (0.13806)	0.011953 (1.11600)
3SLS	0.30610 (0.20372)	0.10626 (0.27151)	0.14631 (0.38881)	0.44275 (0.12494)	0.01187 (0.99991)
MLE	−0.09605 (0.10598)	0.58946 (1.28947)	0.30132 (0.10598)	0.20527 (0.00992)	−0.19868 (0.08333)
True Value	−0.1	0.6	0.3	0.2	−0.2

	β_2					
Method	1	2	3	4	5	6
OLS	0.01161 (0.34187)	0.99511 (0.45344)	0.30722 (0.35782)	0.06045 (0.26436)	−0.44081 (0.51220)	0.17631 (1.61547)
2SLS	0.13040 (0.65685)	1.18940 (0.84460)	0.18570 (0.65150)	0.06526 (0.33942)	−0.63059 (0.83776)	0.12877 (2.05370)
3SLS	0.06629 (0.39583)	0.91494 (0.48521)	0.27159 (0.39631)	0.00944 (0.22490)	−0.32169 (0.52021)	0.08448 (0.99452)
MLE	−0.09605 (0.10598)	−0.09605 (0.10598)	0.30132 (0.10598)	0.58946 (1.28947)	0.30132 (0.10598)	−0.22707 (1.39703)
True Value	−0.1	−0.1	0.3	0.6	0.3	−0.2

	β_3				
Method	1	2	3	4	5
OLS	0.33031 (0.04666)	−0.11585 (0.05995)	0.49058 (0.11202)	0.29817 (0.06658)	−0.03576 (0.42951)
2SLS	0.24204 (0.08676)	0.08264 (0.10739)	0.31187 (0.11406)	0.37296 (0.07158)	0.011284 (0.47727)
3SLS	0.24120 (0.12280)	0.06736 (0.15447)	0.34495 (0.16535)	0.35533 (0.10424)	0.02644 (0.01749)
MLE	−0.09605 (0.10598)	0.30132 (0.10598)	0.58946 (1.28947)	0.20527 (0.00992)	−0.19868 (0.08333)
True Value	−0.1	0.3	0.6	0.2	−0.2

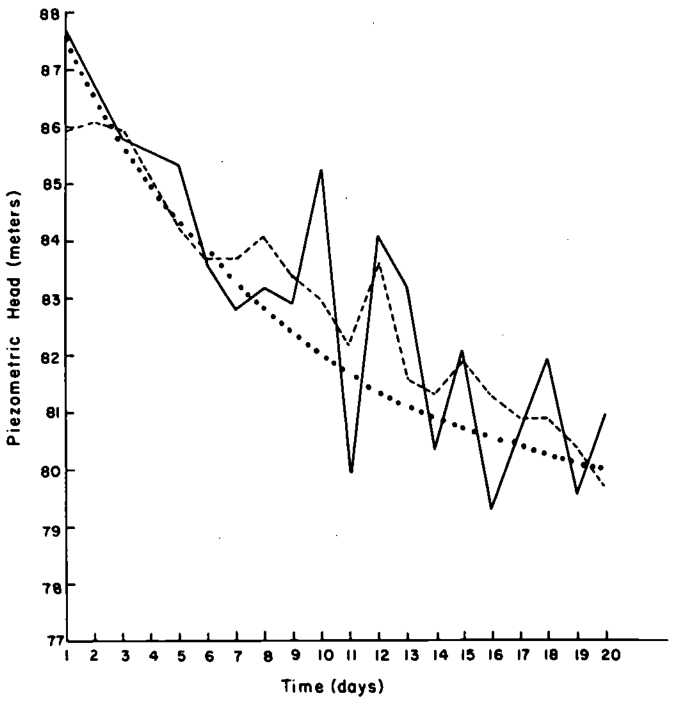

Figure 2. Actual and Simulated Heads (——— actual heads; – – – Φ = XR; . . . $\phi_t = \Pi \underline{x}_t$ (MLE)).

TABLE 5. MLE of Matrix Π.

	G		
K	1	2	3
1	0.56552	0.24930	–0.02395
2	0.24930	0.54157	–0.24930
3	–0.02395	0.24930	0.56552
4	0.20720	–0.02009	0.00193
5	0.00193	–0.02009	0.2072
6	–0.20055	0.01944	–0.00187
7	–0.00187	0.01944	–0.20055
8	0.02222	–0.23133	0.02222

Note: $\Pi = -\psi^{-1}\Gamma$.

TABLE 6. Synopsis of Newton's Search.

Iteration	f	T	S
0	101	250	0.00700
1	94.5	432	0.00787
2	92.5	441	0.00931
3	92.1	449	0.01019
4	91.5	451	0.01045
5	91.0	452	0.01050

f: negative log-likelihood.
T: transmissivity (m^2/day).
S: storativity.
True values: T = 500 m^2/day; S = 0.012.

