Advances in Water Resources 15 (1992) 259–269

Parameter identification of nonsteady groundwater flow systems

Jiannan Xiang* & Derek Elsworth

Department of Mineral Engineering, Pennsylvania State University, Pennsylvania 16802, USA

This paper presents several methods which are based on multiple data sets to reduce the errors caused by the noise in the measured data. The comparisons show that the accuracy of inverse solution depends both on the noise level and on the number of consecutive observations. For low noise levels, both average and least squares methods perform well. When the noise level is high, integration methods based on the trapezoidal rule yield better accuracy. However, when noise dominates the record, all methods may yield an unacceptable error. To reduce noise levels, a Butterworth filter is used. Using filtered data, the accuracy of the estimated parameters is improved. A computed example shows that the errors in transmissivity and storage coefficient are different because they have different derivatives. Application of the inverse methods are demonstrated in a two-dimensional problem.

Key words: parameter estimation, inverse problem, noise, filter.

NOTATION

a_j	Coefficient of the numerator polynomial of the filter function
b_j	Coefficient of the denominator polynomial of the filter function
$\mathbf{b}$	Vector formed by all right terms in eqn (19)
B^n	Matrix representing distribution of head difference at step n
E	Matrix comprising submatrices H^i
$\mathbf{g}$	Filtered data
$\mathbf{g}'$	Modified filtered data
h	Head
h_a, h_b	Head at both ends of aquifer in example problems
h_0, h_1, h_2	Specified function
$\mathbf{H}$	A vector of head
H	A matrix representing the distribution of hydraulic head difference within the domain
k	Number of parameters
K^{ij}	Matrix formed by $\mathbf{H}$ and $\mathbf{B}$
L	Filter function
m	Number of data sets
n	Number of nodes
N	The order of digital filter
$\mathbf{q}^i$	A vector formed by Q^i
Q_w, $\mathbf{Q}$, $\mathbf{Q}^i$	Discharge, discharge vector and vector for step i
R	Flow region
dR	Boundary of aquifer
dR_1, dR_2	Specified head boundary and discharge boundary, respectively
$S, \mathbf{S}, \mathbf{S}^i$	Storage coefficient, storage coefficient vector and vector for step i
$\boldsymbol{S}$	A matrix representing the distribution of storage coefficient in the domain
t, t_i	Time and time level i
Δt	Time step
T	Transmissivity
$\mathbf{T}$, $\mathbf{T}^i$	Vector of transmissivity for steady state and for step i
$\boldsymbol{T}$	Matrix representing the geometric transmissivity of the domain
$\bar{T}$	Average vector of several transmissivity vectors
x, y, x_w, y_w	Coordinates and coordinates at wells
α	Phase shift
β	Gain factor
ε	Relative error
ϵ	Error vector of solution or residual of equation.
ϕ, Φ	Shape functions

*Present address: Bureau of Economic Geology, University of Texas, University Station, Box X, Austin, Texas 78713–7508, USA.

Advances in Water Resources 0309-1708/92/S05.00

1 INTRODUCTION

The development of groundwater management and reservoir withdrawal strategies usually require estimation of parameters describing the system or aquifer. For nonsteady flow problems, a summary of previous methods are given in the review by Yeh[12]. Yeh *et al.*[13] used the least squares method for the nonsteady flow problem in a direct approach. Kriging was used as a presampling filter to reconstruct the head distribution. The distinct advantage of this approach over optimization based approaches was that it involved no iterative searching.

Loaiciaga and Marino[5,6] developed a method for estimating the elements of parameter matrices in the governing equation of flow in a confined aquifer. The estimating methods, named respectively the two-stage least squares and three-stage least squares, were applied to a specific groundwater inverse problem and compared with an ordinary least squares estimator. The three-stage estimator provided a closer approximation to the actual parameter values than the two-stage estimator.

Dagan and Rubin[2] developed a stochastic method to identify aquifer natural recharge, storativity and transmissivity under transient conditions. The aquifer was unbounded and a first-order approximation of the flow equations with slowly varied transients was used. Based on these assumptions, the expected values of transmissivity and head were expressed by analytical equations.

Yoon and Yeh[14] used the finite element method for parameter identification in homogeneous and heterogeneous media. Rosen's gradient projection technique[9] was used to handle constraints combined with the Gauss–Newton method. The computed results showed that ideal estimation of the parameters could be obtained by their proposed method when observations were noise free. The estimations deteriorated in the presence of noise. They also noted that as the dimension of the parameter increased, the ill-posedness of the problem became worse. Using time series observations would not greatly alleviate this problem.

Kool and Parker[4] used the Levenberg–Marquardt algorithm for analysis of the inverse problem in transient unsaturated flow. Their analysis showed that despite the sensitivity to the unknown parameters, their distributions in space and time are important in determining parameter identifiability. As long as a correct constitutive model is used and the directly measurable parameters eliminated, a reasonably good prediction of hysteretic hydraulic properties could be obtained.

Ginn *et al.*[3] presented a novel method for parameter estimation where the model is inverted following the spatial discretization but before the finite difference temporal discretization is attempted contrary to practice in the past. The Laplace transform is used to replace the discretization in time for the system of ordinary differential equations. The existence and stability of the solution is apparently dependent upon selection of the discrete values of the Laplace variables in the transformed model.

The objectives of this paper are to: present several methods for parameter identification in nonsteady-state groundwater systems with noisy data; investigate solution error in each of the proposed methods using numerical examples; and demonstrate the usefulness of the proposed methods. In this paper, the finite element method is used to discretize the physical region and interpolate head and transmissivity functions. The finite difference method is used to discretize in time. For simplicity and uniqueness[7] and to provide additional comparison, the direct method is used exlusively in the inverse solution.[10]

2 PROBLEM DESCRIPTION

Forward and inverse solution of the two-dimensional diffusion equation may be completed through finite element discretization in space. This is described in the following.

2.1 Differential equation

Basic assumptions for the one- or two-dimensional problems of groundwater flow considered in the following are that the aquifer is horizontal, inhomogeneous from point to point and continuous in the region. Only confined problems are considered where the transmissivity is directionally isotropic.

