A nudging data assimilation algorithm for the identification of groundwater pumping

Wei-Chen Cheng,[1] Donald R. Kendall,[2] Mario Putti,[3] and William W.-G. Yeh[1]

Received 21 November 2008; revised 26 May 2009; accepted 10 June 2009; published 26 August 2009.

[1] This study develops a nudging data assimilation algorithm for estimating unknown pumping from private wells in an aquifer system using measured data of hydraulic head. The proposed algorithm treats the unknown pumping as an additional sink term in the governing equation of groundwater flow and provides a consistent physical interpretation for pumping rate identification. The algorithm identifies the unknown pumping and, at the same time, reduces the forecast error in hydraulic heads. We apply the proposed algorithm to the Las Posas Groundwater Basin in southern California. We consider the following three pumping scenarios: constant pumping rates, spatially varying pumping rates, and temporally varying pumping rates. We also study the impact of head measurement errors on the proposed algorithm. In the case study we seek to estimate the six unknown pumping rates from private wells using head measurements from four observation wells. The results show an excellent rate of convergence for pumping estimation. The case study demonstrates the applicability, accuracy, and efficiency of the proposed data assimilation algorithm for the identification of unknown pumping in an aquifer system.

Citation: Cheng, W.-C., D. R. Kendall, M. Putti, and W. W.-G. Yeh (2009), A nudging data assimilation algorithm for the identification of groundwater pumping, *Water Resour. Res.*, *45*, W08434, doi:10.1029/2008WR007602.

1. Introduction

[2] Successful modeling of groundwater flow in aquifer systems relies not only on accurate calibration of model parameters, but also on the correct identification of the factors that force the flow, such as natural recharge and well pumping. While well locations and discharge generally are assumed to be known, estimation of recharge rates may be obtained as closure components of water balance equations for the entire hydrological system [*Scanlon et al.*, 2002; *Sanford*, 2002] or derived from groundwater level measurements [*Healy and Cook*, 2002]. The reliability of the estimations using a water budget obviously is limited by the accuracy of the different terms in the continuity equation, whereas methods that take advantage of head measurements in principle provide better accuracy. These methods generally are based on some kind of optimization, such as minimizing the least squares error where the error is defined as the difference between the model output and observation at selected locations.

[3] Similar to the identification of recharge rates, but much less studied, is the problem of estimating groundwater pumping rates and related uncertainty. In intensively exploited regions, groundwater extraction commonly is used to increase water supply reliability, in particular for agricultural or industrial purposes. Even though well location generally is known, the pumping rate as a function of time is often not available, particularly for private wells. In California, ''property rights do not require land owners in nonadjudicated basins to measure their groundwater pumping rates and publicly disclose them'' [*Harter*, 2003]. In fact, lack of data is not unique to California, but is a common problem across the entire United States and worldwide. A typical situation occurring during the development phase of a groundwater model is the existence of (almost) sufficient information regarding the number and location of production wells, but only partially known pumping rate data [*Ruud et al.*, 2004]. Even if pumping rates are reported, the frequency of reporting is generally insufficient.

[4] Estimation of groundwater pumping rates has been studied in a number of publications [*Koczot*, 1996; *Hanson et al.*, 2003; *Ruud et al.*, 2004; *Tung and Chou*, 2004; *Farrar et al.*, 2006; *Lin and Yeh*, 2008]. Methodologies based upon indirect evidence have been used in particular for estimating past extraction conditions. *Koczot* [1996] developed an estimation technique for agricultural basins in which the unknown pumping rates were estimated on the basis of water demand of the different crops. The technique later was applied successfully by *Hanson et al.* [2003] and *Farrar et al.* [2006]. The approach starts from land use maps and can be very useful in particular for the determination of past conditions. *Ruud et al.* [2004] considered groundwater extraction as a closure term in the global hydrological water balance, leading to an indirect estimation. This water balance–based approach has been applied to the southern San Joaquin Valley in California. As noted by the authors, the closure term approach may contain large errors because measurement data are seldom available and the pumping estimation is intrinsically indirect. *Tung and Chou* [2004] treated the estimation of average groundwater extraction and its spatial distribution as an inverse problem

[1]Department of Civil and Environmental Engineering, University of California, Los Angeles, California, USA.
[2]Calleguas Municipal Water District, Thousand Oaks, California, USA.
[3]Department of Mathematical Methods and Models for Scientific Applications, University of Padua, Padua, Italy.

0043-1397/09/2008WR007602$09.00

of parameter estimation. While this method may be globally accurate in terms of estimating the average extraction, it does not provide information regarding the pumping rate from each well. *Lin and Yeh* [2008] used simulated annealing to identify pumping source location, pumping rate, and pumping period for a single pumping well.

[5] Pumping rates generally are characterized by relatively high temporal variations, with scales ranging from hours to seasons. While hourly frequency is seldom employed in groundwater modeling, the seasonality of the rates must be captured accurately for better groundwater management. Hydraulic head variations represent the system response to time-varying forcing functions, such as extraction, injection, and recharge. Thus hydraulic head observations contain information about these forcing functions, although in an aggregate form. Decomposition of the head measurements into their generating factors is the object of this study, where we seek to estimate the pumping rates by minimizing some appropriate measure of the difference between the simulated and observed heads at specified locations. *Gehrels et al.* [1994] proposed a linear stochastic transfer function to separate natural and artificial components from the groundwater level fluctuations. In our study, we propose a data assimilation technique to separate the influence of pumping from each pumping well on each of the observation wells. We assume that both pumping and observation well locations are known and that a calibrated groundwater model is available. This is a common situation that may occur, for example, when the aquifer system is exploited by a number of private wells, when no discharge reporting is required, or when reporting is infrequent. The goal of our study is to utilize the available head observations and a calibrated groundwater model to estimate the well pumping rates as a function of time and with a frequency that is sufficient for management purposes.

[6] The problem at hand can be formulated in the context of a four-dimensional data assimilation (4DDA) algorithm, where the forecasted solution at a given time is changed in such a way that the new solution (the analysis) minimizes some statistical measure of the difference between the observed and simulated variables [*Nichols*, 2003]. It is a sequential method in that it consists of a forecast step followed by an update step, traditionally called an ''analysis'' [*Ide et al.*, 1997]. Data assimilation (DA) is an active field of research in hydrological sciences. However, it needs to be explored further, as it can be employed successfully in a variety of problems [*Troch et al.*, 2003]. Among the different 4DDA techniques, the simplest is Newtonian relaxation, or ''nudging'' [*Hoke and Anthes*, 1976; *Lorenc*, 1986]. Nudging is a DA technique that extracts information from observations and arbitrarily adds additional source/sink terms into the original model to drive model output toward the observed data. This added relaxation term incorporates four-dimensional (three spatial dimensions and one temporal dimension) interpolation of the measurement data and relaxes the model solution toward the observations [*Hoke and Anthes*, 1976; *Lorenc*, 1986; *Stauffer and Seaman*, 1990; *Stauffer et al.*, 1991]. Grid nudging and observation nudging are two major ways to apply the nudging technique. Grid nudging applies to cases in which observations are available at all model grid points, or a geometrically contiguous subset of model grid points. A typical application of grid nudging is the assimilation of remotely sensed data, generally available at the surface of the model domain. In contrast to grid nudging, observation nudging is applicable when observations are only available at sparse grid points [*Stauffer et al.*, 1991]. These two nudging techniques can be combined together if necessary [*Paniconi et al.* 2003].

[7] Nudging has been applied widely in atmosphere dynamics to assimilate different kinds of data, such as wind, temperature, and water vapor [*Hoke and Anthes*, 1976; *Lorenc*, 1986]; as well as in oceanographic modeling [*Verron*, 1990]. Such applications are given by *Miguez-Macho et al.* [2004], who studied spectral nudging for regional-scale climate simulations; by *Stiles et al.* [2002], who used nudging to obtain simulated ocean surface winds; and by *Pacione et al.* [2001], who implemented nudging to assimilate data on wind, temperature, and the mixing ratio for estimating precipitable water. Applications to hydrological problems are given by *Drusch* [2007], who applied nudging to assimilate satellite-derived surface soil moisture data in order to improve the results of a weather prediction model of temperature and relative humidity; and by *Pauwels et al.* [2001], who used nudging to assimilate remotely sensed soil moisture data for a land-atmosphere model in order to improve discharge prediction at the Zwalm catchment in Belgium. However, applications of DA to groundwater simulations are sparse and mostly related to the assimilation of unsaturated zone measurements. *Houser et al.* [1998] applied nudging to assimilate soil moisture data coming from remote sensing at the Walnut Gulch Experimental Watershed in southeastern Arizona. *Paniconi et al.* [2003] applied nudging to a small hypothetical catchment simulated with a coupled surface-subsurface flow model for the assimilation of both groundwater depth and soil moisture observations. *Hurkmans et al.* [2006] further extended the approach to a real catchment, the Brisy subcatchment located in southeastern Belgium.

[8] The nudging approach offers several attractive features for application to groundwater problems. It is easy to implement the algorithm for any existing groundwater flow simulator and its computational burden is sufficiently low to allow implementation in three-dimensional time-dependent models. Additionally and most importantly for our purposes, it provides a consistent and intuitive physical interpretation of the identified pumping rates, which are represented in the governing equation as sink terms. In this study we develop a nudging scheme to identify unknown pumping using observed hydraulic head data. We first describe the mathematical development of nudging and its application to the groundwater flow model. We then implement the proposed technique to a real world problem where we seek to estimate the unknown private well pumping rates in the Las Posas groundwater basin in Ventura County, California (USA).

