Calibration of transient groundwater models using time series analysis and moment matching

Mark Bakker,[1,2] Kees Maas,[1,2] and Jos R. Von Asmuth[1,2]

Received 8 June 2007; revised 4 January 2008; accepted 16 January 2008; published 15 April 2008.

[1] A comprehensive and efficient approach is presented for the calibration of transient groundwater models. The approach starts with the time series analysis of the measured heads in observation wells using all active stresses as input series, which may include rainfall, evaporation, surface water levels, and pumping. The time series analysis results in the impulse response function of each stress at the observation well. For each impulse response function, the temporal moments M_0 and M_1 may be computed. Both moments fulfill differential equations that are equivalent to the differential equation for steady groundwater flow, with known values along physical boundaries. The model of M_0 may be calibrated for the transmissivity, as it does not depend on the storage coefficient; the computed values of M_0 at the observation wells are used for calibration. The model of M_1 may be calibrated for the storage coefficient, once the transmissivity is known from the M_0 model; the computed values of M_1 at the observation wells are used for calibration. The approach is intended for systems that may be approximated as linear. In summary, our proposed calibration process for transient models reduces to the calibration of only two steady models. Several examples are given to demonstrate the accuracy and efficiency of the proposed approach.

Citation: Bakker, M., K. Maas, and J. R. Von Asmuth (2008), Calibration of transient groundwater models using time series analysis and moment matching, *Water Resour. Res.*, *44*, W04420, doi:10.1029/2007WR006239.

1. Introduction

[2] Hydraulic head variations, also called "groundwater dynamics," are measured in observation wells to study the general behavior of an aquifer and to monitor the consequences of implemented management plans. Many observation wells are monitored automatically, for example, bimonthly or more frequently depending on the specific circumstances. In addition, stresses on the system are measured and recorded, including rainfall, evaporation, pumping, and stream stages. In this paper, we present a new methodology on how to use this wealth of information in the calibration of transient groundwater models in a comprehensive fashion. Our approach is based on the integration of deterministic groundwater modeling with the results of time series analysis.

[3] In time series analysis, a series of observations is simulated with one or more explanatory series through a mathematical technique. In our case, the observations are measured heads in an observation well, and the explanatory series are measured stresses. An important outcome of time series analysis is the impulse response functions of the input series or stresses. Once the impulse response function $\theta(t)$ of a time-varying stress $N(t)$ is known at a point, the head $h(t)$ as a function of time may be computed through convolution [e.g., *Bear*, 1972; *Olsthoorn*, 2008]:

$$h(t) = \int_{-\infty}^{t} N(\tau)\theta(t-\tau)d\tau. \quad (1)$$

Impulse response functions represent the time response of the head due to an impulse of stress on the aquifer system; they are also called transfer functions. Conventional time series analyses, such as the Box-Jenkins method [*Box and Jenkins*, 1970], use a discrete approximation of the impulse response function consisting of a multistep function for early time followed by an exponentially decreasing function for later time. Discrete impulse response functions may also be obtained with nonparametric deconvolution [e.g., *Fienen et al.*, 2006; *Cirpka et al.*, 2007]. Another recent development in time series analysis is the predefined impulse response function in continuous time (PIRFICT) method, which uses specific parametric impulse response functions that are continuous functions of time [*Von Asmuth et al.*, 2002, 2008; *Von Asmuth and Bierkens*, 2005]. The continuous impulse response functions used in the PIRFICT method are less flexible than the discrete functions of the previous two methods, but appropriate choices for geohydrological problems give very accurate results [*Von Asmuth et al.*, 2008; *Von Asmuth and Knotters*, 2004], also when compared to Box-Jenkins results [*Von Asmuth et al.*, 2002]; specific advantages of the PIRFICT method are reviewed in section 2. Impulse response functions can be interpreted as scaled probability density functions and may be characterized by their temporal moments; the first two or three

[1] Water Resources Section, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, Netherlands.
[2] Kiwa Water Research, Nieuwegein, Netherlands.