The governing equation for the groundwater flow system can be written as:

$$\frac{\partial}{\partial x_i}\left(T\frac{\partial h}{\partial x_i}\right) = \sum Q_w(t)\prod \delta(x_i - x_w) + S\frac{\partial h}{\partial t} \qquad (1)$$

where i can be 1, 2 or 3 for one-, two- and three-dimensional problems, respectively. For a two-dimensional problem, the equation can be specialized as:

$$\frac{\partial}{\partial x}\left(T\frac{\partial h}{\partial x}\right) + \frac{\partial}{\partial y}\left(T\frac{\partial h}{\partial y}\right) = \sum Q_w(t)\delta(x - x_w)\delta(y - y_w) + S\frac{\partial h}{\partial t} \qquad (2)$$

and must satisfy the following conditions:

$$h(x, y, 0) = h_0(x, y) \qquad x, y \text{ in } R$$

$$h(x, y, t) = h_1(x, y, t) \qquad x, y \text{ in } dR_1$$

$$T\frac{\partial h}{\partial n}n_x + T\frac{\partial h}{\partial y}n_y = h_2(x, y, t) \qquad x, y \text{ in } dR_2$$

where $h(x, y, t)$ is head at point (x, y); T is transmissivity; S is storage coefficient; Q_w is source–sink term; x, y are coordinates in two-dimensional space; t is time; R is

the flow region; dR is the boundary of the aquifer ($dR_1 U dR_2 = dR$); and h_0, h_1, h_2 are specified head functions. $\delta(x)$ is the delta function and $\delta(x) = \infty$ for $x = 0$; $\delta(x) = 0$ for $x \neq 0$; n_x and n_y are the components of the unit normal vector; and x_w, y_w are the coordinates of a well.

2.2 Governing equation

Using a low-order finite element,[11] eqn (2) may be readily transformed to a system of linear equations. For steady-state conditions, the equation may be written as:

$$\boldsymbol{T}\mathbf{H} = \mathbf{Q}_w + \mathbf{q} = \mathbf{Q} \tag{3}$$

or

$$\mathbf{H}\boldsymbol{T} = \mathbf{Q} \tag{4}$$

where $\boldsymbol{T}$ is a matrix representing the geometric transmissivity of the domain; $\boldsymbol{H}$ is a matrix representing the hydraulic head distribution within the domain; $\mathbf{H}$ is a vector of nodal head magnitudes; $\mathbf{T}$ is a vector of elemental transmissivity magnitudes; $\mathbf{Q}$ is a vector of discharge; $\mathbf{q}$ is a vector of discharge on the boundary; and $\mathbf{Q}_w$ is a vector of discharge at a nodal point representing a well.

For the nonsteady-state condition, eqn (2) can be written as:

$$\boldsymbol{T}\mathbf{H} = \mathbf{Q} + \boldsymbol{S}\frac{\partial \mathbf{H}}{\partial t} \tag{5}$$

or

$$\boldsymbol{H}\mathbf{T} = \mathbf{Q} + \boldsymbol{S}\frac{\partial \mathbf{H}}{\partial t} \tag{6}$$

where $\boldsymbol{S}$ is a matrix representing the storage coefficient distribution within the domain.

2.3 Forward solution

The adequacy of the transient inverse model is verified against forward solution to the same problem with arbitrarily prescribed distributions of the hydraulic parameters. Data for the inverse solution are determined from a forward solution of an arbitrary homogeneous domain in the nonsteady state using eqn (5). The finite difference method is used to discretize the time derivative. For the initial solution, $\boldsymbol{H}^0$ may be obtained from the initial data $h(x, y, 0) = h_0(x, y)$. Therefore, eqn (5) can be written as:

$$\boldsymbol{T}\mathbf{H}^0 = \mathbf{Q}^0 \tag{7}$$

For subsequent head distributions, the central finite difference method may be applied to eqn (5), resulting in the following equation:

$$\boldsymbol{T}(\mathbf{H}^{i+1} + \mathbf{H}^i) = \mathbf{Q}^{i+1} + \mathbf{Q}^i + 2\boldsymbol{S}(\mathbf{H}^{i+1} - \mathbf{H}^i)/\Delta t \tag{8}$$

Alternatively:

$$\boldsymbol{A}\mathbf{H}^{i+1} = \mathbf{Q}^{i+1} + \mathbf{Q}^i - \boldsymbol{C}\mathbf{H}^i \tag{9}$$

where:

$$\boldsymbol{A} = \boldsymbol{T} - 2\boldsymbol{S}/\Delta t$$
$$\boldsymbol{C} = \boldsymbol{T} + 2\boldsymbol{S}/\Delta t \tag{10}$$

$\mathbf{Q}^i$, $\mathbf{Q}^{i+1}$ are discharge vectors at steps i and $i+1$; and Δt is the time-step magnitude.

By evaluating the matrix of eqn (9) and solving eqn (7), the head distribution at successive time levels may be determined.

3 PARAMETER ESTIMATION FOR NOISE-FREE DATA

Unlike the steady state, nonsteady problems involve release from storage and, if strongly transient, enable transmissivity and storage magnitudes to be determined. Methods presented in the following address the esti-mation of both transmissivity and storage coefficient when noise-free time series data are available.

3.1 Inverse solution for transmissivity

Using the finite difference method to discretize in time, eqn (2) may be written as:

$$\frac{\partial}{\partial x_i}\left(T\frac{\partial h^{i+1/2}}{\partial x_i}\right) = \sum Q_w^{i+1/2}(x_i - x_{iw}) + S\frac{\Delta h^{i+1/2}}{\Delta t} \tag{11}$$

For the two-dimensional case, eqn (11) can be expressed, according to time steps i and $i+1$ that bracket the time level of interest, $i + 1/2$, as:

$$\frac{\partial}{\partial x}\left(T\frac{\partial h^{i+1}}{\partial x}\right) + \frac{\partial}{\partial x}\left(T\frac{\partial h^{i}}{\partial x}\right) + \frac{\partial}{\partial y}\left(T\frac{\partial h^{i+1}}{\partial y}\right) + \frac{\partial}{\partial y}\left(T\frac{\partial h^{i}}{\partial y}\right) = -\sum Q_w^{i+1} - \sum Q_w^{i} + S\frac{(h^{i+1} - h^i)}{\Delta t} \tag{12}$$

Using the finite element method to discretize head at time step j:

$$\mathbf{H}^j = \Sigma\phi h_i^j \tag{13}$$

and transmissivity:

$$\mathbf{T} = \Sigma\Phi T_i \tag{14}$$

where ϕ and Φ are shape functions interpolating head and transmissivity, respectively, and need not be the same. Substituting eqns (13) and (14) into eqn (8), one

obtains:

$$(\boldsymbol{H}^{i+1} + \boldsymbol{H}^{i})\mathbf{T} = \mathbf{Q}^{i+1} + \mathbf{Q}^{i} + 2\boldsymbol{S}(\mathbf{H}^{i+1} - \mathbf{H}^{i})/\Delta t \tag{15}$$

Assuming that the vector of storage coefficients, **S** is known, one can formulate $\boldsymbol{S}$, $\boldsymbol{H}^{i}$ and $\boldsymbol{H}^{i+1}$ and prescribed vectors $\mathbf{Q}^{i}$ and $\mathbf{Q}^{i+1}$ using the measured head and discharge data. Subsequently, transmissivity **T** may be determined directly by solution of this equation.