2. Application of Nudging to the Groundwater Equation

2.1. Groundwater Flow Model

[9] The three-dimensional groundwater equation for saturated flow in a confined aquifer can be written as [*Bear*, 1988; *Willis and Yeh*, 1987]

$$S_s \frac{\partial h}{\partial t} = div[K \nabla h] + \overline{q}(x,t) \tag{1}$$

where S_s is the specific storage (L^{-1}), $h(x, t)$ is the piezometric head (L), x (L) is the vector of three-dimensional spatial coordinates, t (T) is time, K is the hydraulic conductivity tensor (LT^{-1}), and $\overline{q}$ (T^{-1}) represents the source or sink term. The effect of pumping and injection wells within the aquifer system may be simulated by representing the wells (or well clusters) as point sources or point sinks. In the case of a number of pumping wells (or well clusters), the sink term $\overline{q}$ (x, t) can be specified as

$$\overline{q}(x,t) = \sum_{i=1}^{n_w} \overline{q}_{w,i}(t)\delta(x_{w,i}) \tag{2}$$

where $\overline{q}_{w,i}$ (t) (T^{-1}) represents the time varying extraction rate of the ith pumping well, located at coordinate $x_{w,i}$; $\delta(x_{w,i})$ is the Dirac delta function, equal to one if $x = x_{w,i}$ and zero otherwise; and n_w denotes the total number of wells (or well clusters). Appropriate initial and boundary conditions complete the formulation of the mathematical model. In this paper we adopt MODFLOW [*Harbaugh et al.*, 2000] to simulate groundwater flow. However, our proposed algorithm is independent of the simulator and can be applied to any other numerical scheme used to discretize equation (1). MODFLOW uses an integrated block-centered finite difference approach to achieve spatial discretization with a first-order backward Euler finite difference scheme for the time discretization [*McDonald and Harbaugh*, 1988]. Following the integrated finite difference approach, equation (1) is solved in an integral form; i.e., it first is integrated in space over the entire domain, and then the intercell fluxes are discretized by finite differences. Formally, we can write the fully discretized system (in space and time) as

$$P\frac{h_{k+1} - h_k}{\Delta t} = Lh_{k+1} + q_{k+1} \tag{3}$$

where $h \in R^n$ is the vector of head cell values, n is the total number of cells in the numerical model; k is the time step index; Δt is the time step; P and L are the discretization operators, with dimension (L^2) and (L^2/T), respectively; and $q_{k+1} \in R^n$ is the sink term vector at time t_{k+1} and has dimension (L^3/T). Following standard MODFLOW terminology, Δt is the time step used in numerical integration; and the stress period (T) is the time during which the source/sink term q_{k+1} is kept constant (it is generally multiples of Δt). To simplify the discussion we assume that only Dirichlet boundary conditions are specified and their implementation are included directly into the matrix L. Other types of boundary conditions can be treated with no additional mathematical complexity. We also assume that there is no other recharge or discharge in the aquifer system except for the specified wells. Inclusion of these terms into the formulation is straightforward but would make the derivations more cumbersome and cause distractions. We note that this is by no means a limitation of the proposed nudging scheme; rather, this assumption simplifies its mathematical development.

2.2. Nudging Equation

[10] Standard nudging uses observations of the state variable to nudge the state variable of the model's output. In our groundwater model, the state variable is the hydraulic head. However, for pumping estimation, instead of nudging the state variable we use observations to nudge the unknown forcing function. We add an artificial nudging term as an additional forcing term in the governing equation. This additional term measures the head difference between the observations and model output, which is caused by the unknown forcing function. Multiple unknown pumpings may contribute to the head difference at a single observation well. We therefore must disseminate the aggregated information of head into the pumping rate at each individual pumping well. In this study, the number of observation wells is smaller than the number of pumping wells. As a result, an inverse problem formulation would be notoriously ill posed [*Yeh*, 1986].

[11] Nudging can be included directly into the discretized groundwater model as an additional sink term so that (3) becomes [*Auroux and Blum*, 2008]

$$P\frac{h_{k+1} - h_k}{\Delta t} = Lh_{k+1} + q_{k+1} + G(h^o - Ch_{k+1}) \tag{4}$$

where G (L^2/T) is an $(n \times n_o)$ matrix (called the gain); n_o is the number of available observations, stored in the n_o-dimensional vector h^o; and C (-) is an appropriate $(n_o \times n)$ matrix used to project h_{k+1} onto the observation space. In the original nudging procedure proposed by *Hoke and Anthes* [1976], the gain G is actually a scalar quantity representing the time of relaxation toward the observed state h^o. In other words, the forecast state will tend toward the observed state in a time proportional to $1/G$. More generally, G may represent the relative strength of the different processes in the mathematical model in order to numerically balance the assimilation of the different variables [*Stauffer and Seaman*, 1990]. A large value of G implies that more weight is given to the observations so that the model solution will quickly match the observations. *Vidard et al.* [2003] formally optimized G by treating it as a parameter estimation problem. The results show that nudging actually can be related to the Kalman filter-based data assimilation technique. In fact, assuming that L is a symmetric and positive definite matrix, it is easy to see that if we take $G = C^T R^{-1}$, where R is the covariance matrix of the observation error, then equation (4) is actually the Euler-Lagrange equation, i.e., the necessary and sufficient condition (because matrix L is symmetric) for the solution of the following energy minimization problem [*Auroux and Blum*, 2008]:

$$\min_h \left[\frac{1}{2}(h - h_k)^T P(h - h_k) - \frac{\Delta t}{2}h^T(Lh + 2q_{k+1}) + \frac{\Delta t}{2}(h^o - Ch)^T R^{-1}(h^o - Ch)\right]. \tag{5}$$

Therefore, solution of (5) is also the solution of (4), which in turn is the numerical approximation of the weak solution of (1) augmented with the nudging sink term $G(h^o - Ch_{k+1})$ [*Quarteroni and Valli*, 1997]. Nudging actually can be thought of as a four-dimensional variational data assimilation scheme as well as a special case of a Kalman filter [*Li and Navon*, 2001; *Vidard et al.*, 2003].

[12] Our goal is to modify the standard nudging approach for pumping estimation using head observations. More

precisely, we want to evaluate the vector of well pumping rates $q_w = \{q_{w,i}\}$, $i = 1, \ldots, n_w$ of equation (2), given the model (3) and the observation vector $h^o = \{h_i^o\}$, $i = 1, \ldots, n_o$. We assume that head measurements are available at times t_u, $u = 1, .., n_t$ (also called the update times), which may be different from both the numerical time steps and the model stress periods.

[13] Before going into the details of the derivation of the nudging term, we briefly review some nomenclature used in the DA literature [*Ide et al.*, 1997]. The n-dimensional vector h_u^f indicates the forecast, i.e., the result from model prediction at time t_u without the contribution of the nudging term; the n-dimensional vector h_u^a indicates the analysis, i.e., the forecast as modified by the DA algorithm; while the n_o-dimensional vector h_u^o is the vector of observations, always at time t_u. The difference between observation and forecast at time t_u at observation well i is grouped into the n_o-dimensional vector d_u, known as the innovation residual:

$$d_u = h_u^o - Ch_u^f. \tag{6}$$

In our case, the operator C is a matrix with elements of either zero or one. The matrix projects the model output onto the observation space. Using this notation, the final nudging equation can be written as (see Appendix A)

$$h_{u+1}^a = F_u(h_u) + \overline{G}\left(h_{u+1}^o - Ch_{u+1}^f\right) = F_u(h_u) + \overline{G}\, d_{u+1} \tag{7}$$

where we assume that m time steps of equation (4) are performed between times t_u and t_{u+1}, possibly encompassing several stress periods. The operator F_u and matrix $\overline{G}$ (-) can be derived from equation (4) (see Appendix A) and includes the sink term vector q_u updated in the previous nudging steps. Note that the forecast appears on the right-hand side of (7) while the analysis appears on the left-hand side, both evaluated at the same time t_{u+1}. Hence the actual DA algorithm described by (7) proceeds sequentially in exactly two steps. In the first step, the forecast h_{u+1}^f is calculated by solving (4) using the previously estimated pumping rates, i.e., without any contribution from the nudging term. In the second step, the forecast is used to determine the innovation residual d_{u+1} and evaluate the pumping mismatch, i.e., the updated nudging term, and finally the analysis h_{u+1}^a is obtained from equation (7).

2.3. Evaluation of the Nudging Term

[14] We start from the formulation known as "observation nudging," in which each grid point is nudged using a weighted average of residuals calculated from observations that are within a specified radius of influence defined in both space and time. In contrast to weather prediction models and because of the dissipative nature of the groundwater flow equation, we are not concerned about the danger of introducing sudden discontinuities into the forcing function, as these discontinuities (if any) will be dissipated exponentially in time and will not generate spurious solutions.

