



Parameter estimation for leaky aquifers using the extended Kalman filter, and considering model and data measurement uncertainties

H.D. Yeh*, Y.C. Huang

Institute of Environmental Engineering, National Chiao-Tung University, 75 Po-Ai Street, Hsinchu 30039, Taiwan

Received 5 November 2002; revised 22 June 2004; accepted 25 June 2004

Abstract

A method using the extended Kalman filter (EKF) is proposed to identify the hydraulic parameters in leaky aquifer systems both with and without considering the aquitard storage. In the case without considering the aquitard storage, Hantush and Jacob's model combined with EKF can optimally determine the parameters for the leaky aquifer when analyzing the drawdown data. Coupled with Neuman and Witherspoon's model, the EKF is also employed to estimate the four parameters of aquifers. The observed drawdown data may be either interpolated using the Lagrangian polynomial or recursively used while implementing the EKF. The proposed method can identify the parameters, using part of the interpolated drawdown data or recursively used data, and obtains results with good accuracy. In the field-pumping test, a long pumping time may not be necessary if the proposed method is implemented on a computer, which is connected to pressure transducers and a data logger. In the process of parameter estimation, the leakage coefficient changes marginally for the first few observations. This phenomenon reflects the fact that there is a time lag between the start of pumping and the leakage effect on the drawdown. The analyses of the data uncertainty demonstrate that the EKF approach is applicable for drawdown data even when it contains white noise or temporal correlated noise. Finally, the choice between Hantush and Jacob's model and Neuman and Witherspoon's model depends on the hydrogeological condition of the aquifer system indicated in the analyses of the model uncertainty. Hantush and Jacob's model is shown to be a good choice for representing the leaky aquifer system if the aquitard storage is comparatively small.

© 2004 Elsevier B.V. All rights reserved.

Keywords: Parameter estimation; Kalman filter; Lagrangian polynomial; Groundwater; Leaky aquifer; Model uncertainty

1. Introduction

Hydrogeologic parameters are very important in site characterization, so groundwater hydrologists

often conduct pumping tests to determine hydrogeologic parameters, such as hydraulic conductivity and storage coefficient. These parameters are necessary information for quantitative and/or qualitative groundwater studies. In a leaky aquifer the semi-pervious bed (also known as the aquitard), although of very low permeability, may yield significant amounts of water to the adjacent pumped aquifer. As time

* Corresponding author. Fax: +886-35-726050.

E-mail address: hdyeh@mail.nctu.edu.tw (H.D. Yeh).

increased, leakage across the semi-impervious bed may become appreciable and flow is not restricted to the pumped aquifer alone. The additional water may be derived from storage of the aquitard and adjacent unpumped aquifers. Therefore, the leaky aquifer must be viewed as part of a complex multiple aquifer system.

Two approaches have been developed for dealing with leaky aquifers, one based on the assumption that the aquitard storage is negligible, and the other considering the aquitard storage. Jacob (1946) developed a partial differential equation for non-steady radial flow in a leaky aquifer, assuming that hydraulic head in the unpumped adjacent aquifer is constant and the storage capacity of the aquitard is negligible. These two assumptions greatly facilitated mathematical treatment of the problem, and Jacob used this approach to develop a large number of solutions to various problems involving flow in aquifers with vertical leakage.

Hantush and Jacob (1955) described non-steady radial flow to a well in a fully penetrated leaky aquifer under a constant pumping rate. In their model, the aquitard is overlain by an unconfined aquifer, and the main aquifer is underlain by an impermeable bed. Their solution is herein called the three-parameter model. Hantush (1960) also presented a modified approach to include the effect of the aquitard storage. Neuman and Witherspoon (1969) gave a solution describing the drawdown of the lower and pumped aquifer in a hydrogeologic system, which is composed of two confined aquifers and one aquitard. Their solution, which considers the effect of aquitard storage and neglects the drawdown in the unpumped aquifer, is called the four-parameter model. Both the three-parameter and four-parameter models are also mentioned in several recent textbooks, for example, Dawson and Istok (1991) and Batu (1998).

In the three-parameter model, the graphical method based on Hantush's or Walton's type curves (Batu, 1998) requires data plotting work and individual judgment during the curve fitting procedure. Therefore, errors may be introduced during the fitting process. In the four-parameter model, the use of the graphical matching method based on the Neuman and Witherspoon's model is practically impossible since there will be several families of type curves. Yeh (1987) used the non-linear least-squares

and finite-difference Newton's method (NLN) for identifying the parameters of the confined aquifer, a method which has the advantages of high accuracy and quick convergence. Yeh and Han (1989) subsequently used NLN to determine the hydraulic parameters of the leaky aquifers.

The Kalman filter was developed by R.E. Kalman in the late 1950s, and its main applications have been in control systems, tracking and navigation of all sorts of vehicles, as well as predictive design of estimation and control systems. Works using the Kalman filter for hydraulic-parameter and water table-related estimations may be divided into two categories. One applies the Kalman filter in a linear system (e.g. Van Geer and Van Der Kloet, 1985; Van Geer and te Stroet, 1990; Van Geer et al., 1990; Lee et al., 2000; Bierkens et al., 2001) and the other deals with non-linear problems using the extended Kalman filter (EKF) (e.g. Chander et al., 1981; Katul et al., 1993; Bierkens, 1998; Cahill et al., 1999).

Chander et al. (1981) estimated the parameters for both non-leaky and leaky aquifers by the iterated EKF. For the leaky aquifer, the measurement equation was a truncated form of Hantush and Jacob's well function (1955), which was suitable for small leakage coefficient and/or large pumping time. Van Geer and Van Der Kloet (1985) presented two linear, filter-based schemes for parameter estimation in groundwater flow problems. An optimal estimate was simultaneously computed for the original state, i.e., heads and the parameter state. Van Geer and te Stroet (1990) combined MODFLOW with a filtering framework and updated the prior estimates of hydraulic parameter values using an off-line procedure, when minimizing the difference between the actual head measurements and those predicted from the MODFLOW–Kalman filter framework. Van Geer et al. (1990) also adopted the idea of using a filter for state estimation in the absence of significant dynamic behavior and studied the applicability of the filter to a relatively quickly reacting groundwater system.

Katul et al. (1993) applied the EKF to test the determination of the hydraulic conductivity function from a field drainage experiment. Bierkens (1998) embedded the stochastic differential equation in the EKF algorithm to calibrate the parameters and noise statistics of the stochastic differential equation on a time series of water table depths. Eigbe et al. (1998)

reviewed the problems of using a filter with groundwater flow models, and identified procedures that would improve its efficiency and convenience for field applications. Cahill et al. (1999) derived optimal parameters for an effective large-scale hydraulic conductivity field that considers both the temporal and spatial variations of the moisture content. Lee et al. (2000) presented a linear model to automatically decide the terminal time of a pumping test and to optimally identify the storage coefficient and anisotropic aquifer transmissivities and of a confined aquifer, using Cooper and Jacob's equation and Papadopoulos' approach. Bierkens et al. (2001) modeled the spatiotemporal variation of shallow water table depth with a regionalized version of an autoregressive exogenous (ARX) time series model. There, the regionalized ARX parameters were estimated by embedding the regionalized ARX model in a space–time Kalman filter.

Beck (1987) reviewed the role of uncertainty in the identification of mathematical models of water quality and in the application of these models to problem of prediction. Eisenberg et al. (1989) selected five broad classes of sources of uncertainty in the area of the geologic repository for radioactive waste. The five types of uncertainty they considered are: (1) systematic and random error in measurement; (2) spatial variations in geologic parameters; (3) conceptual model uncertainties relative to geometrical configuration, major features, and boundary conditions; (4) physicochemical process modeling; and (5) future states of nature. Freeze et al. (1992) presented an overview and examples of the concepts and techniques for dealing with uncertainty in the framework of hydrogeological decision analysis. Wheeler et al. (2000) reviewed the problems of uncertainty in the definition of aquifer properties and the scale-dependence of dispersion with specific regard to well-capture zones. Christiaens and Feyen (2001) compared and evaluated the uncertainties resulting from four ways to obtain soil hydraulic parameters with respect to their resulting uncertainties on different hydrological model outputs. Zheng and Bennett (2002) listed three methods: sensitivity analysis, Monte Carlo method, and the first-order error analysis for evaluating uncertainty.

In this current study, the EKF is used to determine the hydraulic parameters of the leaky aquifer system.

The state vector of the EKF is comprised of the hydraulic parameters and a model, i.e. either Hantush and Jacob's model or Neuman and Witherspoon's model, is used as the measurement equation of the EKF. The unknown parameters can be estimated on-line as the measurement data come in and are interpolated by the Lagrangian polynomial. Thus, this proposed approach can be implemented on a computer, which can be connected with a data logger and a pressure transducer that measures the water level in the observation well during the pumping test. The desired hydraulic parameters can thus be estimated in the field on-line and once stable estimates of the parameters are obtained, then the pumping test may be terminated.