0043-1397/08/2007WR006239$09.00

moments are often sufficient. The values of these temporal moments differ between observation wells and may be represented by a function of the horizontal coordinates. Mathematically, the temporal moments fulfill steady state, Poisson-type differential equations with known values along physical boundaries and may be modeled with standard groundwater models. We propose to calibrate these models against the moments of the response functions obtained from time series analysis. In this fashion, the measured groundwater dynamics and the measured stresses are taken into account in a comprehensive fashion during the calibration process.

[4] Calibration against moments, or moment matching, is a common technique in contaminant transport modeling of breakthrough curves [e.g., *Yu et al.*, 1999; *Luo et al.*, 2006]. It has sporadically been used for the calibration of heads in groundwater models. *Von Asmuth and Maas* [2001] outlined the approach used in this paper, presented the differential equations for the moments (also called the moment-generating functions), and suggested an application to ecohydrological modeling of vegetation. *Lebbink and Rolf* [2003] proposed calibrating steady models against structural groundwater heads, which they defined as the drainage level plus the average stress times the zeroth-order moment; they obtained the zeroth-order moments, or gains, from Box-Jenkins time series analyses and showed a significant improvement of their approach over the use of average heads for an average year. *Li et al.* [2005] presented an approach to determine spatially varying transmissivity and storativity from a pumping test through moment matching and geostatistics. They outlined how to obtain moments from drawdowns for arbitrary pumping regimes and applied their approach to perform geostatistical inversion of artificially generated pump test data, resulting in fields of transmissivity and storativity that agreed fairly well with the true distribution. *Li et al.* [2007] applied the approach to field data based on multiple pumping tests in Germany.

[5] In this paper, we will first briefly review the basics of time series analysis and some specifics of the method we employ. Next, we will discuss the differential equations that govern the moments of the impulse response functions, and we will describe our proposed calibration approach. Four applications to synthetic data are given to demonstrate different aspects of the performance of the approach.

2. Time Series Analysis

[6] In a time series analysis, an output series is modeled with a number of input series. For this paper, the output series is a measured series of heads at an observation well, and the input series are stresses on the aquifer, including rainfall, evaporation, pumping, and changes in surface water levels. The measured heads at an observation well $p(t)$ are written as the sum [e.g., *Von Asmuth et al.*, 2008]

$$p(t) = d + \phi(t) + \rho(t), \tag{2}$$

where d is a parameter that represents the head at the observation well if all stresses are zero, ϕ is the head variation caused by the transient stresses, and $\rho(t)$ is the residual series. The head variation ϕ consists of the sum of the contributions of all transient stresses that are used as input series. The effect of each stress on the head may be characterized with an impulse response function, the effect of a unit impulse of the stress. The effect of a time series of stress on the head may be computed through convolution of the stress with its corresponding impulse response function (1). This approach is suited for systems that are sufficiently linear. The residual series $\rho(t)$ may be interpreted as the result of a white noise process, which facilitates the computation of confidence intervals of the simulations [*Von Asmuth and Bierkens*, 2005].

[7] Heads are measured at intervals, for example, daily, weekly, or monthly. Many stresses are not measured continuously either, but cumulative values are measured for periods of a day, week, or month; average values may be obtained from the cumulative values. For many practical cases it is convenient to replace the convolution integral (1) by the following summation:

$$h(t_i) = \sum_{j=1}^{i} N_j \Theta_j(t_i), \tag{3}$$

where N_j is the average stress from t_{j-1} to t_j and $\Theta_j(t_i)$ is the block response at time t_i due to a unit stress from t_{j-1} to t_j.

[8] In this paper, we will use the PIRFICT method [*Von Asmuth et al.*, 2002] for time series analysis, but other methods, such as Box-Jenkins, may be applied in a similar fashion. In the PIRFICT method, specific parametric functions are selected to represent the impulse response functions of each stress [*Von Asmuth et al.*, 2008]. Use of continuous functions with few parameters (up to four in the current implementation) facilitates batch processing and the handling of irregular or high-frequency data and avoids overfitting [*Von Asmuth et al.*, 2002]. In this paper, we approximate the impulse response function with the following function with four parameters:

$$\begin{aligned} \theta(t) &= 0 \qquad t < 0 \\ \theta(t) &= A t^{\nu} \mathrm{e}^{-at - b/t} \quad t \geq 0, \end{aligned} \tag{4}$$

where $\theta(t)$ is the impulse response function, t is time, and A, ν, a, and b are parameters; the three-parameter functions used by *Von Asmuth et al.* [2002, 2008] are all special cases of this function. Integration of the impulse response function (4) to obtain the block response was carried out using the algorithms of *Chaudhry et al.* [1996], but numerical integration gives accurate results as well.