[15] Before discussing the derivation of the approach, we note that the nudging vector Gd_{u+1} physically represents an additional sink term introduced in the governing equation. Thus the innovation residual can be thought of as the cumulative drawdown due to the (unaccounted) pumping rates at the pumping wells. We thus need to transform this additional drawdown into pumping rates that, when added to the sink term, will drive the analysis closer to the measurements. Because observations and pumping wells are not located within the same computational cell, the definition of the nudging term requires a double spatial interpolation: first we need to separate the contributions of the different pumping wells from the observed drawdown, and then we must apportion the pumping evaluated from the drawdown to the different pumping wells. The innovation residual at observation well i is thus represented as

$$d_{u+1,i} = h_{u+1.i}^o - \sum_{s=1}^{n} c_{i,s} h_{u+1,s}^f = \sum_{j=1}^{n_w} \Delta h_{i,j} \tag{8}$$

where $c_{i,s}$ is equal to one if cell s contains the observation well, and zero otherwise. The value $\Delta h_{i,j}$ (where for simplicity we have omitted the subscript $u + 1$) represents the fraction of the total drawdown, measured at observation well i , that is caused by pumping well j. We evaluate $\Delta h_{i,j}$ by interpolation using a distance-weighted function so that the drawdown observed at observation well i is assumed to be caused only by the pumping wells that are located within a specified radius of influence, i.e., $\Delta h_{i,j} = w'_{i,j}\, d_{u+1,i}$, where $w'_{i,j}$ is the value of the interpolation coefficient, or distribution ratio, calculated using the coordinates of wells i and j. The total observed drawdown mismatch at observation well i is caused by all the pumping wells located within its radius of influence and can be written as

$$d_{u+1,i} = \sum_{j=1}^{n_w} \Delta h_{i,j} = \sum_{j=1}^{n_w} \left(w'_{i,j} \cdot d_{u+1,i}\right). \tag{9}$$

Note that the distribution ratios must be nonnegative and sum up to one, i.e.,

$$\sum_{j=1}^{n_w} w'_{i,j} = 1 \text{ and } 0 \le w'_{i,j} \le 1. \tag{10}$$

The fraction $\Delta h_{i,j}$ is transformed into a pumping rate using the influence coefficient method [*Becker and Yeh*, 1972; *Yeh*, 1986]:

$$\Delta q_{i,j} = \frac{1}{r_{i,j}} \Delta h_{i,j} = \frac{\partial q_j}{\partial h_i} \Delta h_{i,j} \tag{11}$$

where we make use of the inverse of the influence coefficient $r_{i,j} = \partial h_i / \partial q_j$, which indicates the response of the head at cell i because of a unit change of the specific discharge at cell j. The influence coefficient $r_{i,j}$, also known as the response coefficient or Jacobian sensitivity coefficient, can be obtained by a sensitivity analysis [*Yeh*, 1986]. The quantity $\Delta q_{i,j}$ is thus the estimate of the pumping rate increment (or decrement) of pumping well j responsible for the portion of the drawdown mismatch $\Delta h_{i,j}$ at observation i. As a consequence, each portion of the drawdown mismatch of an observation well is converted to the corresponding pumping rate increment (or decrement) for each of the pumping wells located within the radius of influence. Accordingly, multiple

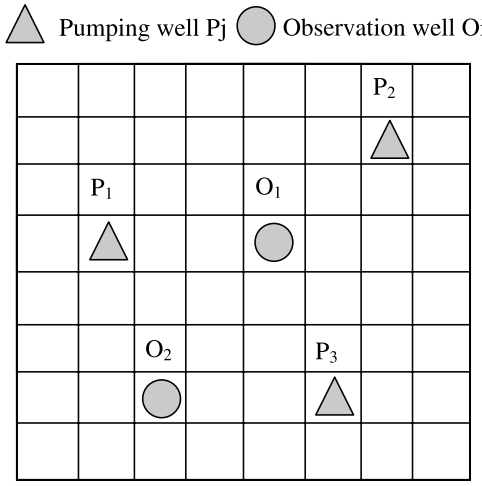

Figure 1. Illustration of pumping wells and observation wells.

observation wells, located within the radius of influence of a pumping well, will generate the same number of pumping estimations for the pumping well. The final estimate $\Delta q'_{u+1,j}$ of the pumping rate mismatch at well j is calculated by an additional weighted average with weights $w_{i,j}$* to integrate all pumping estimations into one value:

$$\Delta q'_{u+1,j} = \sum_{i=1}^{n_o} w^*_{i,j} \Delta q_{i,j}. \tag{12}$$

This n_w-dimensional vector of updated pumping rates $\Delta q'_{u+1,j}$ needs to be projected back to the state space to yield the final expression of the nudging vector of equation (4):

$$G d_{u+1} = g C^T \Delta q'_{u+1} \tag{13}$$

where the scalar g is the scalar gain relaxation time as defined previously.

[16] We present an example to demonstrate the concept of pumping estimation. Figure 1 shows a hypothetical MODFLOW grid with three pumping wells and two observation wells. We assume that the pumping rates of all three pumping wells are unknown. The drawdown at observation wells O_1 and O_2 is due to the extraction from pumping wells P_1, P_2, and P_3. The extraction of each pumping well contributes a portion of the drawdown at the two observation wells. The innovation residual at observation wells O_1 and O_2 can be expressed as

$$\begin{aligned} d_{O1} &= \Delta h_{O1,P1} + \Delta h_{O1,P2} + \Delta h_{O1,P3} \\ &= w'_{O1,P1} \cdot d_{O1} + w'_{O1,P2} \cdot d_{O1} + w'_{O1,P3} \cdot d_{O1} \end{aligned} \tag{14}$$

$$\begin{aligned} d_{O2} &= \Delta h_{O2,P1} + \Delta h_{O2,P2} + \Delta h_{O2,P3} \\ &= w'_{O2,P1} \cdot d_{O2} + w'_{O2,P2} \cdot d_{O2} + w'_{O2,P3} \cdot d_{O2} \end{aligned} \tag{15}$$

where $\Delta h_{i,j}$ is the estimated portion of drawdown mismatch at observation well i due to pumping well j. Note that $w'_{O1,P1} + w'_{O1,P2} + w'_{O1,P3} = 1$ and $w'_{O2,P1} + w'_{O2,P2} + w'_{O2,P3} = 1 (\sum_{j=1} w'_{i,j} = 1)$. The estimated pumping rate $\Delta q_{i,j}$ that causes $\Delta h_{i,j}$ is calculated by equation (11).

$$\Delta q_{O1,P1} = \frac{1}{r_{O1,P1}} \cdot w'_{O1,P1} \cdot d_{O1} \tag{16}$$

$$\Delta q_{O1,P2} = \frac{1}{r_{O1,P2}} \cdot w'_{O1,P2} \cdot d_{O1} \tag{17}$$

$$\Delta q_{O1,P3} = \frac{1}{r_{O1,P3}} \cdot w'_{O1,P3} \cdot d_{O1} \tag{18}$$

$$\Delta q_{O2,P1} = \frac{1}{r_{O2,P1}} \cdot w'_{O2,P1} \cdot d_{O2} \tag{19}$$

$$\Delta q_{O2,P2} = \frac{1}{r_{O2,P2}} \cdot w'_{O2,P2} \cdot d_{O2} \tag{20}$$

$$\Delta q_{O2,P3} = \frac{1}{r_{O2,P3}} \cdot w'_{O2,P3} \cdot d_{O2} \tag{21}$$

Each observation well provides a pumping estimation for each pumping well. For example, equations (16) and (19) provide a pumping estimation for pumping well P1 using observations at O1 and O2, respectively. Each estimated pumping rate increment (or decrement) now must be integrated into a particular well:

$$\Delta q'_{P1} = g \cdot \left(w^*_{O1,P1} \Delta q_{O1,P1} + w^*_{O2,P1} \Delta q_{O2,P1} \right) \tag{22}$$

$$\Delta q'_{P2} = g \cdot \left(w^*_{O1,P2} \Delta q_{O1,P2} + w^*_{O2,P2} \Delta q_{O2,P2} \right) \tag{23}$$

$$\Delta q'_{P3} = g \cdot \left(w^*_{O1,P3} \Delta q_{O1,P3} + w^*_{O2,P3} \Delta q_{O2,P3} \right) \tag{24}$$

We accomplish this by using the weight $w_{i,j}$*, which is calculated using the radius of influence of the pumping well instead of the observation well. The result then is multiplied by the scalar strength factor g.