3.2 Inverse solution to determine additional parameters

If the storage coefficients are unknown, eqn (15) can be rewritten as:

$$(\boldsymbol{H}^{i+1} + \boldsymbol{H}^{i})\mathbf{T} - 2\boldsymbol{B}^{i}\mathbf{S}/\Delta t = \mathbf{Q}^{i+1} + \mathbf{Q}^{i} \tag{16}$$

where $\boldsymbol{B}^{i}$ is a matrix representing distribution of head differences at time n; and **S** is a vector of storage coefficients.

To solve this equation for both **T** and **S**, it is necessary to have one additional data set. Assuming that the system is also monitored at the $i+2$ time step, construction of a similar equation yields:

$$(\boldsymbol{H}^{i+2} + \boldsymbol{H}^{i+1})\mathbf{T} - 2\boldsymbol{B}^{i+1}\mathbf{S}/\Delta t = \mathbf{Q}^{i+2} + \mathbf{Q}^{i+1} \tag{17}$$

By solving eqns (16) and (17) simultaneously, the desired parameters vectors **T** and **S** can be obtained. This procedure may be extended to enable determination of additional parameters such as areal discharge or point discharge. For each additional parameter vector, one more data set must be added. Simultaneously, solution of these equations will yield the required parameters in theory.

4 PARAMETER ESTIMATION FOR NOISY DATA

In reality, most observations contain noise. Consequently, a noisy head distribution measured at different moments in time may result in different parameter estimates when eqns (15)–(17) are used. To minimize the influence of noise on the estimated parameters, the procedures presented in the following may be invoked.

4.1 Solution for transmissivity

Four methods based on time series observations are presented for parameter estimation of transmissivity. They are defined as the average matrix, average parameter, integration and least squares methods.

4.1.1 Average matrix method

As mentioned previously, eqn (15) is based on two independent head distributions representing consecutive time steps separated by the interval Δt. For m data sets, eqn (15) may be simply expanded to the following form:

$$\left(\sum_{i=0}^{m} \boldsymbol{H}^{i}\right)\mathbf{T} = \sum_{i=0}^{m} \mathbf{Q}^{i} + \frac{m}{\Delta t}\boldsymbol{S}(\mathbf{H}^{m} - \mathbf{H}^{0}) \tag{18}$$

for a transmissivity vector **T**. Since this method uses additional data sets, it should give improved results for noisy data.

4.1.2 Average parameter method

According to eqn (15), for all time steps one can formulate a system of equations using any two consecutive data sets as follows:

$$\begin{aligned}
(\boldsymbol{H}^{0} + \boldsymbol{H}^{i})\mathbf{T} &= \mathbf{Q}^{0} + \mathbf{Q}^{1} + \frac{m}{\Delta t}\boldsymbol{S}(\mathbf{H}^{1} - \mathbf{H}^{0}) \\
(\boldsymbol{H}^{1} + \boldsymbol{H}^{2})\mathbf{T} &= \mathbf{Q}^{1} + \mathbf{Q}^{2} + \frac{m}{\Delta t}\boldsymbol{S}(\mathbf{H}^{2} - \mathbf{H}^{1}) \\
(\boldsymbol{H}^{2} + \boldsymbol{H}^{3})\mathbf{T} &= \mathbf{Q}^{2} + \mathbf{Q}^{3} + \frac{m}{\Delta t}\boldsymbol{S}(\mathbf{H}^{3} - \mathbf{H}^{2}) \\
(\boldsymbol{H}^{m-1} + \boldsymbol{H}^{m})\mathbf{T} &= \mathbf{Q}^{m-1} + \mathbf{Q}^{m} + \frac{m}{\Delta t}\boldsymbol{S}(\mathbf{H}^{m} - \mathbf{H}^{m-1})
\end{aligned} \tag{19}$$

By solving this set of equations, m parameter sets may be obtained. The final parameter, transmissivity, can be determined by averaging the recorded transmissivity magnitudes as:

$$\bar{\mathbf{T}} = \frac{\sum_{i=1}^{m} \mathbf{T}^{i}}{m} \tag{20}$$

4.1.3 Integration method

Alternatively, an integration method can be used for noise reduction. Integration of eqn (2) yields:

$$\begin{aligned}
\int_{t_0}^{t_m} \left[\frac{\partial}{\partial x}\left(T\frac{\partial h}{\partial x}\right) + \frac{\partial}{\partial y}\left(T\frac{\partial h}{\partial y}\right)\right] \mathrm{d}t \\
= \int_{t_0}^{t_m} \left[\sum Q_{\mathrm{w}}(t) \prod \delta(x_i - x_{\mathrm{w}})\delta(y_i - y_{\mathrm{w}})\right] \mathrm{d}t \\
+ \int_{t_0}^{t_m} S\frac{\partial h}{\partial t}\,\mathrm{d}t
\end{aligned} \tag{21}$$

where t_0 to t_m represents the time-interval of interest.

To solve this equation for transmissivity, T, various numerical integration methods can be used. Based on the 'composite trapezoidal rule', each term in eqn (21) can be expressed as:

$$\begin{aligned}
&\frac{1}{2}\left[\frac{\partial}{\partial x_j}\left(T\frac{\partial h(t_0)}{\partial x_j}\right) + \frac{\partial}{\partial x_j}\left(T\frac{\partial h(t_m)}{\partial x_j}\right)\right] \\
&+ \sum_{i=1}^{m-1} \frac{\partial}{\partial x_j}\left(T\frac{\partial h(t_i)}{\partial x_j}\right) = \frac{1}{2}\left[Q(t_0) + Q(t_m) + \sum_{i=1}^{m-1} Q(t_i)\right] \\
&\times \prod \delta(x_j - x_{\mathrm{w}k}) + S\frac{h(t_m) - h(t_0)}{\Delta t}
\end{aligned} \tag{22}$$

Since the derivatives of head $\partial h/\partial t$ can only be obtained at the midpoint of each element, the 'composite midpoint rule' for numerical integration may be used.