[9] The objective of a time series analysis is to find values for the parameters in the impulse response functions such that the measured series of heads are explained by the measured stresses in an optimum fashion. The parameter values of the impulse response functions are, of course, unique for each observation well. We will compute the parameters through minimization of an objective function F:

$$F = \sum_{i=1}^{N} (p_i - h_i)^2, \tag{5}$$

where N is the number of observations, p_i is the observation at time t_i, and $h_i = d + \phi(t_i)$ is the head value of the time series model at t_i; examples of this standard least squares approach are given in sections 5–8. Alternatively, optimal parameters may be found through minimization of the

Table 1. Splitting the Differential Equation[a]

BC for h	BC for d	BC for ϕ
Steady $h = h_f$	$d = h_f$	$\phi = 0$
Transient $h = h_t$	$d = 0$	$\phi = h_t$
Steady $\partial h/\partial n = q_f$	$\partial d/\partial n = q_f$	$\partial \phi/\partial n = 0$
Transient $\partial h/\partial n = q_t$	$\partial d/\partial n = 0$	$\partial \phi/\partial n = q_t$
Steady $\partial h/\partial n = \alpha h + \beta_f$	$\partial d/\partial n = \alpha d + \beta_f$	$\partial \phi/\partial n = \alpha \phi$
Transient $\partial h/\partial n = \alpha h + \beta_t$	$\partial d/\partial n = \alpha d$	$\partial \phi/\partial n = \alpha \phi + \beta_t$

[a]BC is boundary condition.

variance of the noise model [*Von Asmuth and Bierkens*, 2005].

[10] In sections 3 and 4 we will discuss how we can use characteristics of the impulse response function directly in the calibration of transient groundwater models, without ever having to perform a transient simulation.

3. Groundwater Modeling

[11] Dupuit-Forchheimer flow in an aquifer is governed by [e.g., *Bear*, 1972]

$$\nabla \cdot (T \nabla h) = S \frac{\partial h}{\partial t} - N, \tag{6}$$

where $\nabla[L^{-1}]$ is the gradient vector, S [dimensionless] is the storativity of the aquifer, T [L^2 T^{-1}] is the transmissivity, and N is the areal recharge. We consider four types of transient stresses: transient areal recharge N_t, boundaries with a specified transient head h_t, a specified transient normal flux q_t, or a transient mixed boundary condition of the form $\partial h/\partial n = \alpha h + \beta_t$, where n is the direction normal to the boundary, α is a constant, and β_t varies with time. In addition, there may be a fixed areal recharge N_f and boundaries with a fixed head h_f, a fixed normal flux q_f, or a fixed mixed boundary condition $\partial h/\partial n = \alpha h + \beta_f$.

[12] The aquifer is approximated as a linear system, such that the head in the aquifer may be written as a steady component plus a transient component, similar to time series analysis (2)

$$h(x, y, t) = d(x, y) + \phi(x, y, t). \tag{7}$$

The steady component d fulfills the steady portion of (6)

$$\nabla \cdot (T \nabla d) = -N_f, \tag{8}$$

and ϕ fulfills (6)

$$\nabla \cdot (T \nabla \phi) = S \frac{\partial \phi}{\partial t} - N_t. \tag{9}$$

The boundary conditions are divided correspondingly, as summarized in Table 1. As a result, the steady component d represents the head in the aquifer when all the transient stresses are set to zero. Note that summation of the second and third columns in Table 1 results in the first column, and thus summation of the models of d and ϕ result in the model of h.