2.4. Distribution Ratios and Weighting Coefficients

[17] It is well known that the drawdown caused by a pumping well tends to approach zero logarithmically with the distance from the well. We approximate this behavior by defining the distribution ratio $w'_{i,j}$ and the weighting coefficient $w_{i,j}$* using Cressman-type (distance-weighted) functions [*Stauffer and Seaman*, 1990; *Houser et al.*, 1998; *Paniconi et al.*, 2003]. If we let $\mathrm{X}_i = [x_i, y_i, z_i]^T$ represent the spatial coordinate of observation well i and $\mathrm{X}_j = [x_j, y_j, z_j]^T$ represent the spatial coordinate of pumping well j, we can write the following expressions for $w'_{i,j}$ and $w_{i,j}$*:

$$w_{i,j} = w(\mathrm{X}_i, \mathrm{X}_j, t) = w_1(D_{i,j}) w_2(z_i, z_j) w_3(t) \tag{25}$$

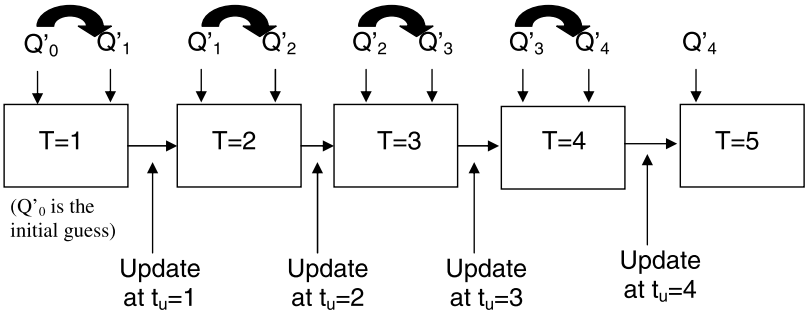

Figure 2. Illustration of sequential simulation model runs and the updating processes.

$$\begin{cases} w_1(D_{i,j}) = \left(R_i^2 - D_{i,j}^2\right)/\left(R_i^2 + D_{i,j}^2\right), & \mathrm{D}_{i,j}^2 \le R_i^2 \\ w_1(D_{i,j}) = 0, & \mathrm{D}_{i,j}^2 > R_i^2 \end{cases} \tag{26}$$

$$D_{i,j}^2 = (x_i - x_j)^2 + (y_i - y_j)^2 \tag{27}$$

$$\begin{cases} w_2(z_i, z_j) = 1, & z_i = z_j \\ w_2(z_i, z_j) = 0, & z_i \neq z_j. \end{cases} \tag{28}$$

$$\begin{cases} w_3(t) = 1, & t = t_u \\ w_3(t) = 0, & \text{Otherwise} \end{cases} \tag{29}$$

$$w'_{i,j} = \frac{w_{i,j}}{\sum_{j=1}^{n_w} w_{i,j}} \tag{30}$$

$$w^*_{i,j} = \frac{w_{i,j}}{\sum_{i=1}^{n_o} w_{i,j}} \tag{31}$$

where R_i (L) is the radius of influence for observation well i, and z_i and z_j are the aquifer layers where the pumping wells and observation wells are located. The definition of w_3 (t) shows that the nudging term is active only at the update time.

2.5. Schematic Representation of the Nudging Update Procedure

[18] The nudging update procedure is illustrated graphically in Figure 2 for a case in which the end of each stress period coincides with the update time. The symbol T in each rectangular box denotes the stress period between two successive updates t_u and t_{u+1}. The update time t_{u+1} is at the end of each stress period T. For example, the update time $t_u = 2$ is at the end of stress period T = 2. We note that each stress period generally contains multiple numerical integration time steps. We assume that the observation data are available at the end of each stress period, or at the update time. One update process includes two steps. The first step estimates the unknown pumping rate Q'_T for the previous stress period T using observation data obtained at the update time t_{u+1}. The second step reruns the simulation model using the estimated pumping rate Q'_T (treated as an additional sink term in equation (4); see also equations (A3) and (A5)) to evaluate the analysis, i.e., the updated hydraulic head distribution at update time t_{u+1}, and the forecast for the next stress period T + 1. In our implementation, the pumping rate estimation is performed offline and at the end of each stress period. Note that nudging can also be carried out in real time once new observations are available. Below we provide a step-by-step procedure for the proposed nudging algorithm.

[19] Step 0 is to calculate influence coefficients, distribution ratios and weighting factors. The distribution ratio $w'_{i,j}$ and weighting factor $w_{I,j}$* can be calculated with equations (30) and (31) using the given radius of influence and the distance between observation well i and pumping well j. The influence coefficient, $r_{i,j}$ ($=\partial h_i/\partial q_j$), can be obtained from a sensitivity analysis. For a nonlinear system, the influence coefficients should be updated at each update time. With the linearity assumption, the influence coefficients remain constant for each stress period and updating is not necessary. In this case, the sensitivity analysis only has to be carried out once. The influence coefficients can be expressed as

$$r_{i,j} = \partial h_{i,t_u=1}/\partial q_{j,T=1} = \partial h_{i,t_u=2}/\partial q_{j,T=2} = \ldots = \partial h_{i,t_u=e}/\partial q_{j,T=e} \tag{32}$$

where e is the total number of stress periods during the planning horizon. The influence coefficient for a given stress period (T) is the head response at observation well i at the end of stress period (t_{u+1}) due to a unit change in the discharge rate at pumping well j during the stress period.

[20] Step 1 is to forecast hydraulic head at observation well for the next update time. The forecast run is conducted with a simulation run using the known pumping data as well as the estimated pumping data. For the first stress period, the initial pumping estimation for the unknown pumping rate is set to zero. The forecast run provides head distribution estimates for the study area.

[21] Step 2 is to calculate the innovation residual. Equation (8) calculates the innovation residual, the head difference between observation and forecast, at observation wells at the update time.

[22] Step 3 is to estimate pumping rates. Using information obtained from the previous steps, step 3 estimates the nudging term, i.e., the new pumping rate correction, using equation (13).

[23] Step 4 is to update hydraulic head at the update time (the analysis). Step 4 performs a simulation run using the

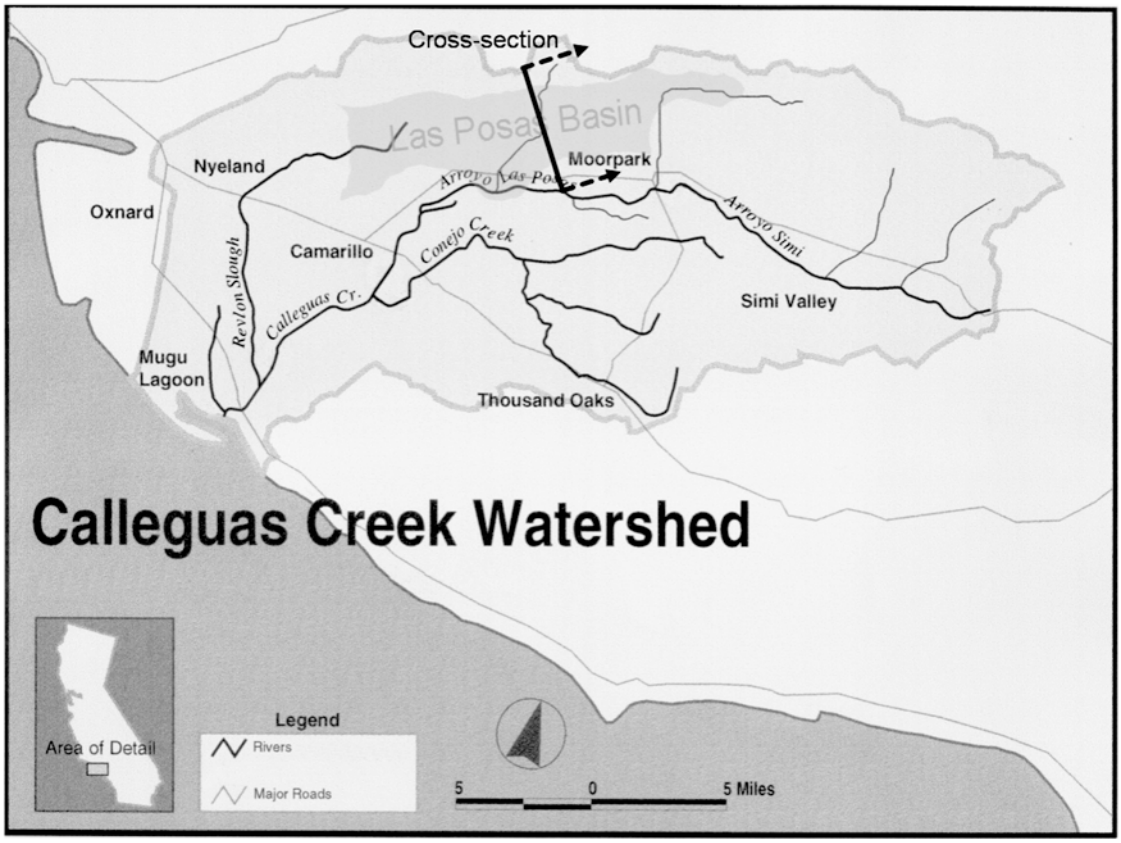

Figure 3. Planar view of the Las Posas Basin.

estimated pumping rates (for the previous stress period) to update the hydraulic head distribution at the update time. Steps 1–4 are repeated until the end of the planning horizon.

3. Application of the Nudging Technique to the Las Posas Groundwater Basin, California

[24] We apply the proposed nudging method to a realistic case study associated with the Las Posas groundwater basin in southern California. We seek to estimate the unknown pumping rates from a set of private wells in the basin. The basin is managed by the Calleguas Municipal Water District (CMWD). Although we select a real groundwater basin in our case study, we use numerical experiments to validate the proposed methodology and demonstrate the rate of convergence and utility of the proposed nudging algorithm. Well locations and pumping rates are imposed on the real aquifer geometry to generate the "true" head distribution that provides the observed head data for the case study. Identification errors can be obtained by calculating the difference between the analysis and the true solution so that the performance of the proposed algorithm can be analyzed. A key advantage of using synthetically generated data on a real system to test the proposed algorithm is that it avoids model errors and other uncertainties that may exist. The results should provide practical information and confidence in real-world applications.

3.1. Model Problem

[25] The Calleguas Municipal Water District (CMWD) operates a conjunctive use project in the Las Posas Basin, located near the city of Moorpark, California (Figure 3). The Las Posas Aquifer Storage and Recovery (ASR) Project stores treated surplus water in a 300 m deep confined aquifer system, which is then used as a strategic reservoir for dry periods. The operational plant is equipped with several dual-purpose injection and extraction wells that are clustered in two different well fields. The CMWD utilizes a calibrated MODFLOW model of the basin to manage its ASR project under a variety of pumping and injection scenarios.