The five objectives of this paper are: (1) to develop a numerical approach based on the EKF method which can identify the hydraulic parameters of leaky aquifer in the field on-line; (2) to examine the accuracy and applicability of Chander et al.'s approach (1981) while using the simplified formula of the Hantush and Jacob's model (1955); (3) to determine four hydraulic parameters using the EKF method based on Neuman and Witherspoon's model (1969); (4) to test the applicability of the EKF approach if the drawdown data contain white noise or temporal correlated noise; (5) to compare the choice between either Hantush and Jacob's model or Neuman and Witherspoon's model for parameter estimation. The data measurement error (herein treated as noise) is related to the problem of measurement data uncertainty, and the choice among various mathematical models for representing a target aquifer belongs to the issue of model uncertainty.

2. Methodology

This section includes two parts: one is the basic framework of the EKF and the other is the Lagrangian polynomial. To implement EKF, the latter approach is applied to interpolate the drawdown data with non-uniform time intervals having a small and uniform time interval.

2.1. Discrete extended Kalman filter

Consider a non-linear dynamical system whose state vector is described as follows

(Grewal and Andrews, 1993)

$$x_k = f(x_{k-1}, k - 1) + w_k \quad (2.1)$$

where x_k is the state vector of system at time step k , $f(x_{k-1}, k-1)$ is the non-linear function of the system, and w_k is the state noise assumed to be normally distributed with zero mean white (uncorrelated) sequence with known covariance Q_k . Note that Q_k is assumed to be a zero vector throughout the filtering steps.

For computing the predicted state estimate, the non-linear implementation equations for the state vector can be described as

$$\hat{x}_k(-) = f(\hat{x}_{k-1}(+), k - 1) \quad (2.2)$$

where $\hat{x}_k(-)$ denotes the a priori estimate at k step and $\hat{x}_k(+)$ represents the a posteriori estimate at $k-1$ step.

A measurement model of system can be described as (Grewal and Andrews, 1993)

$$z_k = h(x_k, k) + v_k \quad (2.3)$$

where $h(x_k, k)$ is the function for the measurement system and z_k is the measurement vector at time step k . The measurement noise v_k is assumed to be a white sequence with known covariance structure R_k . Note that the covariance matrix R_k is assumed constant throughout the filtering process.

As in Eq. (2.2) for computing the predicted measurement, the non-linear implementation equation for the measurement is

$$\hat{z}_k = h(\hat{x}_k(-), k) \quad (2.4)$$

where \hat{z}_k is predicted measurement vector.

This method requires an initial estimate of the process at some point in time step k , and that estimate is based on the knowledge about the process. The error covariance matrix associated with $\hat{x}_k(-)$ is also assumed to be known. The prior estimation error $e_k(-)$ is defined as $x_k - \hat{x}_k(-)$. Initially, the a priori error covariance matrix of $x_k - \hat{x}_k(-)$, denoted as $P_k(-)$, is

$$\begin{aligned} P_k(-) &= E[e_k(-)e_k^T(-)] \\ &= E[(x_k - \hat{x}_k(-))(x_k - \hat{x}_k(-))^T] \end{aligned} \quad (2.5)$$

With the assumption of the a priori error estimate $\hat{x}_k(-)$, the measurement z_k is used to improve

the a priori estimate, as follows

$$\hat{x}_k(+) = \hat{x}_k(-) + \bar{K}_k(z_k - \hat{z}_k) \quad (2.6)$$

where \bar{K}_k is defined as the Kalman gain and $\hat{x}_k(+)$ is the updated estimate at step k .

The Kalman filter uses minimum mean-square error as the performance criterion to find the particular \bar{K}_k that yields an updated estimate. Therefore, the particular Kalman gain that minimizes the mean-square estimation error is

$$\bar{K}_k = P_k(-)H_k^T[H_kP_k(-)H_k^T + R_k]^{-1} \quad (2.7)$$

where the measurement matrix H_k is approximated by the derivative of estimated measurement \hat{z}_k from Eq. (2.4) with respect to each of the state vector. It can be expressed as

$$H_k \approx \left. \frac{\partial h(x, k)}{\partial x} \right|_{x=\hat{x}_k(-)} \quad (2.8)$$

Therefore, \bar{K}_k can be estimated with known $P_k(-)$, H_k , and R_k . Detailed derivations of the minimization process may be found in Grewal and Andrews (1993).

The error covariance matrix derived from Eq. (2.5) can also be expressed as

$$P_k(-) = \Phi_{k-1}P_{k-1}(+)\Phi_{k-1}^T + Q_{k-1} \quad (2.9)$$

where Φ_{k-1} is the state transition matrix. This may be expressed by the derivative of the state vector estimation equation from Eq. (2.2) as

$$\Phi_{k-1} \approx \left. \frac{\partial f(x, k - 1)}{\partial x} \right|_{x=\hat{x}_{k-1}(-)} \quad (2.10)$$

The error covariance matrix associated with the updated (a posteriori) estimate can be written as

$$P_k(+) = \{I - \bar{K}_kH_k\}P_k(-) \quad (2.11)$$

where $P_k(+)$ denotes the error covariance updated by the Kalman gain and the prior error covariance and I represents an identity matrix.

Using Eqs. (2.2), (2.4), (2.6), (2.7), (2.9), and (2.11), the recursive process of EKF is then established.

2.2. Lagrangian polynomial

Data where the x -values are not evenly spaced often occur as the result of experimental observations

or when historical data are examined. The Lagrangian polynomial is perhaps the simplest way used to interpolate the unevenly spaced data.

The general form of Lagrangian polynomial is (Gerald and Wheatley, 1994)

$$P_n(x) = \sum_{i=1}^{n+1} \sum_{\substack{j=1 \\ j \neq i}}^{n+1} \frac{x - x_j}{x_i - x_j} f(x_i) \quad (2.12)$$

It is easy to see that the Lagrangian polynomial passes through each of the points used in its construction. Note that this study utilizes a second-degree Lagrangian polynomial to generate interpolated data points.

3. Application of discrete EKF

This section illustrates how the Kalman filter is coupled with three- and four-parameter models to identify the hydraulic parameters. The difference between the present study and Chander et al. (1981) on the three-parameter model is clearly demonstrated. The three parameter values, i.e. transmissivity T , storage coefficient S , and leakage coefficient L are optimally determined when using the three-parameter model coupled with EKF to analyze the measurement drawdown data. Likewise, the four-parameter model, which has an additional parameter ψ accounting for the effect of the aquitard storage, combined with EKF can estimate four hydraulic parameters in a similar procedure.

3.1. Leaky aquifer without storage effect in an aquitard

3.1.1. Hantush and Jacob's model

With reasonable initial guess values for T , S , L , $P_{k-1}(+)$, and R_k , the algorithm of EKF combined with the Hantush and Jacob's model can determine the best-fit T , S , and L at each identification step. In their model, the aquitard, with a negligible storage, is overlain by an unconfined aquifer, and the main aquifer is underlain by an impermeable bed. The model of Hantush and Jacob describing the drawdown cone within a leaky aquifer in response to

the pumping as a function of radial distance and time is (Hantush and Jacob, 1955, p.98)

$$s = \frac{Q}{4\pi T} W\left(u, \frac{r}{B}\right) \quad (3.1)$$

where s is drawdown, r is the distance between pumping well and observation well, u is dimensionless variable defined as $r^2 S/4Tt$, K' is the vertical conductivity of the aquitard, b' is thickness of the aquitard, r/B is the leakage coefficient L and B is defined as $\sqrt{(K'/b')/T}$, Q is the pumping rate, and $W(u, r/B)$ is the leaky well function.

The leaky well function $W(u, r/B)$ can be expressed as

$$W\left(u, \frac{r}{B}\right) = \int_u^\infty \frac{1}{y} \exp\left[-y - \frac{(r/B)^2}{4y}\right] dy \quad (3.2)$$

where y is a dummy variable. Since the right hand side of Eq. (3.2) is an integral form, a numerical approach is required to evaluate the integration. Both the Laguerre quadrature formula and Gaussian quadrature formula (Carnahan et al., 1969) are employed to evaluate the values of leaky well function with the accuracy to the fourth decimal. The Laguerre integration used to approximate an integral function is usually expressed as

$$\int_0^\infty f(x)e^{-x} dx = \sum_{i=1}^n w_i f(x_i) \quad (3.3)$$

where the w_i is weighting factor and x_i is corresponding to zero of the n th-order Laguerre polynomials. For a small value of u , the Laguerre quadrature formula cannot give the desired accuracy. Therefore, the Gaussian quadrature formula is employed to evaluate the integration of Eq. (3.2) when u is small.