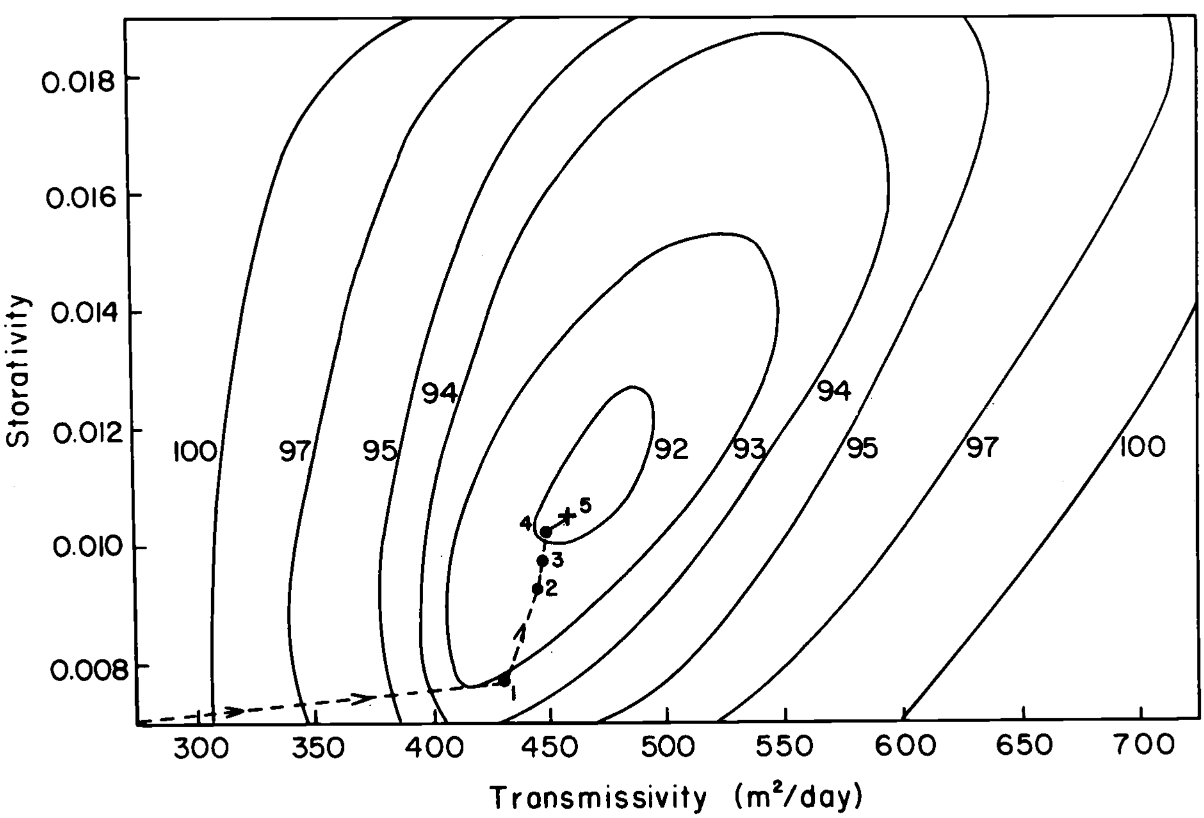

Figure 3. The Negative Log-Likelihood Function and Newton's Search.

Rao, C. R., 1965. Linear Statistical Inference and Its Applications. Wiley, New York, New York.

Sadeghipour, J. and W. W-G. Yeh, 1984. Parameter Identification of Groundwater Aquifer Models: A Generalized Least Squares Approach. Water Resources Research 20(7):971-979.

Theil, H., 1971. Principles of Econometrics. Wiley, New York, New York.

Yakowitz, S. and L. Duckstein, 1980. Instability in Aquifer Identification: Theory and Case Studies. Water Resources Research 16(6): 1045-1064.

Yeh, W. W-G., 1975. Aquifer Parameter Identification. J. Hydraulics Div., Am. Soc. Civil Eng. 101(HY9):1197-1209.

Yeh, W. W-G. and Y. S. Yoon, 1981. Aquifer Parameter Identification with Optimum Dimension in Parametrization. Water Resources Research 17(3):664-672.

Yeh, W. W-G., Y. S. Yoon, and K. S. Lee, 1983. Aquifer Parameter Identification with Kriging and Optimum Parametrization. Water Resources Research 19(1):225-233.