[26] The Las Posas Groundwater Basin is about 32 km long (east–west) and 11 km wide (north–south) (Figure 3). It partially underlies the Las Posas Valley, located in southern Ventura County, California. The basin is bounded to the south by the Camarillo and Las Posas Hills, to the north by the South Mountain and Oak Ridge chains, to the east by the Santa Susana Mountains, and to the west by the Oxnard subbasin of the Santa Clara River Valley. Ground surface elevations range from about 60 m above mean sea level (AMSL) on the western boundary to about 215 m AMSL toward the eastern boundary. Average annual precipitation in this area ranges between 300 and 400 mm/a. The water-bearing geological formations that are of interest for this study include the alluvium, where recharge occurs, the San Pedro formation and the Santa Barbara formation. Productive aquifers include an upper unconfined unit, only sparsely exploited and coinciding with the alluvium, and two deeper confined aquifers that extend throughout the basin [*California State Water Resources Board*, 1956, *Hanson et al.*, 2003] known as the Fox Canyon aquifer and the Grimes Canyon aquifer (Figure 4). These two

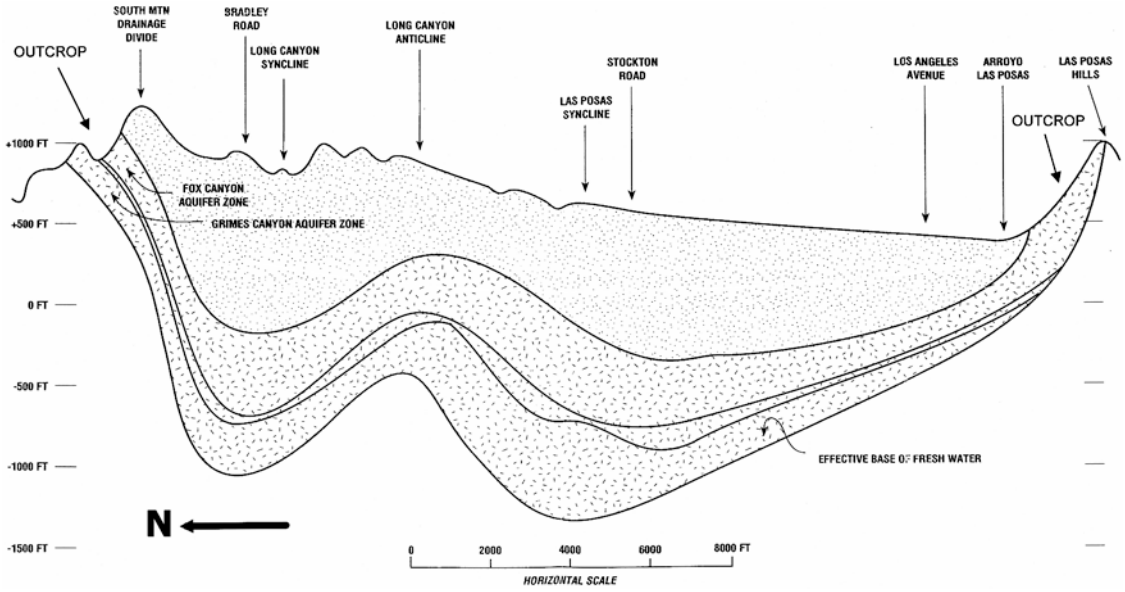

Figure 4. Cross section of the Las Posas Basin.

aquifers are separated by a small, unnamed aquifer. In the confined formations, well yield ranges between 1.5×10^{-2} m^3/s and 4.8×10^{-2} m^3/s [*California State Water Resources Board*, 1956; *Hanson et al.*, 2003]. A southwest to northeast trending fault (the central Las Posas fault) practically subdivides the aquifer into two distinct subsystems: the Eastern and the Western Las Posas basins. A distinct change in groundwater levels in the western-central part of the basin is noticeable in the wellhead data, and is likely due to a further buried and unnamed fault. Recharge to the basin comes predominantly from percolation of precipitation water across the outcrops of the Fox Canyon and Grimes Canyon gravels that are located at the northern and southern border of the basin, and from infiltration of imported water released into the Arroyo Las Posas by treatment plants.

[27] The simulation model consists of three confined layers and is defined aerially by 51×104 square cells, each with an area of 305.4×305.4 m^2. Fixed head boundary conditions are set on the east and west boundaries for the first layer, while no flow conditions are imposed on the east and west boundaries for the second and third layers as well as the north and south boundaries for all three layers. The first (shallowest) layer is a leaky aquifer, simulated by means of the general head boundary condition package provided by MODFLOW. In addition to leakage, surface recharge areas are positioned at the north and south boundaries at the location of the aquifer outcrops (Figures 3 and 4). The central Las Posas fault is represented as a horizontal-flow barrier with a hydraulic characteristic determined during model calibration. The computational domain contains 14 zones (values) of calibrated hydraulic conductivities, characterized by a vertical anisotropy ratio of 1/10. The top layer contains 14 zones of hydraulic conductivities. There are three zones in the second layer and two zones in the third layer. The values of conductivities in the second and third zones coincide with zones in the top layer. This model is a subset of a larger model of the Santa Clara–Calleguas basin developed by *Hanson et al.* [2003], to which the reader is referred for more details.

3.2. Setup of the Test Cases

[28] We distribute six well clusters (P1 to P6) and four observation wells (O1 to O4) on both sides of the fault and use nine 6-month stress periods as the simulation horizon. Figure 5 shows the spatial locations of the pumping and observation wells as well as the ASR well field (a cluster of wells) within the calibrated MODFLOW model. The northeast–southwest line represents the fault which, as noted, acts as a hydraulic barrier practically separating the aquifer system into two parts. The radius of influence for the nudging algorithm is set at $R = 7620$ m (equivalent to 25 cells). This setting implies that the identification of the pumping rates for private wells P1, P2, and P3 is based on the measurements taken at observation wells O1 and O2, while the measurements from O3 and O4 are used for identifying the pumping rates for wells P4 to P6. From the spatial distribution of the wells we see that in the western part of the domain the two observation wells (O1 and O2) are spaced uniformly away from the pumping wells: O1 is approximately equidistant from wells P1 and P2, while O2 is located approximately the same distance from wells P2 and P3. To the east of the fault, observation well O3 is centrally located with respect to all wells, but observation well O4 is situated at the edge of one of the outcrops, where the recharge rate is highest and thus potentially may interfere with well dynamics. In fact, we anticipate that the poorly designed location of this observation well may have a negative impact on the accuracy of identification for all wells in the eastern Las Posas basin, but especially for P4 and P6, the farthest away from O4.

[29] We designed four hypothetical test cases. Case A implements a constant pumping rate for all wells. Cases B and C address spatially and temporally varying pumping rates, respectively. Finally, Case D considers measurement errors within the case A scenario.

[30] Here are the details of the four test cases.

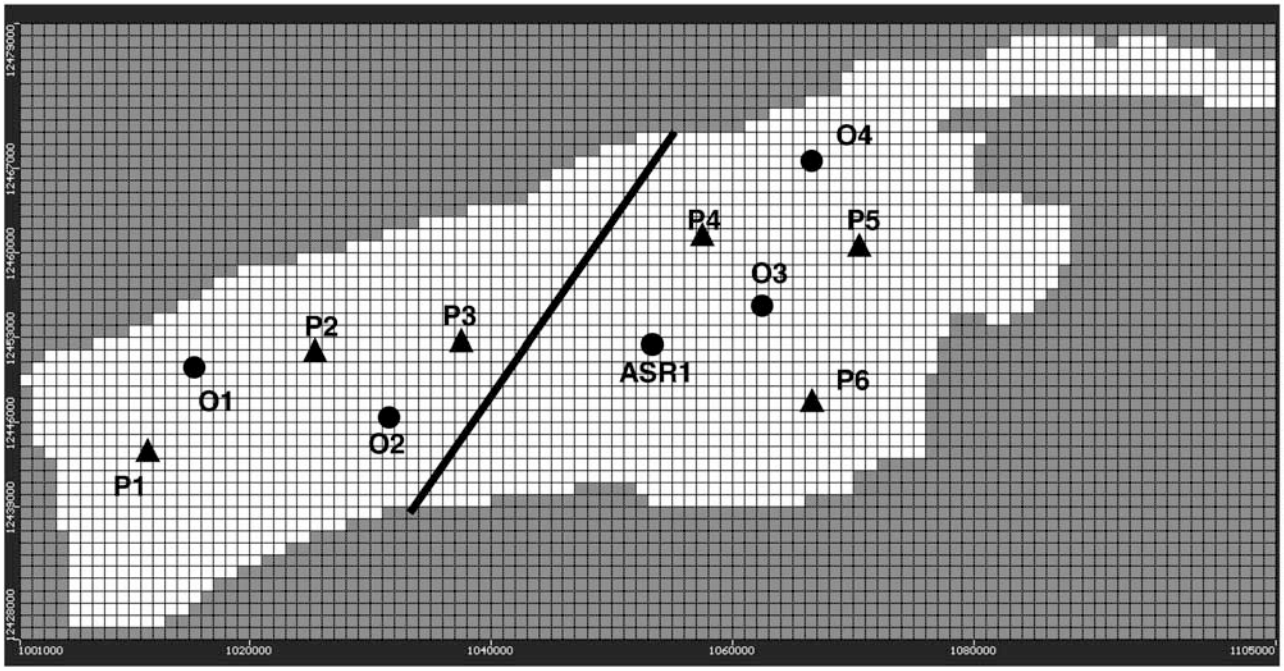

Figure 5. Planar view of the active cells (white color) of the Las Posas Basin model.