This transforms the last term of eqn (21) as:

$$\int_0^L f(x)\,dx = [f(x_1) + f(x_2) + f(x_3) + \cdots f(x_m)]\Delta x \tag{23}$$

where x_i are coordinates at the midpoint of element i, Δx is the length of the element, $\phi(x_i)$ are integrands and L is the length of the domain.

Using the finite element method to discretize eqn (22) yields:

$$\left(\frac{1}{2}[\boldsymbol{H}(t_0) + \boldsymbol{H}(t_m)] + \sum_{i=1}^{m-1} \boldsymbol{H}(t_i)\right)\mathbf{T} = \frac{1}{2}[\mathbf{Q}(t_0) + \mathbf{Q}(t_m)] + \sum_{i=1}^{m-1} \mathbf{Q}(t_i) + \frac{\mathbf{S}}{\Delta t}[\mathbf{H}(t_m) - \mathbf{H}(t_0)] \tag{24}$$

Other methods such as those based on the 'composite Simpson's rule' may also be invoked to represent the integrals in eqn (21).

4.1.4 The least squares method

Equation (19) is exact at any time level and may be compactly written as:

$$\boldsymbol{E}\mathbf{T} = \mathbf{b} \tag{25}$$

where $\boldsymbol{E}$ is an $(n+m) * k$ matrix comprising submatrices $\boldsymbol{H}^i$, and k is the number of parameters desired from inverse solution; n and m have the same definitions as before, and $\mathbf{b}$ is an $(n+m)$ vector formed by all terms on the right-hand side of eqn (19). Assuming the residual error in eqn (25) is ϵ, and using the least squares method, the solution of eqn (25) can be written as:

$$\mathbf{T} = (\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{H}^T\mathbf{b} \tag{26}$$

where the error ϵ is minimized as $\partial(\epsilon^T\epsilon)/\partial T = 0$.

4.2 Solution for transmissivity and storage coefficient

In principle, all methods described previously can be used to determine the magnitude of the unknown storage coefficient vector. Where storage distributions are desired, the number of measured data sets must be doubled. For example, eqn (16) can be extended for m sets of observation as:

$$\begin{bmatrix} \boldsymbol{K}^{11}\boldsymbol{K}^{21} \\ \boldsymbol{K}^{12}\boldsymbol{K}^{21} \\ \cdots \\ \boldsymbol{K}^{1m}\boldsymbol{K}^{2m} \end{bmatrix} \begin{Bmatrix} \mathbf{T} \\ \mathbf{S} \end{Bmatrix} = \begin{Bmatrix} \mathbf{q}^1 \\ \mathbf{q}^2 \\ \cdots \\ \mathbf{q}^m \end{Bmatrix} \tag{27}$$

where $\boldsymbol{K}^{1i} = \boldsymbol{H}^{i+1} + \boldsymbol{H}^i$, $\boldsymbol{K}^{2i} = \boldsymbol{B}^i/\Delta t$ and $\mathbf{q}^i = \mathbf{Q}^{i+1} + \mathbf{Q}^i$.

Using the least squares method, as mentioned above, the vectors of transmissivity $\mathbf{T}$ and storage coefficient $\mathbf{S}$ can be evaluated.

4.3 Noise reduction

For noisy data sets, parameter estimation becomes very difficult. To improve results, some curve-fitting methods such as polynomial, exponential or nonlinear interpolation may be used to smooth the data. Simple smoothing methods may actually smear the data and an alternative is to use a digital filter to reduce the noise level. Low pass filters are most appropriate since noise is usually of high frequency. The Butterworth filter is a suitable low pass filter and may be expressed as:

$$L(z) = \frac{\sum_{j=0}^{N} a_j z^{-j}}{1 - \sum_{j=1}^{N} b_j z^{-j}} \tag{28}$$

where N is the order of the linear constant coefficient difference equation or filter; and a_j and b_j are coefficients defined by the order, N, cut-off frequency or input data.

The relationship between the input (noise) data $h(t)$ and the output data $h'(t)$ can be expressed by:

$$h'(t_i) = -\sum_{j=1}^{N} a_j h'(t_{i-j}) + \sum_{j=0}^{N} b_j h(t_{i-j}) \tag{29}$$

It is necessary to choose both a reasonable number to represent N and a suitable cut-off frequency. When the curve representing head variation with time is very variable, the number N should be appropriately increased. For more practical applications, a small number N and a middle range value of cut-off frequency should be initially chosen. After computing a certain error, one can use a trial-and-error method to compare these errors until a suitable cut-off frequency is found which gives the least error.

5 EXAMPLES

Two example problems are considered. The first one is a one-dimensional problem with a linearly distributed transmissivity. All methods identified in Table 1 are tested on this simple problem and the relative accuracy of prediction is determined. The second problem geometry is two-dimensional and is used to illustrate potential applications of the methods and establish the generality of the methods to multi-dimensional domains.

5.1 One-dimensional example

The example problem is one of unidirectional flow in a heterogeneous aquifer of uniform unit width.

The constraints are:

$$h_b(t) = 3{\cdot}0 - 0{\cdot}01t$$

$$q_a(t) = -0{\cdot}24$$

$$T_6 = 1{\cdot}916667$$

The analytical solution for the steady condition is:

$$h(x,t)|_{t=0} = h(x,0) = 2{\cdot}8854 \ln(x/6+1) + 1$$

$$T(x) = x/6 + 1$$

The storage coefficient S is constant in space and given by $S = 1$. h_a and h_b are the heads; T_a and T_b are transmissivities at both ends of the aquifer and x is the longitudinal coordinate.

To determine the influence of noise on the accuracy of the estimated parameter distribution, three cases are considered as follows:

case 1: noise-free;
case 2: 5% randomly positive noise in the head distribution;
case 3: 10% randomly positive noise in the head distribution;
case 4: 20% randomly positive noise in the head distribution.