3.1.2. Dynamic and measurement models for Hantush and Jacob's model

Parameters T , S , and L , considered as time invariants during the pumping test, constitute the state vector which is to be estimated at each time step. The state vector in Eq. (2.2) may be expressed as

$$\hat{x}_k(-) = [T \quad S \quad L]^T \quad (3.4)$$

After each time step, the renewed state vector, Eq. (3.4), can be substituted into Hantush and Jacob's model, therefore forming the estimated

drawdown. The estimated measurement from Eq. (2.4) may be combined with Eqs. (3.1) and (3.2) and computed as

$$\hat{z}_k = \frac{Q}{4\pi T} \int_u^\infty \frac{1}{y} \exp\left[-y - \frac{(r/B)^2}{4y}\right] dy \quad (3.5)$$

3.1.3. Linear approximation equations for Hantush and Jacob's model

The first-order approximation for the transition matrix for Eq. (2.10) is given as

$$\Phi_{k-1} \approx \begin{bmatrix} \frac{\partial T}{\partial T} & \frac{\partial T}{\partial S} & \frac{\partial T}{\partial L} \\ \frac{\partial S}{\partial T} & \frac{\partial S}{\partial S} & \frac{\partial S}{\partial L} \\ \frac{\partial L}{\partial T} & \frac{\partial L}{\partial S} & \frac{\partial L}{\partial L} \end{bmatrix} \quad (3.6)$$

Based on Eq. (3.13), $P_k(-)$ can be estimated with known transition matrix Φ_{k-1} . Since the state of the system is described by the parameters, which are mutually independent, thus, Φ_{k-1} in Eq. (3.6) is an identity matrix. Assuming that Q_{k-1} is equal to zero, Eq. (2.9) can be reduced to

$$P_k(-) = P_{k-1}(-) \quad (3.7)$$

To update the hydraulic parameters in Eq. (2.7), the Kalman gain \bar{K}_k , estimated by known H_k and the prior covariance matrix $P_k(-)$, is required. The input z_k in Eq. (2.6) is the drawdown data generated by the Lagrangian polynomial. The measurement matrix H_k from Eq. (2.8) consists of three elements, which are the partial derivatives of the estimated drawdown \hat{z}_k with respect to T , S , and L , respectively; that is

$$H_k \approx \left[\frac{\partial \hat{z}_k}{\partial T} \quad \frac{\partial \hat{z}_k}{\partial S} \quad \frac{\partial \hat{z}_k}{\partial L} \right]^T \quad (3.8)$$

The posterior covariance matrix $P_k(+)$ is then calculated by inserting the Kalman gain and prior covariance matrix $P_k(-)$. The posterior covariance matrix is used recursively as the prior covariance at the next time step.

One of the criteria to stop this recursive process is

$$TOL_T = |T(k+1) - T(k)| \quad (3.9)$$

where TOL_T is the tolerance criterion for T . The tolerance criterion TOL_S for S and TOL_L for L can be expressed in a similar way. When all

the criteria are met, the recursive process is terminated and the optimal parameters may then be determined.

3.2. Leaky aquifer with storage effect in aquitard

Here, the EKF method along with Neuman and Witherspoon's model (1969) is applied to leaky aquifer considering the effect of aquitard storage. Given the initial guesses of parameters, error covariance, and measurement noise, the four best-fit parameters are identified when the convergence criteria are met.

Neuman and Witherspoon presented a closed-form solution for the problem of flow to a well in a confined infinite radial system composed of two confined aquifers that are separated by an aquitard (Neuman and Witherspoon, 1969). Differing from Hantush and Jacob's work (1955), Neuman and Witherspoon's model includes the effect of the aquitard storage on the drawdown of the pumping aquifer. Their mathematical model may be written as

$$s = \frac{Q}{2\pi T} \int_0^\infty \frac{1}{y} [1 - \exp(-y^2 \bar{t}_D)] J_0[w(y)] dy \quad (3.10)$$

where $\bar{t}_D = t_D L^2 / 16\psi^2$, $t_D = Tt/r^2 S$, $L = r/B$, $\psi = \beta/L$, $B = \sqrt{Tb'/K'}$, $\beta = r\sqrt{S'/4B\sqrt{S}}$, and $w^2(y) = L^2 y^2 / 16\psi^2 - L^2 y \cot y$. Note that Eq. (3.10) is valid for all values of time intervals and the Bessel function $J_0[w(y)]$ must be set to zero when $w^2(y) < 0$.

The state vector of Eq. (2.2) at each time step is

$$\hat{x}_k(-) = [T \quad S \quad L \quad \psi]^T \quad (3.11)$$

Note that those four parameters are also taken as time invariants during the pumping test. After each time step, the renewed state vector $\hat{x}_k(-)$ from Eq. (3.11) can be substituted into Neuman and Witherspoon's model, consequently forming the estimated drawdown \hat{z}_k . The estimated measurement \hat{z}_k from Eq. (2.4) is expressed as

$$\hat{z}_k = \frac{Q}{2\pi T} \int_0^\infty \frac{1}{y} [1 - \exp(-y^2 \bar{t}_D)] J_0[w(y)] dy \quad (3.12)$$

The measurement \hat{z}_k estimated by Eq. (3.12) is used in Eq. (2.6) to obtain the updated state vector. The first-order approximation for the transition

matrix of Eq. (2.10) is given as

$$\Phi_{k-1} \approx \begin{bmatrix} \frac{\partial T}{\partial T} & \frac{\partial T}{\partial S} & \frac{\partial T}{\partial L} & \frac{\partial T}{\partial \psi} \\ \frac{\partial S}{\partial T} & \frac{\partial S}{\partial S} & \frac{\partial S}{\partial L} & \frac{\partial S}{\partial \psi} \\ \frac{\partial L}{\partial T} & \frac{\partial L}{\partial S} & \frac{\partial L}{\partial L} & \frac{\partial L}{\partial \psi} \\ \frac{\partial \psi}{\partial T} & \frac{\partial \psi}{\partial S} & \frac{\partial \psi}{\partial L} & \frac{\partial \psi}{\partial \psi} \end{bmatrix} \quad (3.13)$$

The relation between $P_k(-)$ and $P_{k-1}(+)$ can also be described by Eq. (3.7). Both H_k and $P_k(-)$ should be calculated in advance by Eqs. (2.8) and (2.9), respectively. The input z_k is the measurement data produced by the Lagrangian polynomial. The measurement matrix H_k is composed of $\partial \hat{z}_k / \partial T$, $\partial \hat{z}_k / \partial S$, $\partial \hat{z}_k / \partial L$, and $\partial \hat{z}_k / \partial \psi$, which can be approximated by the finite-difference formula.

The posterior covariance matrix $P_k(+)$ is then updated by the Kalman gain and the prior covariance matrix $P_k(-)$. The updated covariance $P_k(+)$ can function as the prior covariance in the next time step for the recursive process.

The tolerance criteria of the four-parameter model are the same as those of the three-parameter model, except that the additional criterion, e.g. TOL_ψ , is for the parameter ψ .

3.3. Evaluation of H_k

In Hantush and Jacob' model (1955), the value of the drawdown depends on the decision variables, T , S , and L , which may be written as

$$s = \frac{Q}{4\pi T} w(T, S, L) \quad (3.14)$$

where the leaky well function $w(T, S, L)$ represents the integral in Eq. (3.2).

The derivatives $\partial \hat{z}_k / \partial T$, $\partial \hat{z}_k / \partial S$, $\partial \hat{z}_k / \partial L$, are, respectively,

$$\frac{\partial \hat{z}_k}{\partial T} = -\frac{Q}{4\pi T^2} w + \frac{Q}{4\pi T} \frac{\partial w}{\partial T} \quad (3.15)$$

$$\frac{\partial \hat{z}_k}{\partial S} = \frac{Q}{4\pi T} \frac{\partial w}{\partial S} \quad (3.16)$$

$$\frac{\partial \hat{z}_k}{\partial L} = \frac{Q}{4\pi T} \frac{\partial w}{\partial L} \quad (3.17)$$

where $\partial w / \partial T$ is approximated by forward differencing as

$$\frac{\partial w}{\partial T} = \frac{w(T + \Delta T, S, L) - w(T, S, L)}{\Delta T} \quad (3.18)$$

and the other partial derivatives can also be expressed in a similar manner. The increment shown in the denominator of Eq. (3.18) may be approximated by the parameter value times a factor of 10^{-3} or less, e.g. $\Delta T = 10^{-3} T$

In the four-parameter model, the values of drawdown depend on the decision variables, T , S , L , and ψ , which may be written as

$$s = \frac{Q}{2\pi T} G(T, S, L, \psi) \quad (3.19)$$

where the function $G(T, S, L, \psi)$ represents the integral in Eq. (3.10).