[31] 1. Case A has a constant pumping rate. All the wells (including the ASR) are pumping at a constant rate of 2832 m^3/d.

[32] 2. Case B has a spatially varying pumping rates. The ASR and the P1 wells are pumping at a rate of $q = 2832$ m^3/d. Pumping rates assigned to wells P2 to P6 range from 90% (P2) to 50% (P6) of q with four intervals of 10% q.

[33] 3. Case C has a temporally varying pumping rates. In this case the ASR well is pumping at a constant rate of 2832 m^3/d, while the pumping rates of the other six wells vary in time starting from 2832 m^3/d during the first stress period, decreasing uniformly at each stress period by 283 m^3/d (10% of the initial value) until period 5. During period 6 the pumping rates start to increase by the same amount and reach the initial value at the end of the simulation (stress period 9).

[34] 4. Case D has errors in the observed heads. This scenario is the same as in Case A, but we add measurement errors to the observations. Specifically, we corrupt the observations with Gaussian noise with zero mean and standard deviations of 0.03 m and 0.3 m, respectively. Note that a standard deviation of 0.3 m is unusually large when in practice the head measurement error seldom exceeds 0.03 m [*Prinos et al.*, 2004].

3.3. Numerical Results

[35] We evaluate the results of the numerical experiments by measuring the difference between the analysis and the true solution by means of the root mean square error (RMSE) and the average error in percentage, defined for the uth stress period as

$$RMSE(t_u) = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(\frac{f_{j,u}^{a} - f_{j,u}^{t}}{f_{j,u}^{t}}\right)^2} \tag{33}$$

$$\varepsilon^{(f)}(t_u) = \frac{1}{n}\sum_{j=1}^{n}\left|\frac{f_{j,u}^{a} - f_{j,u}^{t}}{f_{j,u}^{t}}\right| \times 100 \tag{34}$$

where f can be hydraulic head ($f = h$) or drawdown ($f = s = h - \overline{h}$, with $\overline{h}$ being the reference head), in which case the summation runs over the observation wells with $n = n_o$; or f can represent pumping rates ($f = q$), in which case the summation encompasses the pumping wells with $n = n_w$. To better highlight the error reduction achieved by the nudging algorithm, in the case of drawdown we plot the average error in percentage terms before and after the update for each stress period. When considering the pumping rate identification errors, we also look at the global pumping error ($\varepsilon_I^{(q)}(t_u)$), calculated using equation (34) but without the absolute value. The nudged simulations consider the ASR pumping rates as known and seek to estimate the unknown rates at the six pumping wells using head measurements from the four observation wells. The initial guess for the pumping rate at the six pumping wells always is assumed to be zero. The total simulation time is 54 months and is divided into nine stress periods.

[36] Figure 6 shows the RMSE as a function of time for both the analyzed drawdowns at the observation wells and the identified pumping rates of the pumping wells. In Figure 7, we show the behavior of the average drawdown error before and after the update step. Finally, Figure 8 (top) shows the temporal behavior of the average error in percentage for the identified pumping rate for each well, while Figure 8b indicates the total pumping error.

[37] The results show that the errors in the drawdown decrease drastically after the first update for Cases A, B and C, both in terms of the RMSE and the average relative error (Figures 6 (top) and 7 (bottom)). For Case D, where residuals are bounded by the imposed measurement errors, the successive updates continue to show slow but steady improvement, with relative errors approaching 1%. Corre-

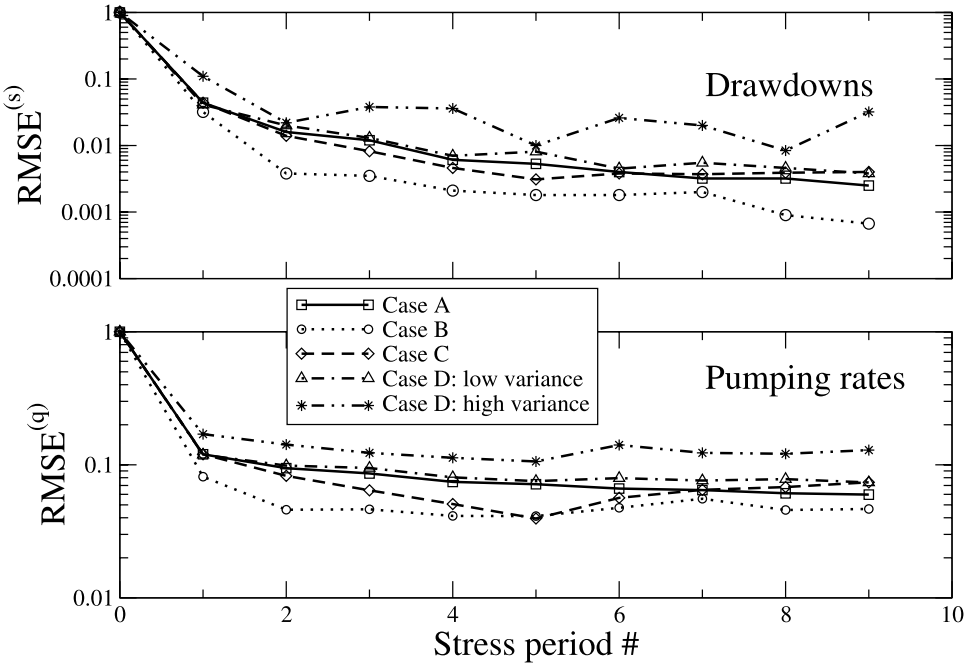

Figure 6. Temporal behavior of RMSE (top) for the drawdowns and (bottom) for the identified pumping rates for all test cases.

spondingly, the errors in the identified pumping rates show a similar behavior, but display larger values in both the RMSE and the relative errors. This is due to the fact that nudging minimizes observation errors, but the pumping rates are estimated indirectly via the influence coefficient. The overall results clearly show that the proposed approach works as well as expected in recovering the head distribution at the observation wells, but in some instances may produce a relatively larger error in pumping rate identification. This may be due to data insufficiency or using predefined and fixed spatial interpolation ratios of $w'_{i,j}$ and $w_{i,j}$* to estimate the dynamically changing, unknown pumping. The errors also can be attributed in part to the accuracy of the influence coefficient, particularly in certain areas of the domain where the *h-q* relationship is nonlinear (e.g., close to the recharge zones). This situation can be improved by performing iterations within each stress period [*Lorenc*, 1986; *Yeh*, 1986], or by increasing the time interval in the

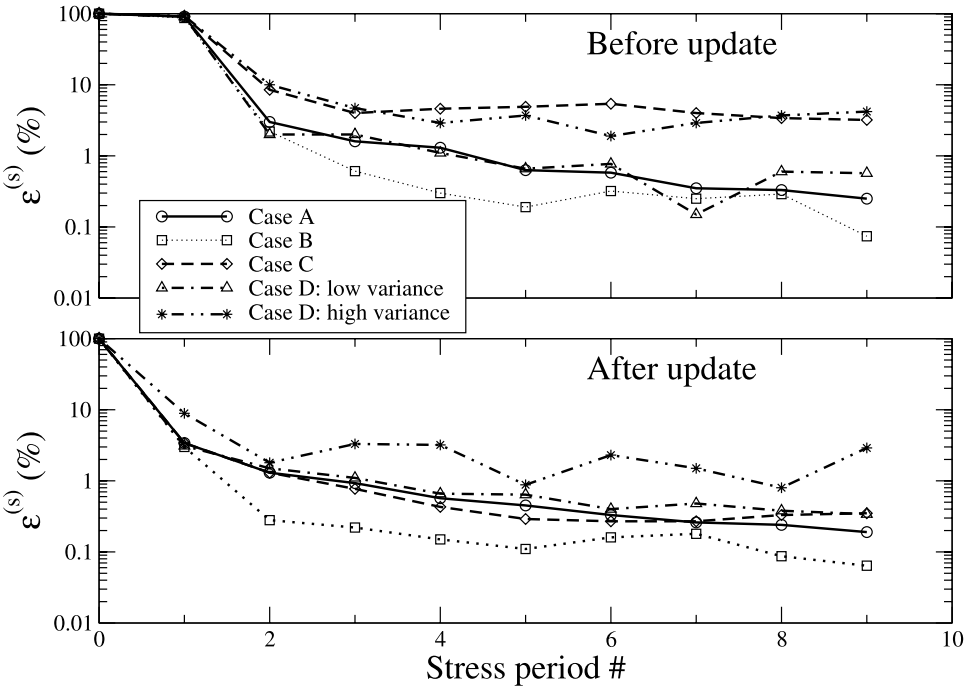

Figure 7. Temporal behavior of average percentage error for the drawdowns (top) before and (bottom) after the update for all test cases.