[13] The transient part ϕ is divided further into separate components for each transient stress. First, consider a transient recharge N_t. The differential equation for this component of ϕ is given by (9). Boundary conditions are given in Table 1 with $h_t = q_t = \beta_t = 0$. The solution for this component of ϕ is obtained by determining the solution for a unit impulse of recharge (the impulse response function), after which $\phi(x, y, t)$ may be obtained for an arbitrary recharge through convolution of the impulse response function with the recharge (Duhamel's principle [*Bear*, 1972; *Zwillinger*, 1997]). The impulse response function θ for an impulse of recharge fulfills (9) where the recharge is replaced by the Dirac delta function $\delta(t)$:

$$\nabla \cdot (T \nabla \theta) = S \frac{\partial \theta}{\partial t} - \delta(t). \tag{10}$$

Boundary conditions for θ are identical to boundary conditions for this component of ϕ, as stated above.

[14] Treatment of other transient stresses is similar. The component of ϕ for any of the other transient stresses fulfills (9) with $N_t = 0$. For each component, one of the transient stresses h_t, q_t, or β_t is specified, while the others are zero. The impulse response function for each component also fulfills (9) with $N_t = 0$, while the nonzero transient stress is replaced by the Dirac delta function. Once the impulse response function for a stress has been determined, the response for an arbitrarily varying stress may be obtained through convolution.

[15] In summary, we have defined a differential equation with boundary conditions for the impulse response function of each transient stress. In addition, we can determine the impulse response function of each transient stress at observation wells by using time series analysis. Our objective is to calibrate the models of the impulse response functions against the impulse response functions obtained with time series analysis. One possible way to compute the impulse response function is to solve a transient groundwater model for an impulse of the stress, using a small enough time step [e.g., *Olsthoorn*, 2008]. Calibration of such a model against the impulse response function obtained from time series analysis at observation wells can be achieved by computing head values at the observation wells at a number of times and minimizing the sum of squares of the residuals (e.g., by using a parameter estimation package such as PEST (J. Doherty, Manual for PEST, 5th edition, available at www.sspa.com/pest)). Head values at early times may be weighted differently to emphasize a good fit of the early or late response. We will present an alternative approach that only requires the solution of steady models and allows for the incorporation of the measured time series at observation wells in a comprehensive fashion.

4. Moment Matching

[16] As stated in section 1, impulse response functions may be viewed as scaled probability density functions and may be characterized by their temporal moments. The kth moment (sometimes called the kth raw moment) of an impulse response function is defined as

$$M_k = \int_{-\infty}^{\infty} t^k \theta \mathrm{d}t. \tag{11}$$

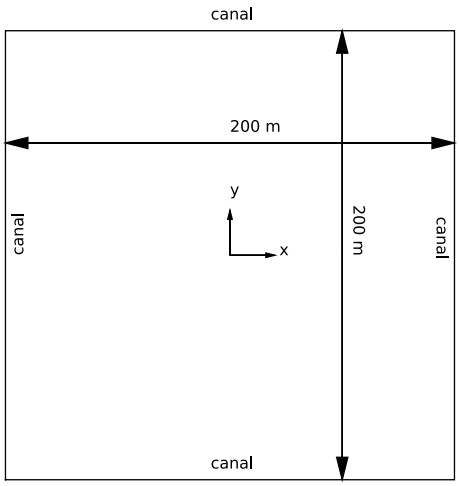

Figure 1. Square model area with canals on all four sides.

Temporal moments are related to common statistical terms. M_0 represents the area under the impulse response function, $\mu = M_1/M_0$ is the mean of the impulse response function, and $\sigma^2 = M_2/M_0 - \mu^2$ is the variance. From a hydrological perspective, M_0 is the latetime response to a unit step input. Hence, if M_0 is multiplied with the mean value of the stress, the mean effect on the head is obtained (recall that we have restricted ourselves to linear systems here).