In this model, three derivatives $\partial \hat{z}_k / \partial T$, $\partial \hat{z}_k / \partial S$, and $\partial \hat{z}_k / \partial L$ are, respectively, given in Eqs. (3.14)–(3.16) and the additional derivative $\partial \hat{z}_k / \partial \psi$ is

$$\frac{\partial \hat{z}_k}{\partial \psi} = \frac{Q}{2\pi T} \frac{\partial G}{\partial \psi} \quad (3.20)$$

where $\partial G / \partial T$ is approximated by forward difference as

$$\frac{\partial G}{\partial T} = \frac{G(T + \Delta T, S, L, \psi) - G(T, S, L, \psi)}{\Delta T} \quad (3.21)$$

and the other partial derivatives $\partial G / \partial S$, $\partial G / \partial L$, and $\partial G / \partial \psi$ are also expressed in a similar manner.

4. Two practical problems

4.1. Problems in Chander et al. (1981)

Chander et al. (1981) used the Kalman filter to identify the parameters of confined and leaky aquifers. For the leaky aquifer, they combined the Kalman filter with approximate drawdown function without considering the effect of the aquitard storage. Their well function was a truncated form of Eq. (3.2) and may be only suitable for long pumping time or/and small

leakage coefficient. The drawdown equation that they used was

$$s = \frac{Q}{4\pi T} \left[2K_0 \left(\frac{r}{L} \right) - w \left(\frac{Tt}{L^2 S} \right) \right] \quad (4.1)$$

This equation may be expanded as

$$s = \frac{Q}{4\pi T} \times 2 \left\{ \left[1 + \left(\frac{r}{2L} \right)^2 + \left(\frac{1}{4} \right) \left(\frac{r}{2L} \right)^4 + \dots \right] \times \ln \left[\frac{1.123}{\frac{r}{L}} \right] + \left(\frac{r}{2L} \right)^2 + \left(\frac{3}{8} \right) \left(\frac{r}{2L} \right)^4 + \dots \right\} - \frac{Q}{4\pi T} \times \left[-0.5572 - \ln \left(\frac{Tt}{L^2 S} \right) + \left(\frac{Tt}{L^2 S} \right) - \left(\frac{1}{4} \right) \left(\frac{Tt}{L^2 S} \right) \right] \quad (4.2)$$

The results from applying Chander et al.'s method may have significant errors when the pumping time is short or/and the leakage coefficient is large. In addition, there are two drawbacks in their study. First, laborious works are required to estimate the initial guesses before applying the EKF. The guess values of T and S were determined by using a truncated Theis equation (Chander et al., 1981, Eq. (4)) to the first two drawdown data. The guess value of L was calculated by using the last available observation and applying the approximate steady-state drawdown relationship for the leaky aquifer as used in the Hantush I method (Kruseman and de Ridder, 1970). These strategies may not yield better initial estimates and lose the superiority of on-line estimation of EKF over other parameter estimation methods. Second, Chander et al. (1981) did not specify stopping criteria for their EKF algorithm. It seems to us that Figs. 1 and 2 in their manuscript indicate that the EKF identification process terminated at the time when running out of the observation data. Under this circumstance, the estimated parameters are not warranted to have desired accuracy. The EKF usually takes more than 200 steps to get convergent results, as shown later in this study. In reality, the number of steps needed when using the EKF to determine the hydraulic parameters depends on those guess values mentioned above. The available observed pumping data given in Chander et al.'s study is no more than 12. Thus, the EKF may run out of data very quickly before the hydraulic parameters converge to satisfactory results.

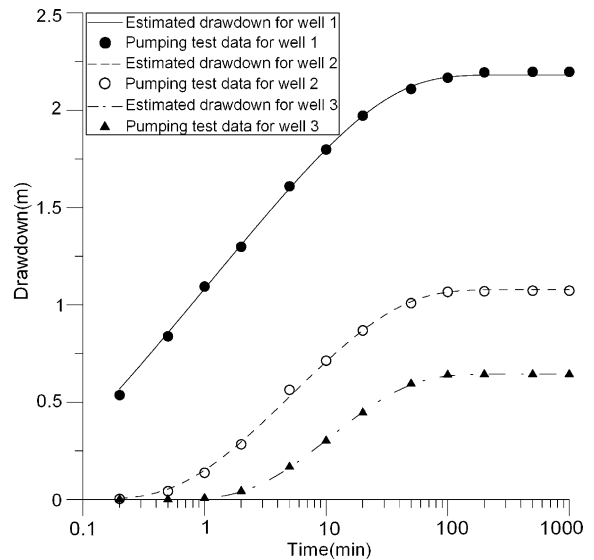


Fig. 1. The pumping test data of Cooper (1963) and the estimated drawdown based on Hantush and Jacob's model (1955) for the leaky aquifer without considering the storage effect.

4.2. Problem of small amount of drawdown data

Usually, the changes for the estimated parameters and the elements of error covariance matrix

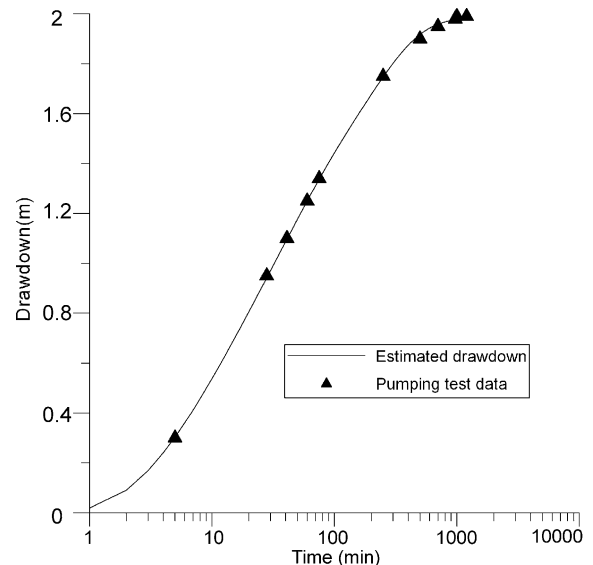


Fig. 2. The pumping test data of Sridharan et al. (1987) and the estimated drawdown based on Neuman and Witherspoon's model (1969) for the leaky aquifer with considering the storage effect.

of the state vector in each step of the EKF are fairly small. Therefore, the parameter estimation using EKF usually requires the preparing action of a large amount of the drawdown data. However, the number of field measurements is usually limited, say less than 50. Two approaches suggested here to solve this problem are: (1) applying the Lagrangian polynomial to interpolate the field measurement data with smaller time intervals, and (2) recursively using the field measurement data during the step-wise identification process. The first approach is suitable to be used for on-line measurement of the field-pumping test, whereas the second approach is recommended to be applied for an existing drawdown data set.

5. Data analyses and discussion

5.1. Two sets of pumping test data

Two sets of pumping test data are chosen, one for the three-parameter model and the other for the four-parameter model. Three time–drawdown data sets measured from observation wells, as reported in Cooper (1963) and cited by Lohman (1972, p.31, Table 11), are selected for data analyses. The distances between the pumping well and the observation wells 1, 2, and 3, are respectively, 30.48, 152.4, and 304.8 m. The pumping rate Q is 5450.98 m³/day, the thickness of the aquitard is 30.48 m and total pumping time is 1000 min (16.67 h). The data set for the four-parameter model is taken from Sridharan et al. (1987). The distance between the observation well and pumping well is 29.0 m and the pumping rate Q is 136.26 m³/day. The total pumping time is 1200 min (20.00 h). These two sets of the pumping test data are shown in Tables 1 and 2 and interpolated using the Lagrangian polynomial.

The time intervals of the pumping test data are normally non-uniform. A small and uniform time interval, say 6 s (i.e. 0.1 min), was chosen when applying the Lagrangian polynomial to interpolate the data. In order to smooth the interpolation results, the time coordinate (t) was transferred to the logarithmic scale ($\log t$) before interpolation and then the second-degree Lagrangian polynomial was employed to interpolate drawdown data for the drawdown versus $\log t$. At the beginning of

Table 1

Time–drawdown data obtained from three observation wells (Cooper, 1963, p. 31)

Time (min)	Observation well (m)		
	1	2	3
0.2	0.536	0.003	0.000
0.5	0.838	0.043	0.000
1	1.094	0.137	0.006
2	1.298	0.284	0.043
5	1.609	0.536	0.168
10	1.798	0.713	0.302
20	1.972	0.869	0.445
50	2.109	1.009	0.594
100	2.167	1.067	0.640
200	2.195	1.070	0.643
500	2.198	1.073	0.643
1000	2.198	1.073	0.643

The unit of drawdown is meter.

the pumping test, the first three drawdown data points were utilized to generate interpolated data between the first and third data points. Once a new drawdown point is measured, the second, third, and the new measured data point (i.e. the fourth data point) were used to generate interpolated data between the third and new data points. Such an interpolation procedure can keep going on in the field until the estimated parameters meet the convergent criteria. For the data sets from Cooper (1963) and Sridharan et al. (1987), the total numbers of the interpolated drawdown data generated by the Lagrangian polynomial are 10,000 and 12,000, respectively. The interpolated time–drawdown data are then ready for EKF identification process.