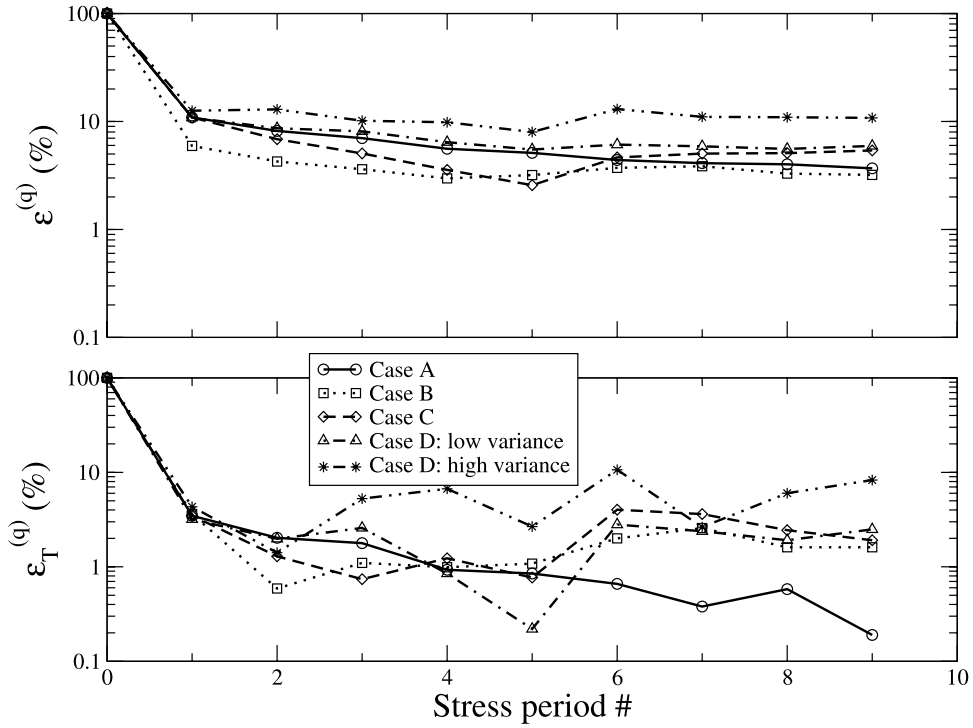

Figure 8. Temporal behavior of average percentage error (top) in the identified pumping rates and (bottom) for the global pumping for all test cases.

weighting functions from one to two or more stress periods. Note that nudging achieves greater accuracy in the identification of the total pumping rate (sum over all wells), as shown in Figure 8 (bottom), where the total error percentage remains consistently below five percent, except again for the high variance of Case D, where the measurement errors obviously affect the accuracy of the identification.

[38] Finally, Figure 9 shows the single well pumping rate identification error (equation (34)) for Cases A, B, and C. As expected, the performance of the proposed nudging algorithm is not optimal for pumping wells P4 and P6

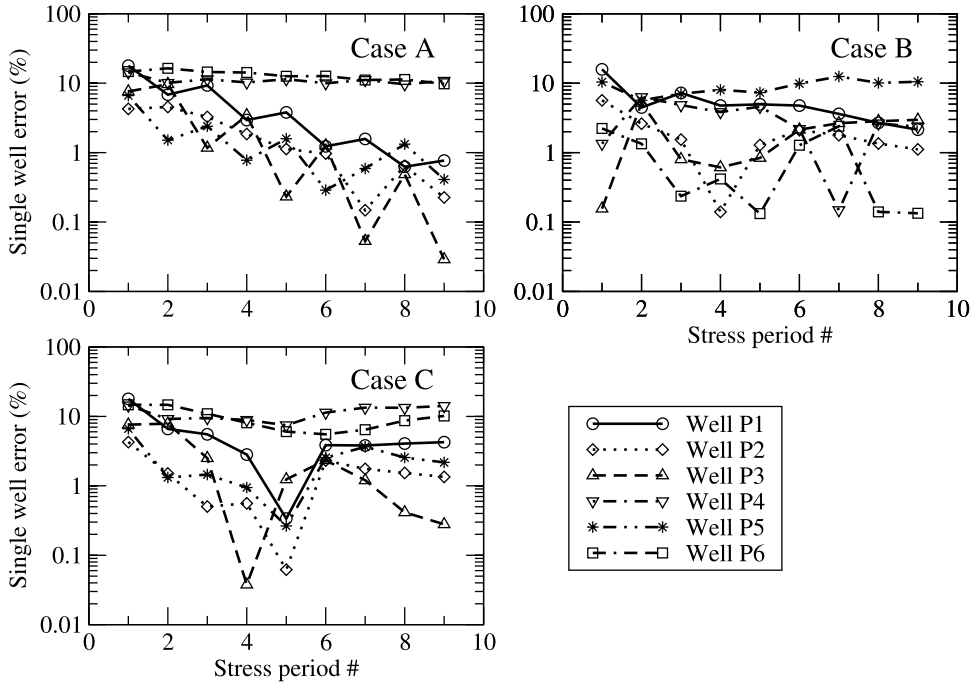

Figure 9. Temporal behavior of average percentage error in the identified pumping rates for each single well P1 to P5, for cases A, B, and C.

because of their relatively poor spatial location away from the observation wells, thus showing the importance of selecting observation well locations [*Yeh*, 1992]. Optimal experimental design algorithms, such as those discussed by *McPhee and Yeh* [2006], could be employed to define the optimal location of the observation wells.

4. Conclusions

[39] We have proposed a Newtonian relaxation (nudging) approach for estimating unknown pumping in a groundwater basin using observed hydraulic heads at a number of selected locations. The method of nudging relaxes the model state (hydraulic head) toward the observed state by adding an artificial sink term in the governing equation so that the model solution is nudged toward individual observations at each update time when new observations on the state variable are available. For purposes of pumping identification, the sink term corresponds to the unknown pumping. This formulation provides a consistent physical interpretation for identifying pumping rates without solving the inverse problem.

[40] We applied the proposed approach to the Las Posas Groundwater Basin in southern California in order to identify pumping from private wells. We utilized a calibrated MODFLOW model for flow simulation and numerical experiments. The influence coefficient method was used to calculate the incremental drawdown due to a unit change in pumping. With a linearity assumption, the influence coefficients are fixed constants for each stress period. We also assumed that there is no other recharge or discharge in the aquifer system except for the specified wells. This assumption simplifies algorithm development and avoids the distraction of other forcing terms. If needed, inclusion of these terms into the simulation model is easy and straightforward.

[41] We modified the observation nudging procedure to add forcing terms only at pumping well locations using a weighted average of residuals calculated from observation wells located within a predefined radius of influence. We tested the robustness of the proposed methodology under four different scenarios. The case studies show that the proposed approach is (1) efficient for estimating the unknown pumping from private wells, (2) able to distinguish different pumping rates from different pumping wells, and (3) able to provide accurate estimation of total extraction from private wells. Additionally, because of nudging and accurate estimation of pumping, the simulation model produces a more accurate solution for head distribution. In this study, observations are available at the end of each stress period. It is worthwhile to further explore the approach in order to consider the case in which whether observations at different locations are available at different times.

Appendix A

[42] The derivation of equation (7) from equation (4) follows. Algebraic manipulation of (4) leads to

$$\left(\frac{1}{\Delta t}P - L\right)h_{k+1} = \frac{1}{\Delta t}Ph_k + q_{k+1} \tag{A1}$$

where for simplicity we assume a constant Δt, but variable step sizes can be accommodated with no difficulties. Since the system matrix $M = (\frac{1}{\Delta t}P - L)$ is positive definite and thus invertible, we can stipulate

$$h_{k+1} = \frac{1}{\Delta t}M^{-1}Ph_k + M^{-1}q_{k+1} = Ah_k + Bq_{k+1}. \tag{A2}$$

Let m be the number of time steps between times t_u and t_{u+1}. To advance the time of the previous nudging update t_u to the time when the next observation is available, t_{u+1}, we need to perform m MODFLOW time steps, so that $t_{u+1} = t_u + m\Delta t$. The forecast head at t_{u+1}, which is the solution of (A2) for the m time steps, is formally represented as

$$h^f_{u+1} = A^m h_u + \sum_{b=1}^{m} A^{b-1}Bq_b = F_u(h_u) \tag{A3}$$

where the operator F_u characterized as

$$F_u(x) = A^m x + \sum_{b=1}^{m} A^{b-1}Bq_b. \tag{A4}$$

The analysis vector thus is formulated as:

$$h^a_{u+1} = F_u(h_u) + \overline{G}\left(h^o_{u+1} - h^f_{u+1}\right) = F_u(h_u) + \overline{G}\, d_{u+1} \tag{A5}$$

where matrix $\overline{G}$ is dimensionless. We use the subscript u in F_u to indicate that m is a function of the current update interval $[t_u, t_{u+1}]$.

[43] **Acknowledgments.** This study was supported by the Calleguas Municipal Water District (http://www.calleguas.com/index.html). We would like to thank three anonymous reviewers for their in-depth and constructive reviews.

References

Auroux, D., and J. Blum (2008), A nudging-based data assimilation method: The Back and Forth Nudging (BFN) algorithm, *Nonlinear Processes Geophys.*, *15*, 305–319.

Bear, J. (1988), *Dynamics of Fluids in Porous Media*, Dover, New York.

Becker, L., and W. W.-G. Yeh (1972), Identification of parameters in unsteady open channel flows, *Water Resour. Res.*, *8*, 956–965, doi:10.1029/WR008i004p00956.

California State Water Resources Board (1956), *Ventura County Investigation: Bulletin 12*, Sacramento, Calif.