NOTATION

A	$G \times G$ matrix in the continuous-time groundwater flow equation.
B	$G \times G$ matrix in the continuous-time groundwater flow equation.
C	$G \times G$ matrix in the continuous-time groundwater equation; the value of its characteristic roots determine the stability of the continuous-time flow process.
$\underline{e}_t$	$G \times 1$ error term vector in the discretized groundwater equation.
F	pumping rate.
$\underline{F}_t$	$G \times 1$ vector of inputs in the discretized groundwater flow equation at time t.
f	negative log-likelihood function.
G	number of structural equations.
G_k	Hessian matrix
G^{Δ}	number of nonzero structural parameters in matrix ψ of any structural equation.
j	index to denote anyone of the structural equations, $j = 1, 2, \ldots, G$.
K	column dimension of the structural matrix Γ.
K^*	number of nonzero parameters in matrix Γ of any structural equation.
L	likelihood function.
n	number of time periods.
$\underline{p}_k$	step vector in Newton's method.
R	$K \times G$ matrix of parameters in the multivariate regression of heads ($R = \Pi^T$).
$\hat{R}$	$K \times G$ estimator matrix of R.
$\hat{R}_j$	$K \times (G^{\Delta}-1)$ submatrix of R in the jth structural equation.
S	storativity.
$S^{(i)}$	storativity within the ith element.
s_{ij}	estimator of σ_{ij}.
T	transmissivity.
$T^{(i)}$	transmissivity within the ith element.
t	time index, $t = 1, 2, \ldots, n$.
$\underline{u}_t$	$G \times 1$ error vector in the discretized flow equation at time t.
$\underline{u}_j$	$n \times 1$ error term in the jth structural equation.
V	$n \times G$ error matrix in the multivariate regression model for heads.
$\hat{V}$	$n \times G$ estimator for V.
$\underline{w}_j$	$n \times 1$ error term in the transformed jth structural equation of the 2SLS method.
$\underline{w}$	$nG \times 1$ error term in the 3SLS method.
X	$n \times K$ regression matrix in the multivariate regression model for heads.
X_j	$n \times K^*$ regression submatrix in the jth structural equation.
$\underline{x}_t$	$K \times 1$ vector of predetermined variables in the discretized flow equation at time t.
Z_j	$n \times (G^{\Delta} - 1 + K^*)$ regression matrix in the jth structural equation.
$\hat{Z}$	$nG \times (\sum_{j=1}^{G} G_j^{\Delta} - 1 + K_j^*)$ regression matrix in the 3SLS method.
$\hat{Z}_j$	$n \times (G^{\Delta} - 1 + K^*)$ transformed regression matrix in the jth structural equation of the 2SLS method.
α	step length factor in Newton's method.
$\underline{\beta}$	$(\sum_{j=1}^{G} (G_j^{\Delta} - 1 + K_j^*)) \times 1$ parameter vector in the 3SLS method.
$\underline{\beta}_j$	$(G^{\Delta} - 1 + K^*) \times 1$ vector of parameters in the jth structural equation.
$\hat{\underline{\beta}}_j$	$(G^{\Delta} - 1 + K^*) \times 1$ estimator of $\underline{\beta}_j$ in the OLS method.
$\tilde{\underline{\beta}}_j$	$(G^{\Delta} - 1 + K^*) \times 1$ estimator of $\underline{\beta}_j$ in the 2SLS method.
$\bar{\underline{\beta}}$	$(\sum_{j=1}^{G} (G^{\Delta} - 1 + K_j^*)) \times 1$ estimator of $\underline{\beta}$ in the 3SLS method.
Γ	$G \times K$ matrix of structural parameters.
γ_j	$K^* \times 1$ vector of parameters (the jth row of the matrix Γ).
γ_{ij}	the ijth element of Γ.
η	weighting factor in the discrete-time flow equation; $\frac{1}{2} \leqslant \eta \leqslant 1$.
$\underline{\theta}_k$	parameter vector.
Π	$G \times K$ matrix of parameters in the discrete-time flow equation.
Σ	$G \times G$ covariance matrix of $\underline{u}_t$.
σ_{ij}	the ijth element of the covariance matrix Σ.
Φ	$n \times G$ matrix of piezometric heads in the multivariate regression model.
Φ_j	$n \times (G^{\Delta}-1)$ matrix of piezometric heads in the jth structural equation.
$\underline{\phi}_t$	$G \times 1$ vector of piezometric heads at time t.
$\dot{\underline{\phi}}_t$	$G \times 1$ time derivative of vector $\underline{\phi}_t$.

$\underline{\phi}_j$	n × 1 vector of piezometric heads in the jth structural equation.
ψ	G × G matrix of structural parameters.
ψ_{ij}	the ijth element of ψ.
$\underline{\psi}_j$	$(G_j^{\Delta} - 1) \times 1$ vector of parameters (the jth row of the matrix ψ).
$\tilde{\Omega}$	covariance matrix of 2SLS estimator.
$\bar{\Omega}$	covariance matrix of 3SLS estimator.
Ω_{MLE}	covariance of MLE's.

Acronyms:

ML	maximum likelihood.
OSL	ordinary least squares.
2SLS	two-stage least squares.
3SLS	three-stage least squares.