Table 2

Pumping test data (Sridharan et al., 1987, p. 170)

Time (min)	Drawdown (m)
5	0.3
28	0.95
41	1.1
60	1.25
75	1.34
250	1.75
500	1.9
700	1.95
970	1.98
1000	1.99
1200	1.99

$Q = 136.26$ m³/day; $r = 29.00$ m.

5.2. Assessment for prediction errors

To assess the accuracy of the estimated parameters or the goodness-of-fit of the estimated drawdown to the observed drawdown, two error criteria, mean error (ME) and standard error of estimate (SEE), are used to calculate the prediction errors between the observed and predicted drawdowns. The ME is defined as

$$ME = \frac{1}{n} \sum_{i=1}^n e_i \tag{5.1}$$

The principle of least-squares assumes that the errors are normally distributed with zero mean and constant variance (McCuen, 1985). When the ME value is equal to or very close to zero, the assumption that errors having zero mean will be satisfied.

The SEE is defined as

$$SEE = \sqrt{\frac{1}{\nu} \sum_{i=1}^n e_i^2} \tag{5.2}$$

where ν is the degree of freedom, which is equal to the number of observed data points minus the number of estimate parameters.

5.3. Parameter identification for leaky aquifer without considering the aquitard storage

The initial guesses of the aquifer parameters T , S , and L are given in Table 3. For the initial error covariance matrix, $P_{k-1}(+)$, the diagonal elements could be assigned as, for example, 55,000, 1.0×10^{-9} and 1.0×10^{-6} , for T , S , and L , respectively; while and the off-diagonal elements of $P_{k-1}(+)$ are set to zero. There is no specific rule to decide the initial

Table 3

Initial guess values for EKF and the estimated parameters and prediction errors when using EKF, NLN, and graphical method to analyze Cooper’s data (Cooper, 1963) for leaky aquifer without considering the effect of aquitard storage

Three-parameter model							
Case no.	Initial guesses for hydraulic parameters			Initial error covariance matrix for hydraulic parameters			Step
	T	S	L	T	S	L	
1	1000	2.00×10^{-4}	1.00×10^{-2}	55,000	1.00×10^{-9}	1.00×10^{-3}	
2	1000	2.00×10^{-4}	1.00×10^{-1}	55,000	1.00×10^{-9}	2.00×10^{-2}	
3	1000	2.00×10^{-4}	1.00×10^{-1}	55,000	1.00×10^{-9}	1.00×10^{-2}	
Case no.	Estimated parameters			Prediction errors			Step
	T	S	L	ME	RMSE	SEE	
EKF on interpolated data							
1	1257.9	9.09×10^{-5}	4.82×10^{-2}	-6.53×10^{-4}	1.46×10^{-2}	1.69×10^{-2}	236
2	1311.4	9.29×10^{-5}	2.28×10^{-1}	3.72×10^{-3}	7.46×10^{-3}	8.62×10^{-3}	1508
3	1228.0	1.00×10^{-4}	5.08×10^{-1}	-2.44×10^{-4}	3.42×10^{-3}	4.09×10^{-3}	7603
EKF on recursively used data							
1	1239.4	9.78×10^{-5}	4.94×10^{-2}	1.56×10^{-4}	1.15×10^{-2}	1.33×10^{-2}	14,656
NLN							
1	1239.1	9.80×10^{-5}	4.94×10^{-2}	-1.10×10^{-4}	1.15×10^{-2}	1.33×10^{-2}	
2	1242.1	9.80×10^{-5}	2.52×10^{-1}	4.98×10^{-4}	4.93×10^{-3}	5.71×10^{-3}	
3	1215.2	9.70×10^{-5}	5.12×10^{-1}	-1.90×10^{-4}	3.22×10^{-3}	3.72×10^{-3}	
Graphical method							
1	1234.5	1.00×10^{-4}	5.00×10^{-2}	-2.15×10^{-4}	1.19×10^{-2}	1.38×10^{-2}	
2	1234.5	1.00×10^{-4}	2.50×10^{-1}	2.55×10^{-3}	9.10×10^{-3}	1.05×10^{-2}	
3	1234.5	1.00×10^{-4}	5.00×10^{-1}	2.33×10^{-3}	4.13×10^{-3}	4.77×10^{-3}	

The convergence criteria for EKF are $TOL_T = 1.0 \times 10^{-4}$, $TOL_S = 1.0 \times 10^{-8}$, and $TOL_L = 1.0 \times 10^{-6}$.

diagonal values. However, since $P_{k-1}(+)$ is equal to the expected value of squared estimation error $E[(\hat{x}_{k-1}(+) - x_{k-1})(\hat{x}_{k-1}(+) - x_{k-1})^T]$, a reasonable range of initial guess for $P_{k-1}(+)$ can be easily made. With the computation of EKF method to analyze the measured drawdown data obtained from observation well 1, the values of $P_{k-1}(+)$ are minimized, as are the estimation errors at each time step. The measurement covariance R_k is set to 10^{-5} and kept constant throughout the identification process. Reasonable magnitudes of the measurement error covariance range from 10^{-4} to 10^{-6} . Thus, the median value of R_k is selected in this study. Given the initial guess values for T , S , and L , along with the initial error covariance matrix, EKF coupled with Hantush and Jacob's model can identify the hydraulic parameters in a step-wise manner. The identification process is terminated when the specified convergence criteria are all satisfied, i.e. $TOL_T \leq 10^{-4}$, $TOL_S \leq 10^{-8}$, and $TOL_L \leq 10^{-6}$. The parameters, i.e. transmissivity, storage coefficient, and leakage coefficient, are successfully determined and achieve acceptable accuracy. The estimated three hydraulic parameters are $T=1257.9 \text{ m}^2/\text{day}$, $S=9.09 \times 10^{-5}$, $L=4.82 \times 10^{-2}$, and the related SEE value is 1.69×10^{-2} when analyzing the pumping test data in the observation well 1, $T=1311.4 \text{ m}^2/\text{day}$, $S=9.29 \times 10^{-5}$, $L=2.28 \times 10^{-2}$, and the related SEE value is 8.62×10^{-3} when analyzing the pumping test data in the observation well 2, and $T=1228.0 \text{ m}^2/\text{day}$, $S=100 \times 10^{-4}$, $L=5.08 \times 10^{-2}$, and the related SEE

value is 4.09×10^{-3} when analyzing the pumping test data in the observation well 3. The comparisons for the results when using EKF, NLN, and the graphical method to analyze drawdown data are listed in Table 3.

5.4. Parameter identification for leaky aquifer with considering the aquitard storage

The initial guesses of the aquifer parameters T , S , and L are listed in Table 4. With the off-diagonal elements set to zero, the diagonal elements of initial error covariance matrix $P_{k-1}(+)$ could have the values, for example, of 100, 10^{-10} , 10^{-4} and 4.0×10^{-7} , for T , S , L , and ψ , respectively. The values of error covariance matrix $P_{k-1}(+)$ can be assigned similarly to the descriptions in Section 5.3. The measurement covariance R_k is also set to 10^{-5} , and is kept constant throughout the identification process. With initial guesses for the hydraulic parameters, the application of EKF method to Neuman and Witherspoon's model can identify the values of those four hydraulic parameters step-by-step.

The four hydraulic parameters estimated by EKF on interpolated data, also shown in Table 4, are $T=22.6 \text{ m}^2/\text{day}$, $S=1.73 \times 10^{-4}$, $L=1.42 \times 10^{-1}$ and $\psi=3.16 \times 10^{-4}$; while the related prediction error of the SEE value is 1.37×10^{-2} . On the other hand, the estimated transmissivity, which is the most important hydraulic parameter, by the NLN is $23.3 \text{ m}^2/\text{day}$ and its SEE value is 1.06×10^{-2} .

Table 4

Initial guess values for EKF and estimated parameters and prediction errors when using EKF and NLN to analyze data reported in Sridharan et al. (1987) for leaky aquifer with considering the effect of aquitard storage

Method	Initial guesses for hydraulic parameters				Initial error covariance matrix for hydraulic parameters			
	T	S	L	ψ	T	S	L	ψ
	50	2.00×10^{-4}	2.00×10^{-1}	5.00×10^{-4}	100	1.00×10^{-10}	1.00×10^{-4}	4.00×10^{-7}
	Estimated parameters				Prediction errors			
	T	S	L	ψ	ME	RMSE	SEE	Step
EKF on interpolated data	22.6	1.73×10^{-4}	1.42×10^{-1}	3.16×10^{-4}	1.49×10^{-3}	1.08×10^{-2}	1.37×10^{-2}	839
EKF on recursively used data	23.3	1.64×10^{-4}	1.33×10^{-1}	3.19×10^{-4}	-5.28×10^{-5}	9.04×10^{-3}	1.13×10^{-2}	2001
NLN	23.3	1.65×10^{-4}	1.34×10^{-1}	7.04×10^{-4}	-1.54×10^{-3}	8.50×10^{-3}	1.06×10^{-2}	

The convergence criteria for EKF are $TOL_T=1.0 \times 10^{-5}$, $TOL_S=1.0 \times 10^{-9}$, $TOL_L=1.0 \times 10^{-6}$, and $TOL_\psi=1.0 \times 10^{-9}$.