Drusch, M. (2007), Initializing numerical weather prediction models with satellite-derived surface soil moisture: Data assimilation experiments with ECMWF's Integrated Forecast System and TMI soil moisture data set, *J. Geophys. Res.*, *112*, D03102, doi:10.1029/2006JD007478.

Farrar, C. D., L. F. Metzger, T. Nishikawa, K. M. Koczot, and E. G. Reichard (2006), Geohydrologic characterization, water-chemistry, and groundwater flow simulation model of the Sonoma valley area, Sonoma County, California, *U.S. Geol. Surv. Sci. Invest. Rep., 2006–5092*.

Gehrels, J. C., F. C. van Geer, and J. J. de Vries (1994), Decomposition of groundwater level fluctuations using transfer modeling in an area with shallow to deep unsaturated zones, *J. Hydrol. Amsterdam*, *157*, 105–138, doi:10.1016/0022-1694(94)90101-5.

Hanson, R. T., P. Martin, and K. M. Koczot (2003), Simulation of ground-water/surface-water flow in the Santa Clara–Calleguas Ground-Water Basin, Ventura County, California, *U.S. Geol. Surv. Open File Rep., 02-4136*.

Harbaugh, A. W., E. R. Banta, M. C. Hill, and M. G. McDonald (2000), MODFLOW-2000, The U.S. Geological Survey modular ground-water model–user guide to modularization concepts and the ground-water flow process, *U.S. Geol. Surv. Open File Rep., 00–92*.

Harter, T. (2003), *Legal Control of California's Water Resources*, Div. of Agric. and Nat. Resour., Univ. of Calif., Oakland.

Healy, R. W., and P. G. Cook (2002), Choosing appropriate techniques for quantifying groundwater recharge, *Hydrogeol. J.*, *10*, 91–109, doi:10.1007/s10040-001-0178-0.

Hoke, J., and R. A. Anthes (1976), The initialization of numerical models by a dynamic initialization technique, *Mon. Weather Rev.*, *104*, 1551–1556, doi:10.1175/1520-0493(1976)104<1551:TIONMB>2.0.CO;2.

Houser, P. R., W. J. Shuttleworth, J. S. Famiglietti, H. V. Gupta, K. H. Syed, and D. C. Goodrich (1998), Integration of soil moisture remote sensing and hydrologic modeling using data assimilation, *Water Resour. Res.*, *34*, 3405–3420, doi:10.1029/1998WR900001.

Hurkmans, R., C. Paniconi, and P. A. Troch (2006), Numerical assessment of a dynamical relaxation data assimilation scheme for a catchment hydrological model, *Hydrol. Processes*, *20*, 549–563, doi:10.1002/hyp.5921.

Ide, K., P. Courtier, M. Ghil, and A. C. Lorenc (1997), Unified notation for data assimilation: Operational, sequential and variational, *J. Meteorol. Soc. Jpn.*, *75*, 181–189.

Koczot, K. M. (1996), Estimating rates and patterns of historical groundwater pumpage for irrigation using a geobased numerical model, M.A. thesis, San Diego State Univ., San Diego, Calif.

Li, Z., and I. M. Navon (2001), Optimality of variational data assimilation and its relationship with the Kalman filter and smoother, *Q. J. R. Meteorol. Soc.*, *127B*, 661–684, doi:10.1002/qj.49712757220.

Lin, Y.-C., and H.-D. Yeh (2008), Identifying groundwater pumping source information using simulated annealing, *Hydrol. Processes*, *22*, 3010–3019, doi:10.1002/hyp.6875.

Lorenc, A. C. (1986), Analysis methods for numerical weather prediction, *Q. J. R. Meteorol. Soc.*, *112*, 1177–1194, doi:10.1002/qj.49711247414.

McDonald, M. G., and A. W. Harbaugh (1988), A modular three-dimensional finite-difference ground-water flow model, *U.S. Geol. Surv. Tech. Water Resour. Invest., Book 6, Chap. A1*, 586 pp.

McPhee, J., and W. W.-G. Yeh (2006), Experimental design for groundwater modeling and management, *Water Resour. Res.*, *42*, W02408, doi:10.1029/2005WR003997.

Miguez-Macho, G., G. L. Stenchikov, and A. Robock (2004), Spectral nudging to eliminate the effects of domain position and geometry in regional climate model simulations, *J. Geophys. Res.*, *109*, D13104, doi:10.1029/2003JD004495.

Nichols, N. K. (2003), Data assimilation: Aims and basic concepts, in *Data Assimilation for the Earth System, Nato ASI Sci. Ser., Ser. IV, Earth and Environmental Sciences*, vol. 26, edited by R. Swinbak et al., pp. 9–21, Kluwer Acad., Dordrecht, Netherlands.

Pacione, R., C. Sciarretta, C. Faccani, R. Ferretti, and F. Vespe (2001), GPS PW assimilation into MM5 with the nudging technique, *Phys. Chem. Earth Part A*, *26*, 481–485, doi:10.1016/S1464-1895(01)00088-6.

Paniconi, C., M. Marrocu, M. Putti, and M. Verbunt (2003), Newtonian nudging for a Richards equation-based distributed hydrological model, *Adv. Water Resour.*, *26*, 161–178, doi:10.1016/S0309-1708(02)00099-4.

Pauwels, V. R. N., R. Hoeben, N. E. C. Verhoest, and F. P. D. Troch (2001), The importance of the spatial patterns of remotely sensed soil moisture in the improvement of discharge predictions for small-scale basins through data assimilation, *J. Hydrol. Amsterdam*, *251*, 88–102, doi:10.1016/S0022-1694(01)00440-1.

Prinos, S., R. Irvin, and M. Byrne (2004), Water resources data Florida water year 2004, volume 2B: South Florida—Groundwater, *U.S. Geol. Surv. Water Data Rep., FL-04–2B*.

Quarteroni, A., and A. Valli (1997), *Numerical Approximation of Partial Differential Equations*, *Springer Ser. Comput. Math.*, vol. *23*, Springer, Berlin.

Ruud, N., T. Harter, and A. Naugle (2004), Estimation of groundwater pumping as closure to the water balance of a semi-arid, irrigated agricultural basin, *J. Hydrol. Amsterdam*, *297*, 51–73, doi:10.1016/j.jhydrol.2004.04.014.

Sanford, W. (2002), Recharge and groundwater models: An overview, *Hydrogeol. J.*, *10*, 110–120, doi:10.1007/s10040-001-0173-5.

Scanlon, B. R., R. W. Healy, and P. G. Cook (2002), Choosing appropriate techniques for quantifying groundwater recharge, *Hydrogeol. J.*, *10*, 18–39, doi:10.1007/s10040-001-0176-2.

Stauffer, D. R., and N. L. Seaman (1990), Use of four-dimensional data assimilation in a limited-area mesoscale model: Part I: Experiments with synoptic-scale data, *Mon. Weather Rev.*, *118*, 1250–1277, doi:10.1175/1520-0493(1990)118<1250:UOFDDA>2.0.CO;2.

Stauffer, D. R., N. L. Seaman, and F. S. Binkowski (1991), Use of four-dimensional data assimilation in a limited-area mesoscale model: Part II: Effects of data assimilation within the planetary boundary layer, *Mon. Weather Rev.*, *119*, 734–754, doi:10.1175/1520-0493(1991)119<0734:UOFDDA>2.0.CO;2.

Stiles, B. W., B. D. Pollard, and R. S. Dunbar (2002), Direction interval retrieval with thresholded nudging: A method for improving the accuracy of QuikSCAT winds, *IEEE Trans. Geosci. Remote Sens.*, *40*, 79–89, doi:10.1109/36.981351.

Troch, P., C. Paniconi, and D. McLouglin (2003), Catchment-scale hydrological modeling and data assimilation, *Adv. Water Resour.*, *26*, 131–135, doi:10.1016/S0309-1708(02)00087-8.

Tung, C.-P., and C.-A. Chou (2004), Pattern classification using tabu search to identify the spatial distribution of groundwater pumping, *Hydrogeol. J.*, *12*, 488–496, doi:10.1007/s10040-004-0344-2.

Verron, J. (1990), Altimeter data assimilation into an ocean circulation model: Sensitivity to orbital parameters, *J. Geophys. Res.*, *95*, 11,443–11,459, doi:10.1029/JC095iC07p11443.

Vidard, P. A., F.-X. Le Dimet, and A. Piacentini (2003), Determination of optimal nudging coefficients, *Tellus, Ser. A*, *55*, 1–15.

Willis, R., and W. W.-G. Yeh (1987), *Groundwater Systems Planning and Management: Planning and Management*, Prentice-Hall, Englewood Cliffs, N. J.

Yeh, W. W.-G. (1986), Review of parameter identification procedures in groundwater hydrology: The inverse problem, *Water Resour. Res.*, *22*, 95–108, doi:10.1029/WR022i002p00095.

Yeh, W. W.-G. (1992), Systems analysis in ground-water planning and management, *J. Water Resour. Plann. Manage.*, *118*, 224–237, doi:10.1061/(ASCE)0733-9496(1992)118:3(224).

W.-C. Cheng and W. W.-G. Yeh, Department of Civil and Environmental Engineering, University of California, Los Angeles, CA, USA. (williamy@seas.ucla.edu)

D. R. Kendall, Calleguas Municipal Water District, 2100 Olsen Road, Thousand Oaks, CA 91362, USA.

M. Putti, Department of Mathematical Methods and Models for Scientific Applications, University of Padua, Padua I-35131, Italy.