To obtain the forward solution, a two-node finite element was used. The conventional Gauss method was applied for solution of the system of equations. The forward solution $\mathbf{H}_i$, $i = 0, 1, 2, \ldots m$, for parameter identification was obtained by solving eqn (5) subject to the boundary conditions defined above. Using a time step of $\Delta t = 10$, the distributions of head at $t = 0$, 20, 50 and 100 relative time units for the noise-free case are illustrated in Fig. 1. It is noted that the curves in this figure decrease uniformly and appear to have reached a steady state. As time step is reduced from 10 to 1, the transient state in the early time becomes apparent as illustrated in Fig. 2 where a discharge $q_a = -0{\cdot}12$ is used. The head distributions for cases 2, 3 and 4 in the previous near-steady condition are illustrated in Figs 3, 4 and 5, respectively. These figures show that for the noise-free case, the curves are very smooth; at the 5% level noise has only slight influence, but at the 10% level the noise becomes a significant component. Consequently, head distributions with 10% or 20% noise may have a large effect on parameter estimation.

Table 1. The sums of square errors between the theoretical and computed transmissivities for different noise levels

Time step	Noise-free	5% noise	10% noise	20% noise
1	7·3241E-10	0·015275	0·06651	0·334453
2	1·1500E-9	0·414756	1·36951	4·284167
3	7·3792E-10	0·070710	0·41162	2·787507
4	8·7878E-11	0·011092	3·35765	3·941245
5	5·4380E-10	0·032245	0·10781	0·381388
6	9·5356E-11	0·152362	1·07671	106·0827
7	4·2010E-10	0·243445	0·81751	2·364574
8	2·0156E-10	0·420950	3·28718	703·2505
9	2·2958E-10	0·044945	0·27531	4·814765
10	1·6915E-11	0·012524	0·05655	0·279624

5.2 Results for comparison

For inverse solution, a constant element is used. In order to reduce the residual error in the discretized form of eqn (15), the least squares method is used together with Gauss elimination for solution of the system of equations. The computed hydraulic head distributions are used directly to determine transmissivity for the noise-free case. For each of the two subsequent sets of head distribution data, the transmissivity vector **T** can be solved using eqn (15). For a total of 10 steps, 10 sets of **T** were obtained. In fact, all these transmissivity values should be the same. However, it is almost impossible to obtain a constant solution as a result of noise. Table 1 shows the sum of square errors between the theoretical solution and computed results for the

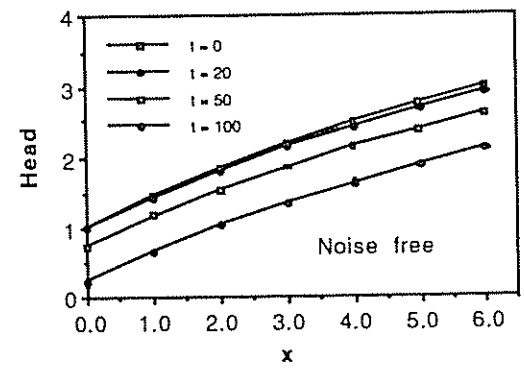


Fig. 1. The head distribution along the aquifer at different times for noise-free case.

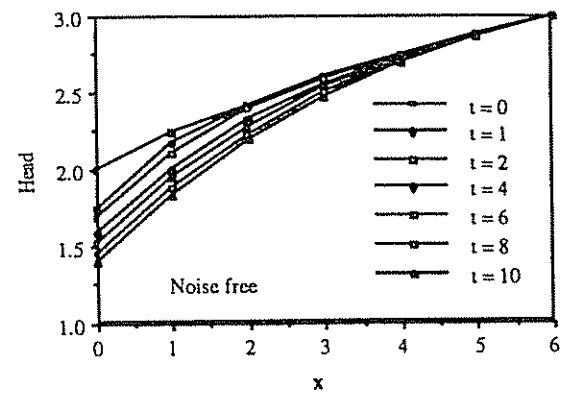


Fig. 2. The head distribution along the aquifer at different times in early time for the noise-free case (case I).

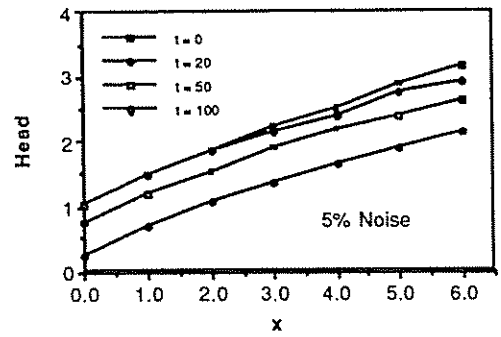


Fig. 3. The head distribution along the aquifer at different times for case 2.

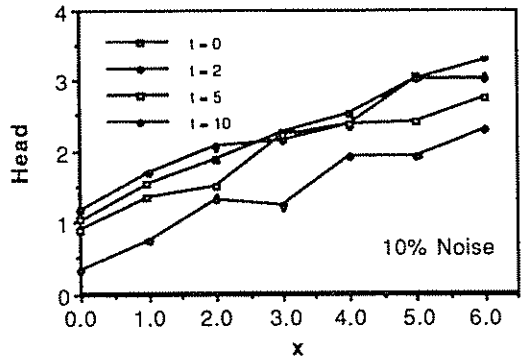


Fig. 4. The head distribution along the aquifer at different times for case 3.

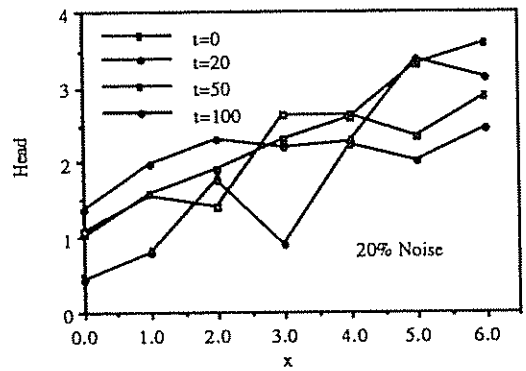


Fig. 5. The head distribution along the aquifer at different times for case 4.

noise-free case. It is noted that for the noise-free case, the estimated parameters are very close to the theoretical solution, although the errors vary with different time steps.

The noise in head distribution will largely affect the accuracy of the estimated parameters. Different computations for cases 2–4 were made using corresponding head distributions as shown in Figs 3–5. The square errors in the estimated parameters for cases 2–4 were obtained using the same method and are shown in Table 1. It is apparent that the square errors increase when the noise level increases. The magnitude of square error increases 8–10 orders of magnitude when noise level increases from 0% to 5%, and the magnitude increases one or two orders when transiting from 5% to 10% or from 10% to 20%, depending on the methods used.