5.5. Parameter identification while analyzing the pumping data being recursively used

As mentioned above, the EKF method can be applied to analyze the existing pumping data set in a repeated manner. In other words, the pumping data can be repeatedly used once the EKF has gone through all the available data. Again, the first data set of Cooper (1963) and the pumping data of Sridharan et al. (1987) are chosen for parameter estimations using the three-parameter and four-parameter models, respectively. The analyzed results are also demonstrated in Tables 3 and 4.

5.6. Discussion

Table 3 lists the estimated parameters and prediction errors when using EKF, NLN, and graphical method to analyze Cooper's data (1963) for leaky aquifer without storage effect. Compared with NLN and graphical method, EKF on interpolated data can quickly identify the parameters and achieve good accuracy, using only part of the measured drawdown data. The parameters, i.e. transmissivity, storage coefficient, and leakage coefficient, are successfully determined and achieve acceptable accuracy, compared with those estimated by NLN as indicated in Table 3 when analyzing Cooper's three drawdown data sets (1963). Note that both computer methods, EKF and NLN, give slightly better estimation for the hydraulic parameters judging from the prediction error of the SEE value. The parameters for observation well 1 is quickly determined and the identification process takes less than 236 steps, i.e. 23.6 min of the pumping time to get convergent results. An aquifer test with long pumping time is usually required for leaky aquifers by conventional graphical methods. The total pumping time for the recorded data are 1000 min (16.67 h). However, a long period of pumping time may not be necessary if the EKF is employed. In this case, EKF can save 97.6% of the time required to obtain results from the pumping test. The total numbers of steps for EKF to get convergent values of parameters when analyzing the drawdown data of the observation wells 2 and 3 are significantly larger than that of the observation well 1. This reflects that a larger distance from the observation well to the pumping well requires a longer

observation time. Furthermore, it is found that the estimate leakage coefficient for observation well 1 varies only marginally at the first few observations, but changes dramatically later on during the identification process. This may be attributed to the fact that there is some time lag between the start of pumping and the leakage effect influencing the drawdown.

Table 4 exhibits the estimated parameters and prediction errors when using EKF and NLN to analyze data reported in Sridharan et al. (1987) for the leaky aquifer considering the aquitard storage. The EKF on interpolated data gives results very close to the estimated parameters and prediction errors to those of NLN. There is a total of 839 steps used in the identification process, i.e. 83.9 min of the pumping time. On the other hand, the total pumping time for the observed data are 1200 min (20 h). Clearly, the EKF method can save 93.0% of the pumping test time in this case.

Fig. 1 shows the observed drawdown and the estimated drawdown of those three observation wells based on Hantush and Jacob's model. This indicates that the estimated drawdowns fit the pumping test data quite well. Fig. 2 exhibits that the plot of the observed drawdown and the estimated drawdown based on Neuman and Witherspoon's model. The figure implies that the parameters estimated by the EKF method are suitable to represent the hydraulic characteristic of the aquifer.

Tables 3 and 4 also display the results of estimated parameters and prediction errors while measurement data are recursively used in the EKF identification process using the three-parameter and four-parameter models, respectively. Both tables indicate that the EKF method gives slightly more accurate results, even though many more time steps are needed, due to the measurement data that is recursively used. The state vector defined in Eq. (2.2) is assumed to be noise-free and independent of time, so the state transition matrix, Eq. (2.10), is time-invariant and is an identity matrix. The associated measurement model, Eq. (2.3), assuming known error covariance structure, can be evaluated at any time. Thus, the EKF method can be applied for any sampling time, even for analyzing recursively used data. Nevertheless, there are two advantages involved if applying the EKF on the interpolated data with a smaller time interval. First, the measurement of the field-pumping test and the estimation of the hydraulic

Table 5

The relative error (%) of truncated leaky well function for various values of u and leakage coefficient L

u	L								
	0.01	0.02	0.05	0.1	0.2	0.25	0.5	1	
0.00001	0.08	–	–	–	–	–	–	–	
0.00005	0.22	0.22	–	–	–	–	–	–	
0.0001	0.24	0.26	93.57	–	–	–	–	–	
0.0002	0.26	0.27	0.58	–	–	–	–	–	
0.0005	0.29	0.29	0.33	27.01	–	–	–	–	
0.001	0.33	0.33	0.35	0.16	–	–	–	–	
0.002	0.39	0.38	0.39	0.43	37.40	–	–	–	
0.005	0.52	0.52	0.50	0.50	0.51	1.39	–	–	
0.01	0.74	0.73	0.70	0.66	0.66	0.68	–	–	
0.02	1.18	1.17	1.12	1.03	0.90	0.87	1.81	–	
0.05	2.80	2.78	2.70	2.50	2.09	1.90	1.63	–	
0.1	6.43	6.40	6.27	5.94	5.14	4.73	3.36	7.09	
0.2	17.18	17.14	16.91	16.33	14.80	13.94	10.27	12.25	
0.5	82.79	82.69	82.13	80.62	76.37	73.80	61.09	49.21	
1	–	–	–	–	–	–	–	–	

Symbol ‘–’ represents that the relative error is larger than 100%.

parameters can be done with the same on-line measurements in the field. Second, the EKF method usually takes significantly less steps in the identification process to get convergent results.

Chander et al. (1981) combined the truncated leaky well function with EKF to identify three hydraulic parameters, so the applicable ranges of the pumping time and leakage coefficient L for using their method may be limited. Table 5 shows the relative errors of truncated leaky well function versus various values of u and L , indicating that when the leakage coefficient is large, the applicable range of the truncated leaky well function is small if the relative error is restricted to be less than 5%. On the other hand, the relative errors are larger than 5% when $u > 0.1$ and $L < 0.2$. Clearly, Chander et al.’s method may yield significant errors in parameter estimation when the pumping test data do not fall in the appropriate range, as indicated in Table 5.

6. Analyses for model and data measurement uncertainties

In this study, the uncertainties of the measurement data are represented by white noise and temporally correlated noise, and sensitivity analyses for the EKF method are then performed for data with those two types of noise. The model uncertainty in the parameter

estimation is assessed for the case of employing both three-parameter and four-parameter models for analyzing three data sets.

6.1. Measurement data uncertainty

The MATLAB function `randn(m, n)` with $m = 200$ and $n = 1$ is first chosen to generate a realization of white noise (The MathWorks, 1995). The elements in this realization are normally distributed random numbers with zero mean and unit variance. Each element is then multiplied by 4.09×10^{-3} , which is taken from the case of the least prediction error (SEE values) by EKF for observation well 3 in the three-parameter model. Finally, 130 data points are taken from this realization and divided into 10 data sets, so each data set contains 13 elements (i.e. 13 data points). Hantush and Jacob’s model along with the given parameters, say $T = 1000 \text{ m}^3/\text{day}$, $S = 10^{-4}$, and $L = 0.1$, was employed to generate 13 synthetic sets of drawdown data. Thus, a set of synthetic drawdown data was generated by simply adding the adjusted noise data to the predicted drawdown data one by one. Accordingly, 10 sets of synthetic drawdown data with uncorrelated noise were obtained.

The original realization of white noise is employed to generate temporally correlated noise. The MATLAB

Table 6
The estimated parameters and prediction errors for drawdown data with correlated noises

Case	Estimated parameters			Prediction errors		
	T	S	L	ME	SEE	STEP
1	1033.97	9.26×10^{-5}	0.092	2.05×10^{-3}	1.02×10^{-2}	673
2	1019.79	9.49×10^{-5}	0.095	4.95×10^{-3}	7.62×10^{-3}	1655
3	1023.91	9.39×10^{-5}	0.094	5.39×10^{-3}	8.83×10^{-3}	911
4	1041.93	9.29×10^{-5}	0.089	6.05×10^{-3}	1.20×10^{-2}	300
5	1021.01	9.54×10^{-5}	0.095	3.89×10^{-3}	6.56×10^{-3}	352
6	1042.00	9.31×10^{-5}	0.090	2.72×10^{-3}	1.02×10^{-2}	312
7	1029.79	9.29×10^{-5}	0.092	6.06×10^{-3}	1.02×10^{-2}	346
8	1016.93	9.29×10^{-5}	0.096	7.35×10^{-3}	1.26×10^{-2}	918
9	1020.55	9.47×10^{-5}	0.095	5.61×10^{-3}	8.14×10^{-3}	1173
10	1039.74	9.22×10^{-5}	0.090	5.22×10^{-3}	1.09×10^{-2}	697

function $\text{Hamming}(n)$ with $n=5$ is used to produce five coefficients of a Hamming window (The MathWorks, 1995). The MATLAB function $\text{conv}(a, b)$ is applied to convolve vectors a and b (The MathWorks, 1995), where vector a represents the original realization and vector b represents the coefficients of the Hamming window. Algebraically, convolution is an operation which multiplies two polynomials with coefficients containing the elements of a and b . The product of the convolution is adjusted by a factor of 2.88×10^{-3} , which is calculated based on the data variance and the SEE value for the result of observation well 3. The result after the adjustment represents a new realization with temporal correlated noise. Thus, 10 sets of synthetic drawdown data with temporally correlated noise can be formed in a similar manner.