All results in Table 1 also show that by using head distributions at different time steps, the estimated parameters vary greatly at each time step when noise exists. For high noise levels, the error difference at different steps becomes worse. In order to reduce the effects of noise, five methods, mentioned previously, were used, as shown in Table 2.

The computed results, all of which are from the same data, are illustrated in Table 3. From Table 3, it is noted that all methods can give reasonable parameters for the noise-free case. However, the accuracy of solution is dependent on both the method used and noise level when noise is contained in the head distribution. The comparisons of all methods are as follows.

Method 1 (average transmissivity) works well for low noise levels. This method may perform very well when the computed transmissivity profile is flat, enabling an estimate to be made that is close to the real value. Otherwise a large error may develop. However, this method presents a reasonable and simple procedure for parameter estimation of nonsteady problems. Method 2 (expanding method) performs very well for high noise

Table 2. The methods used for parameter estimation in the nonsteady flow condition

Method	Description
1	Average transmissivity
2	Expanding method
3	Integration based on trapezoidal rule
4	Integration based on Simpson's rule
5	Least squares method

Table 3. The sums of square errors by different methods

Method	Noise-free	5% noise	10% noise	20% noise
1	0·412910E-9	0·369869E-2	0·129113E-1	11·242634
2	0·186678E-1	0·333209E-1	0·129067	0·7727143
3	0·266153E-1	0·879319E-2	0·318225E-1	0·1021264
4	0·154700E-4	0·717740E-1	0·365445	4·8859186
5	0·991349E-10	0·636231E-1	0·610535	4·2753668

levels, although when noise-free, this method does not work well compared with the other procedures. In addition, this method is not like methods 1, 4 and 5 which are extremely sensitive to the noise level. Method 3 (trapezoidal integration) is the best method among all of them for noisy data since the square errors in cases 3 and 4 are the smallest. However, method 1 yields better results for the noise-free case. The reason that method 1 performs so well at high noise levels may result from the equally weighted integration applied to eqn (4). Method 4 (Simpson's rule) performs better than methods 2 and 3 for noise-free data, although the inverted results have a large error for the high noise cases. This may result from the nonequal weight applied to each term in eqn (4). When excessive noise exists at integration points with more weight, this method will yield larger errors. Method 5 (least squares) performed best for the noise-free case. As the noise level is increased, the square error becomes larger and the method appears very sensitive to noise level.

In general, for low noise level data, the least squares and average methods are recommended. For high noise levels, the expanding matrix methods or integration methods based on the trapezoidal rule are a good choice.

5.3 Results for transmissivity and storage coefficients

As mentioned before, for the case where storage coefficients and transmissivity are unknowns, both parameter vectors **T** and **S** can be obtained by expanding the methods for transmissivity to both parameters. For demonstration purposes, the one-dimensional problem with a linear distribution of storage coefficient was solved using methods 1 and 5. Only the noise-free case was considered for both the near-steady condition and the transient condition. The results are shown in Table 4. From Table 4, it is apparent that the square errors in the near-steady condition are larger than those in the transient condition. This results from the higher head gradient apparent in the transient condition. Close to the steady state, S has no influence on the resulting head distribution. If S is increased or Δt is decreased, transient response is extended enabling the parameters S and T to be better defined, as apparent in Table 4 for both methods 1 and 5.

Table 4 also shows that the square errors in storage coefficient are much larger than those in transmissivity. The reason is apparent from eqn (1). The parameters h, T and S are represented by different orders of differentiation. Head is represented by a second-order differential, transmissivity by a first-order differential and storage is constant. When eqn (1) is solved numerically, h is determined with the highest accuracy and S with the lowest accuracy.[8,10] Consequently, for the same problem, the same accuracy cannot be obtained in the determination of different parameters.

Table 4. The sums of square errors for transmissivity and storage coefficient in the near-steady condition and transient condition by methods 1 and 5

Method	Transmissivity	Storage coefficient	Condition
1	0·341098E-3	0·4458875	Near-steady
5	0·139542E-5	0·232971E-2	Near-steady
1	0·166410E-3	0·112158	Transient
5	0·491881E-6	0·109948E-3	Transient

5.4 Noise reduction

From Tables 1 and 3, it is apparent that all the methods examined are very sensitive to noise level. Parameters may be better estimated if noise levels can be reduced, but this intrinsically involves data smoothing and therefore data loss. For the one-dimensional example, using the coefficients of eqn (28), noise within the head record can be effectively filtered. An example for noise levels equal to 20% is illustrated in Figs 6, 7, 8 and 9 for time $t = 40$, 80, 130 and 180 relative units, respectively, using eqn (29). Phase shift and scale change may be amended through:

$$g' = \alpha + \beta g \tag{30}$$

where α is the phase shift, β is the gain factor, which can be determined by comparing the original data with the filtered data, g are the filtered data and g' are the modified data.

From Figs 6, 7, 8 and 9, it is apparent that the high frequency noise can be filtered out, resulting in smoother distribution profiles. Similarly, results can be improved by increasing the number of time steps.

Using the filtered record, and applying the same procedures and methods as previous, improved results were obtained as documented in Table 5. Comparing Table 3 with Table 5, it is apparent that by methods 1, 4 and 5 the square errors are reduced for the high noise ratio filtered data, but using methods 2 and 3 degrades the results.

Square error magnitudes are largely dependent on the choice of filter parameter. For a low pass digital filter, a

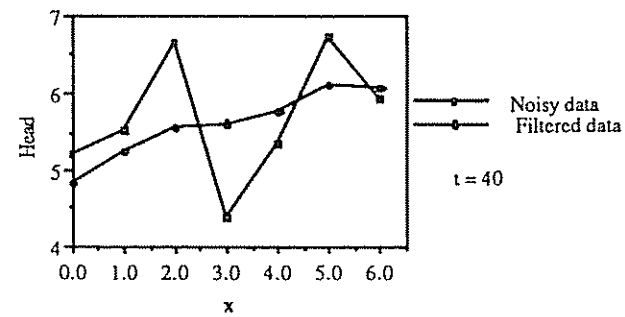


Fig. 6. The head distributions for noisy data and filtered data at time $t = 40$.

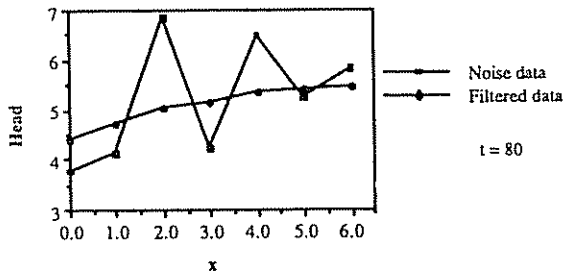


Fig. 7. The head distributions for noisy data and filtered data at time $t = 80$.

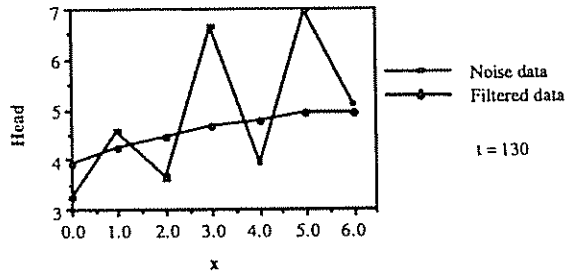


Fig. 8. The head distributions for noisy data and filtered data at time $t = 130$.

small number N should be chosen, and the cut-off frequency parameter CF may be determined by trial-and-error. Figure 10 shows the relation between the square error and CF values for the five different methods.

It is apparent that small CF values yield small square error magnitudes and the square error increases for larger CF magnitudes. It is also noted that for some methods the error increases irregularly with CF as apparent for methods 1, 2 and 4. Comparing all methods, method 5 appears most predictable and gives a small error when the CF value is small and method 4 is not particularly robust. Accordingly, method 5 is the best choice when this method is combined with a digital filter since the square error is small and relatively stable with increase of CF.

5.5 Two-dimensional example

A similar problem to that attempted by Carrera and

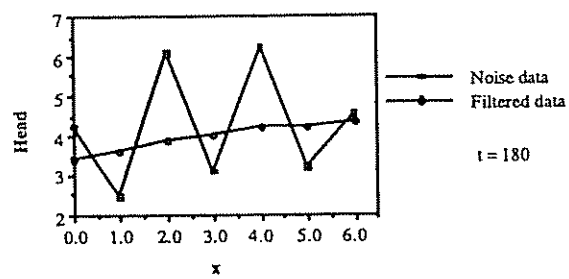


Fig. 9. The head distributions for noisy data and filtered data at time $t = 180$.

Table 5. The sums of square errors using digital filter to reduce the noise levels for different methods

Method	5% noise	10% noise	20% noise
1	0·51366E-1	0·17623	0·66699
2	0·33452	0·59663	1·59750
3	0·43390E-1	0·16797	0·80764
4	0·40231E-1	0·15399	0·77563
5	0·47641E-1	0·24670	1·80481

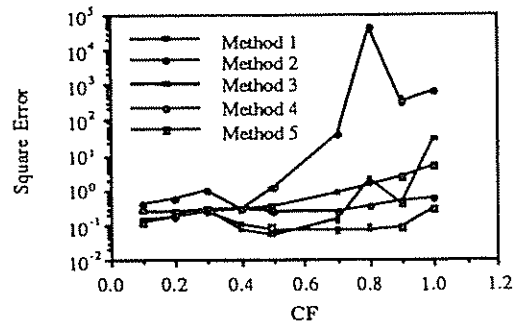
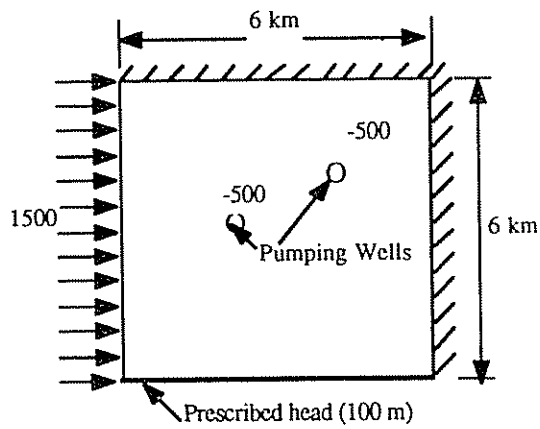


Fig. 10. The relationship between square error and cut-off frequency CF for different methods.

Neuman[1] is used as an example. The idealized aquifer is square with an area of 36 km^2. The boundary conditions and pumping wells are shown in Fig. 11. For simplicity, no areal recharge is considered. The whole flow region is divided into 36 square elements (6×6 mesh) for forward computation of the head distribution.

In the inverse solution, only a 3×3 mesh is used. A four-node constant element is used for both forward and inverse solutions. The head distribution at steady state is computed using eqn (7) and is shown in Fig. 12 for $t = 0$. The transient condition is computed using eqn (9) with a time step of 1 day. The contours of head, computed at t from 0 to 10 days, are shown in Fig. 12 for $t = 2$, 4, 6 and 10 days. A total of 10 steps are considered and the computed heads are used for parameter estimation based on a 3×3 mesh (nine elements). The relative error of each zone for different methods is shown in Table 6. It is apparent that the relative errors in zones 5, 6 and 9 are very large and those in the remaining zones are quite small. From Fig. 12, it is noted that the variation of head gradients in zones 5 and 6 is very small. The reason for this correspondence between high error magnitudes and small gradient is that small magnitude terms in matrix H, formed by head gradients, result in a singularity in eqn (9). The error in zone 6 is due to an inaccurate head distribution around the well and may be improved by a larger element density.

The sum of square relative errors for each method are shown at the bottom of Table 6. Comparing them, it is apparent that methods 3, 4 and 5 work best. Unfortunately, method 1 gives a very large square error, although it performs well for the noise-free case. If a



a

50	15	5
150	50	15
150	150	50

b

Fig. 11. Confined aquifer. (a) Geometry and boundary condition (the discharge and recharge in m^3/day); (b) transmissivity per zone in m^2/day.