Table 6 lists the results of data uncertainty analysis for the case with correlated measurement noise.

The estimated parameter T ranges from 1016.93 to 1042 m³/day, S ranges from 9.22×10^{-5} to 9.54×10^{-5} , and L ranges from 8.88×10^{-2} to 9.60×10^{-2} . The SEE values of all cases are on the same order of magnitude. Table 7 shows similar results for data with the white noise. Thus, the effect of data with either white noise or temporal noise is negligible in the identification procedure.

6.2. Model uncertainty

The third type of uncertainty listed in Eisenberg et al. (1989) is the conceptual model uncertainty regarding to the geometrical configuration, major features, and boundary conditions. Generally speaking, field hydrogeologic information is never known in sufficient detail. Also, the development of a mathematical model usually depends on some

Table 7
The estimated parameters and prediction errors for drawdown data with white noises

Case	Estimated parameters			Prediction errors		
	T	S	L	ME	SEE	STEP
1	1033.97	9.24×10^{-5}	0.092	5.71×10^{-3}	1.06×10^{-2}	1033
2	1019.79	9.49×10^{-5}	0.094	5.89×10^{-3}	8.29×10^{-3}	316
3	1023.91	9.15×10^{-5}	0.090	8.60×10^{-3}	1.31×10^{-2}	400
4	1041.93	9.57×10^{-5}	0.096	1.75×10^{-3}	6.30×10^{-3}	496
5	1021.01	9.44×10^{-5}	0.095	2.39×10^{-3}	8.68×10^{-3}	1154
6	1042.00	9.62×10^{-5}	0.096	1.76×10^{-3}	5.01×10^{-3}	1587
7	1029.79	9.31×10^{-5}	0.092	5.53×10^{-3}	9.71×10^{-3}	1293
8	1016.93	9.42×10^{-5}	0.096	2.60×10^{-3}	9.50×10^{-3}	592
9	1020.55	9.65×10^{-5}	0.097	3.42×10^{-3}	5.73×10^{-3}	1132
10	1039.74	9.49×10^{-5}	0.095	5.65×10^{-3}	8.03×10^{-3}	1001

Table 8
Pumping test data (Batu, 1998, p. 265)

Time (min)	Drawdown (m)
0.00	0.000
48.96	0.055
61.63	0.062
75.60	0.066
101.09	0.073
145.44	0.090
192.96	0.098
270.72	0.108
361.44	0.116
449.28	0.122
576.00	0.135
652.32	0.136
767.52	0.137
881.28	0.138

$$Q = 625 \text{ m}^3/\text{day}, r = 105 \text{ m.}$$

assumptions and/or simplifications. Thus, the selection of a model for describing a target aquifer system is always subject to some degree of uncertainty.

Hantush and Jacob (1955) deals with a confined aquifer underlain by an impermeable bed and overlain by an unconfined aquifer; and an aquitard with a negligible storage exists between those two aquifers. In contrast, Neuman and Witherspoon's model (1969) is composed of two confined aquifers that are separated by an aquitard. Both of these models assume that the hydraulic head of the top aquifer maintains constant during the pumping, although Neuman and Witherspoon's model considers the effect of the aquitard storage on the drawdown of the pumped aquifer.

Three different data sets obtained from pumping tests in leaky aquifers are chosen to assess the difference due to model uncertainty in parameter estimation. The first two data sets, also mentioned above, are taken from Cooper (1963) and Sridharan et al. (1987), and the third data set is taken from Batu (1998, p. 265). The estimated parameters based on the three-parameter model are listed in Table 3, and those based on the four-parameter model were listed in Table 4. Accordingly, the storage coefficient of the aquitard, S' , can be calculated with known values of ψ for the four-parameter model; and the hydraulic conductivity of the aquitard K' can be estimated based on the known values of b' and L for both models, as listed in Table 10.

For Batu's pumping test data (1998), the confined aquifer is pumped at a uniform flow rate of $625 \text{ m}^3/\text{day}$. The observation well fully penetrating the aquifer completely is located 105 m away from the pumped well, and its drawdown data are given in Table 8. The thicknesses of the aquifer and aquitard are estimated to be 80 and 28 m, respectively. Walton's type-curve method was used to determine the hydraulic conductivity values of the aquifer and aquitard. Two different match points were selected, point A was arbitrarily chosen and point B was a point on the overlapping part of the curve, as described in Batu (1998). The estimated values of K , K' , and S , by Walton's type-curve method and the EKF method using the three-parameter model are shown in Table 9. The prediction errors of SEE and ME clearly indicate that the EKF has the better accuracy for the estimated hydraulic parameters than the graphical approach of Walton's type-curve method.

Table 9
The estimated parameters and prediction errors using EKF to analyze data reported in Batu (1998, p. 265)

Method	Three-parameter model					
	Estimated parameters			Prediction errors		
	T	S	L	ME	RMSE	SEE
Walton's type—curve A	875	2.7×10^{-3}	0.30	1.26×10^{-2}	1.38×10^{-2}	1.55×10^{-2}
Walton's type—curve B	899	2.8×10^{-3}	0.30	9.50×10^{-2}	1.06×10^{-2}	1.19×10^{-2}
EKF	907	3.5×10^{-3}	0.34	-2.68×10^{-3}	4.41×10^{-3}	4.97×10^{-3}

Method	Four-parameter model						
	Estimated parameters				Prediction errors		
	T	S	L	ψ	ME	RMSE	SEE
EKF	893	3.43×10^{-3}	0.34	4.88×10^{-3}	-1.51×10^{-3}	3.89×10^{-3}	4.61×10^{-3}

Table 10
Estimated hydraulic conductivity and storage coefficient for aquifers and aquitards

Model	Estimated parameters					Prediction errors	
	T	K	K'	S	S'	ME	SEE
(Cooper's data (1963, p. 54) ^a)							
Three-parameter	1257.9	N/A	9.59×10^{-2}	9.09×10^{-5}		-6.53×10^{-4}	1.99×10^{-2}
Four-parameter	1242.7	N/A	7.89×10^{-2}	1.16×10^{-4}	2.95×10^{-9}	-1.81×10^{-3}	5.79×10^{-2}
Sridharan's data (1987, p.170)							
Three-parameter	21.60	N/A	N/A	1.69×10^{-4}		-1.95×10^{-2}	2.73×10^{-2}
Four-parameter	22.62	N/A	N/A	1.73×10^{-4}	2.76×10^{-10}	1.49×10^{-3}	1.37×10^{-2}
Date reported in Batu (1998, p. 265) ^b							
Three-parameter	907	11.34	0.26	3.5×10^{-3}		-2.68×10^{-3}	4.97×10^{-3}
Four-parameter	893	11.16	0.26	3.43×10^{-3}	1.31×10^{-6}	-1.51×10^{-3}	4.61×10^{-3}

N/A represents that the data are not available because the aquifer thickness (b) or/and the aquitard thickness (b') is unknown.

^a $b' = 30.480$ m.

^b $b = 80$ m and $b' = 28$ m.

Again, Batu's data (1998) is analyzed by the EKF method for the four-parameter model. The estimated values of K , K' , S , and S' are also reported in Table 10. The four-parameter model gives slightly better fit than the three-parameter model for Sridharan's and Batu's pumping data sets, judging from the prediction errors shown in Table 10. From the viewpoint of curve-fitting, the four-parameter model is more flexible to fit the measurement data than the three-parameter model. However, the three-parameter model yields fewer prediction errors than the four-parameter model in the case of analyzing Cooper's data. This discrepancy may be caused by the fact that the top aquifers of the leaky aquifer system for Hantush and Jacob's model (1955) and Neuman and Witherspoon's model (1969) are different. The top aquifer in Hantush and Jacob's model is an unconfined aquifer, whereas in Neuman and Witherspoon's model it is a confined one. Different geological settings may produce different response of the groundwater flow system while aquifers are under pumping. For the four-parameter model, each estimated value of S' is at least three orders of magnitude less than the value of S in the same aquifer system for those three data sets as shown in the table. The SEE values for both the three-parameter and four-parameter models are on the same order of magnitude suggesting that the effect of S' is so small as to be negligible. Clearly, the three-parameter model is a good choice for representing a leaky aquifer system if $S'/S < 10^{-3}$.