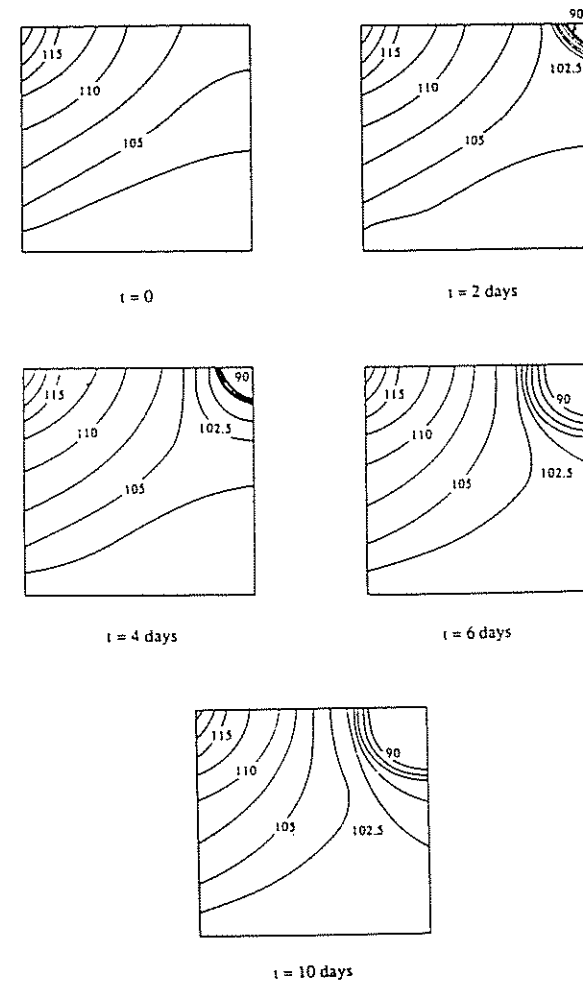


Fig. 12. The head distribution for the steady ($t = 0$) and the nonsteady state ($t = 2$, 4, 6 and 50 days).

fine mesh is used in forward solution, this method may give better results.

6 CONCLUSIONS

Although major comparisons are completed for a simplified one-dimensional problem, they illustrate the principal advantages and disadvantages of the methods presented. The following points may be drawn from the computations and comparisons:

(1) For the noise-free case, a direct solution based on two subsequent data sets may be used for parameter estimation in the nonsteady state. However, different noise levels affect the accuracy of the estimated parameters.
(2) The various methods perform differently, and to some degree unpredictably, at variable noise levels. No definitive choice is apparent although the least squares method and the average method perform well at low noise levels, and the integration method based on the trapezoidal rule may be more suitable for higher noise levels.
(3) All methods presented are capable of being expanded to determine storage coefficients and areal discharge. However, the accuracy with

Table 6. The relative errors and sum of square relative error between the real and computed parameter for different methods

Zone	Method 1	Method 2	Method 3	Method 4	Method 5
4	0·0140	0·0167	0·0139	0·0145	0·0140
5	0·189	0·196	0·17	0·169	0·172
6	2·496	0·403	0·587	0·625	0·507
7	0·009	0·051	0·011	0·012	0·011
8	0·004	0·0860	0·036	0·031	0·018
9	0·224	0·928	0·248	0·234	0·254
$\Sigma\epsilon^2$	6·316	1·805	0·436	0·476	0·351

which different parameter sets may be determined depends on the magnitudes of storage coefficient relative to time-step resolution in the data set.

(4) Using a filter to reduce the noise level is an acceptable procedure, and a small number N and a small cut-off frequency should be used (in this work, $N = 2$ was used). The comparison shows that the least squares method combined with a digital filter gives superior results.

(5) A two-dimensional example shows that conclusions drawn from the one-dimensional example may largely be applied to two-dimensional problems. Accurate recording of the head distribution improves the fidelity of the resulting parameter estimation.

ACKNOWLEDGMENT

This research has been supported by the US Department of the Interior through the National Mine Land Reclamation Center under agreement CO3880-26. The authors also thank reviewers for their valuable comments.

REFERENCES

1. Carrera, J. & Neuman, S., Estimation of aquifer parameters under transient and steady state conditions: Application to synthetic and field data. *Water Resour. Res.*, **22** (1986) 228–42.
2. Dagan, G. & Rubin, Y., Stochastic Identification of recharge, transmissivity, and storativity in aquifer transient flow: A quasi-steady approach. *Water Resour. Res.*, **24** (1988) 1698–710.
3. Ginn, T.R., Cushman, J.H. & Houck, M.H., A continuous-time inverse operator for groundwater and contaminant transport modeling: deterministic case. *Water Resour. Res.*, **26** (1990) 241–52.
4. Kool, J.B. & Parker, J.X., Analysis of the inverse problem for transient unsaturated flow. *Water Resour. Res.*, **24** (1988) 817–30.
5. Loaiciga, H.A. & Marino, M.A., The inverse problem for confined aquifer flow: Identification and estimation with extensions. *Water Resour. Res.*, **23** (1987) 92–104.
6. Loaiciga, H.A. & Marino, M.A., Parameter estimationing groundwater: Classical, Bayesian, and deterministic assumptions and their impact on management policies. *Water Resour. Res.*, **23** (1987) 1027–35.
7. Richter, G.R., An inverse problem for the steady state diffusion equation. *SIAM J. Appl. Math.*, **41** (1981) 210–21.
8. Richter, G.R., Numerical identification of a spatially varying diffusion coefficient. *Math. Comp.*, **36** (1981) 375–86.
9. Rosen, J.B., The gradient projection method for nonlinear programming. Part I, Linear constraints. *J. Soc. Indust. Appl. Math.*, **8** (1960) 181–217.
10. Xiang, J. & Elsworth, D., A comparison of several optimization methods for inverse solution of groundwater flow systems. In *Proc. 5th Canadian and American Conf. on Hydrogeology*, Canada, September, ed. Stefan Buchu, National Water Well Association, Dublin, Ohio, 1990, pp. 276–88.
11. Xiang, J. & Elsworth, D., Low-order finite elements for parameter identification in groundwater flow. *Applied Mathematical Modeling*, **15**(5), (1991) 256–66.
12. Yeh, W.W.-G., Review of parameter identification procedures in groundwater hydrology: The inverse problem. *Water Resour. Res.*, **22** (1986) 95–108.
13. Yeh, W.W.-G., Yoon, Y.S. & Lee, K.S., Aquifer parameter identification with Kriging and optimum parametrization. *Water Resour. Res.*, **19** (1983) 225–33.
14. Yoon, Y.S. & Yeh, W.W.-G., Parameter identification in an inhomogeneous medium with the finite-element method. *S. Petroleum Engineer J. Trans. AIME*, **261** (1976) 217–26.