7. Concluding remarks

A method using EKF and the Lagrangian polynomial is proposed to identify the hydraulic parameters in leaky aquifer systems. Hantush and Jacob's model combined with EKF can optimally determine the parameters for leaky aquifers without considering the aquitard storage step-by-step in the identification process. Neuman and Witherspoon's model can also be employed in a similar manner in leaky aquifer with considering the aquitard storage. The drawdown data with non-uniform time intervals from the pumping test are interpolated by the Lagrangian polynomial to obtain the data with a smaller time interval. This interpolation approach facilitates the implementation of the EKF to estimate the hydraulic parameters on-line in a field-pumping test.

The proposed approach can quickly identify the parameters, using only a few observed drawdown data, and the obtained parameters are shown to achieve good accuracy. At present, pumping tests are usually performed in the field with a system having pressure transducers installed in the observation wells to measure the water level and a data logger to store the measured data transmitted from the pressure transducers. Such a system may be linked to a computer in which the EKF coupled with the aquifer model is implemented and executed simultaneously. Accordingly, the hydraulic parameters can be determined on-line in the field. Once stable estimates of the hydraulic parameters have been reached, the pumping test may be terminated.

The analyses on Chander et al.'s truncated leaky well function show that their approach of using EKF to estimate the hydraulic parameters is only suitable for cases where the pumping time is large or/and the leakage coefficient is small. The identification process of EKF has been shown to be able to reflect the physical nature of the leaky aquifer. The slow response of the leakage coefficient may be attributed to the fact that there is some time lag between the start of pumping and the leakage effects influencing the drawdown.

Finally, the sensitivity test for the measurement noise indicates that the effect of both the white noise and the temporal correlated noise in measured drawdown is insignificant when employing the EKF to identify the hydrogeologic parameters of the leaky aquifers. The choice between Hantush and Jacob's model and Neuman and Witherspoon's model for representing the leaky aquifer system depends on the hydrogeological condition of the system indicated in the analyses of the model uncertainty. However, Hantush and Jacob's model is suggested for use if the ratio of the aquitard storage to the aquifer storage is less than 10^{-3} .

Acknowledgements

This study was partly supported by the Taiwan National Science Council under grant NSC90-2621-Z-009-003. The authors would like to thank the two anonymous reviewers for their valuable and constructive comments.

References

- Batu, V., 1998. *Aquifer Hydraulics*. Wiley, New York.
- Beck, M.E., 1987. Water quality modeling: a review of the analysis of uncertainty. *Water Resour. Res.* 23 (8), 1393–1442.
- Bierkens, M.F.P., 1998. Modeling water table fluctuations by means of a stochastic differential equation. *Water Resour. Res.* 34 (10), 2485–2499.
- Bierkens, M.F.P., Knotters, M., Hoogland, T., 2001. Space-time modeling of water depth using a regionalized time series model and the Kalman filter. *Water Resour. Res.* 37 (5), 1277–1290.
- Cahill, A.T., Ungaro, F., Parlange, M.B., Mata, M., Nielsen, D.R., 1999. Combined spatial and Kalman filter estimation of optimal soil hydraulic properties. *Water Resour. Res.* 35 (4), 1079–1088.
- Carnahan, B., Luther, H.A., Wilkes, J.O., 1969. *Applied Numerical Methods*. Wiley, New York.
- Chander, S., Kapoor, P.N., Goyal, S.K., 1981. Aquifer parameter estimation using Kalman filters. *J. Irrig. Drain. Div., Am. Soc. Civ. Eng.* 107 (IR1), 25–33.
- Christiaens, K., Feyen, J., 2001. Analysis of uncertainties associated with different methods to determine soil hydraulic properties and their propagation in the distributed hydrological MIKE SHE model. *J. Hydrol.* 246, 63–81.
- Cooper Jr., H.H., 1963. Type curves for nonsteady radial flow in an infinite leaky artesian aquifer, in: Bentall, Ray, compiler, Shortcuts and special problems in aquifer tests. US Geol. Surv. Water-Supply Pap. 1545-C 1963;.
- Dawson, K.J., Istok, J.D., 1991. *Aquifer Testing: Design and Analysis of Pumping and Slug Tests*. Lewis Publishers, Chelsea, MI.
- Eigbe, U., Beck, M.B., Wheeler, H.S., Hirano, F., 1998. Kalman filtering in groundwater flow modeling: problems and prospects. *Stochastic Hydrol. Hydraul.* 12, 15–32.
- Eisenberg, N.A., Rickertsen, L.D., Voss, C., 1989. Performance assessment, site characterization, and sensitivity and uncertainty methods: their necessary association for licensing, in: Buxton, B.E. (Ed.), *Proceedings of the Conference on Geostatistical, Sensitivity, and Uncertainty Methods for Ground-Water Flow and Radionuclide Transport Modeling*, CONF-870971. Battelle Press, Columbus, OH.
- Freeze, R.A., James, B., Massmann, J., Sperling, T., Smith, L., 1992. Hydrogeologic decision analysis. 4. The concept of data worth and its use in the development of site investigation strategies. *Ground Water* 30 (4), 574–588.
- Gerald, C.F., Wheatley, P.O., 1994. *Applied Numerical Analysis*, fifth ed. Addison-Wesley, Reading, MA.
- Grewal, M.S., Andrews, A.P., 1993. *Kalman Filtering: Theory and Practice*. Prentice-Hall, New Jersey.
- Hantush, M.S., 1960. Modification of the theory of leaky aquifers. *J. Geophys. Res.* 65 (11), 3713–3724.
- Hantush, M.S., Jacob, C.E., 1955. Non-steady radial flow in an infinite leaky aquifer. *Trans. Am. Geophys. Union* 36, 95–100.
- Jacob, C.E., 1946. Radial flow in a leaky artesian aquifer. *Trans. Am. Geophys. Union* 27, 198–208.
- Katul, G.G., Wendroth, O., Parlange, M.B., Puente, C.E., Folegatti, M.V., Nielsen, D.R., 1993. Estimation of in situ hydraulic conductivity function from nonlinear filtering theory. *Water Resour. Res.* 29 (4), 1063–1070.
- Kruseman, G.P., de Ridder, N.A., 1970. Analysis and evaluation of the pumping test data, Bulletin 11. Institute for Land Reclamation and Improvement, Wageningen, The Netherlands pp. 80–94.
- Lee, Y.P., Chin, C.V., Yen, P.H., 2000. Application of Kalman Filter for Real-time Identifying Aquifer Anisotropy, in: *Proceedings of the 11th Hydraulic Engineering Conference*, Taipei, Taiwan pp. D81–D86.
- Lohman, S.W., 1972. *Ground-water hydraulics*. US Geol. Surv. Prof. Pap. 1972, 708.
- McCuen, R.H., 1985. *Statistical Methods for Engineers*, Prentice Hall, Englewood Cliffs, New Jersey.

- Neuman, S.P., Witherspoon, P.A., 1969. Theory of flow in a confined two aquifer system. *Water Resour. Res.* 5 (4), 803–816.
- Sridharan, K., Ramaswamy, R., Lakshmana Rao, N.S., 1987. Identification of parameters in semiconfined aquifers. *J. Hydrol.* 93, 163–173.
- The MathWorks, 1995. The Student Edition of MATLAB, Version 4, User's Guide. Prentice-Hall, Englewood Cliffs, NJ.
- Van Geer, F.C., Van Der Kloet, P., 1985. Two algorithms for parameter estimation in groundwater flow problems. *J. Hydrol.* 77, 361–378.
- Van Geer, F.C., te Stroet, C.B.M., 1990. A Kalman filter approach to the quantification of the reliability of a groundwater model, in: Kovar, K. (Ed.), *Calibration and reliability in groundwater modeling ModelCare 90*, IAHS Publication No. 195, pp. 467–476.
- Van Geer, F.C., te Stroet, C.B.M., Bierkens, M.F.P., 1990. Groundwater modeling in relation to the system's response time using Kalman filtering, in: Gambolati, G., Rinaldi, A., Brebbia, C.A., Gray, W.G., Pinder, G.F. (Eds.), *Proceedings of the Eighth International Conference on Computational Methods in Water Resources and Surface Hydrology, Part A. Computational Mechanics Publication, Venice, Italy*, pp. 23–30.
- Wheater, H.S., Tompkins, J.A., van Leeuwen, M., Butler, A.P., 2000. Uncertainty in groundwater flow and transport modeling—a stochastic analysis of well-protection zone. *Hydrol. Process.* 14, 2019–2029.
- Yeh, H.D., 1987. Theis' solution by nonlinear least-squares and finite-difference Newton's method. *Ground Water* 25 (6), 710–715.
- Yeh, H.D., Han, H.Y., 1989. Numerical identification of parameters in leaky aquifers. *Ground Water* 27 (5), 655–663.
- Zheng, C., Bennett, G.D., 2002. *Applied Contaminant Transport Modeling*, second ed. Wiley/Interscience, New York.