[17] Temporal moments fulfill differential equations that are equivalent to differential equations for steady groundwater flow. The differential equation of the zeroth-order moment of the impulse response function for areal recharge is obtained through integration of both sides of (10) from time is minus to plus infinity, which gives [e.g., *Von Asmuth and Maas*, 2001; *Li et al.*, 2005; *Bakker et al.*, 2007]

$$\nabla \cdot (T \nabla M_0) = -1. \tag{12}$$

Equation (12) is identical to the differential equation for steady groundwater flow with a unit areal recharge. Similarly, the differential equation for the zeroth-order moment of the impulse response functions of the other transient stresses is

$$\nabla \cdot (T \nabla M_0) = 0 \tag{13}$$

which is identical to the differential equation for steady flow with zero areal recharge. Boundaries along which $\theta = \delta$, $\partial\theta/\partial n = \delta$, or $\partial\theta/\partial n = \alpha\theta + \delta$ may be integrated as well, which results in $M_0 = 1$, $\partial M_0/\partial n = 1$, or $\partial M_0/\partial n = \alpha M_0 + 1$, respectively.

[18] The differential equations for moments of order k ($k > 0$) are obtained through multiplication of both sides of the differential equation (10) with t^k and integration from minus to plus infinity, which gives

$$\nabla \cdot (T \nabla M_k) = -kSM_{k-1} \quad k = 1, 2, \ldots. \tag{14}$$

These differential equations (14) are also equivalent to the differential equation for steady groundwater flow but with a spatially varying recharge of the form kSM_{k-1}. Boundaries along which $\theta = \delta$, $\partial\theta/\partial n = \delta$, or $\partial\theta/\partial n = \alpha\theta + \delta$ may also be multiplied with t^k and integrated, which gives $M_k = 0$, $\partial M_k/\partial n = 0$, or $\partial M_k/\partial n = \alpha M_k$, respectively.

[19] The differential equations for the moments of the impulse response functions are exact. The only approximations that have been made are the common approximations of Dupuit-Forchheimer flow (although the analysis can be repeated for 3-D flow) and a linear aquifer system of the form (7). Differential equations (12), (13), and (14) may be solved with any standard groundwater model. *Li et al.* [2005] solved similar equations for M_0 and M_1 for a pumping well without areal recharge using a finite element code. *Bakker et al.* [2007] solved for M_0 and M_1 with areal recharge using analytic elements.

[20] We propose to calibrate transient groundwater models by constructing models of M_0 and M_1 and to calibrate against the values of M_0 and M_1 obtained at observation wells with time series analysis. Moment n of the impulse response function (4) used in the time series analysis is

$$M_n = 2A\left(\frac{b}{a}\right)^{\frac{n+\nu+1}{2}} \mathrm{K}_{n+\nu+1}\left(2\sqrt{ab}\right), \tag{15}$$

where $\mathrm{K}_{n+\nu+1}$ is the modified Bessel function of the second kind and order $n + \nu + 1$. The models of M_0 and M_1 are steady models. The model of M_0 is used to calibrate for the transmissivity as it does not depend on the storage coefficient (equations (12) and (13)). The model of M_1 is used to calibrate for the storage coefficient, using the calibrated transmissivity from the M_0 model and using M_0 itself in the recharge term (see equation (14)). In the remainder of this paper, we will evaluate different aspects of the performance of the proposed procedure by solving a number of hypothetical cases with combinations of different stresses. We consider flow in an unconfined aquifer, and we approximate the transmissivity as uniform, which is reasonable when fluctuations of the saturated thickness are relatively small. It is important to note, however, that the approach is applicable to fields with continuously varying transmissivities.

5. Varying Recharge

[21] We first consider a case where the head varies because of temporal variations in rainfall and reference evaporation. Consider a hypothetical case of a square agricultural plot with sides of $2L$, surrounded by drainage canals with a head fixed to 0.35 m (Figure 1). The aquifer is unconfined. For simplicity, the transmissivity is approximated as uniform, and the storage coefficient is homogeneous. The response of the hydraulic head may be characterized by the characteristic time τ:

$$\tau = \frac{L^2 S}{T}. \tag{16}$$

Initially, the head in the model area is equal everywhere to the constant water level in the drainage canals. The response due to a constant and uniform recharge, called the step response, may be computed with a numerical model. A block response, the response due to a unit recharge for a period of 1 d, is obtained through superposition of the step response starting at time $t = 0$ minus the step response

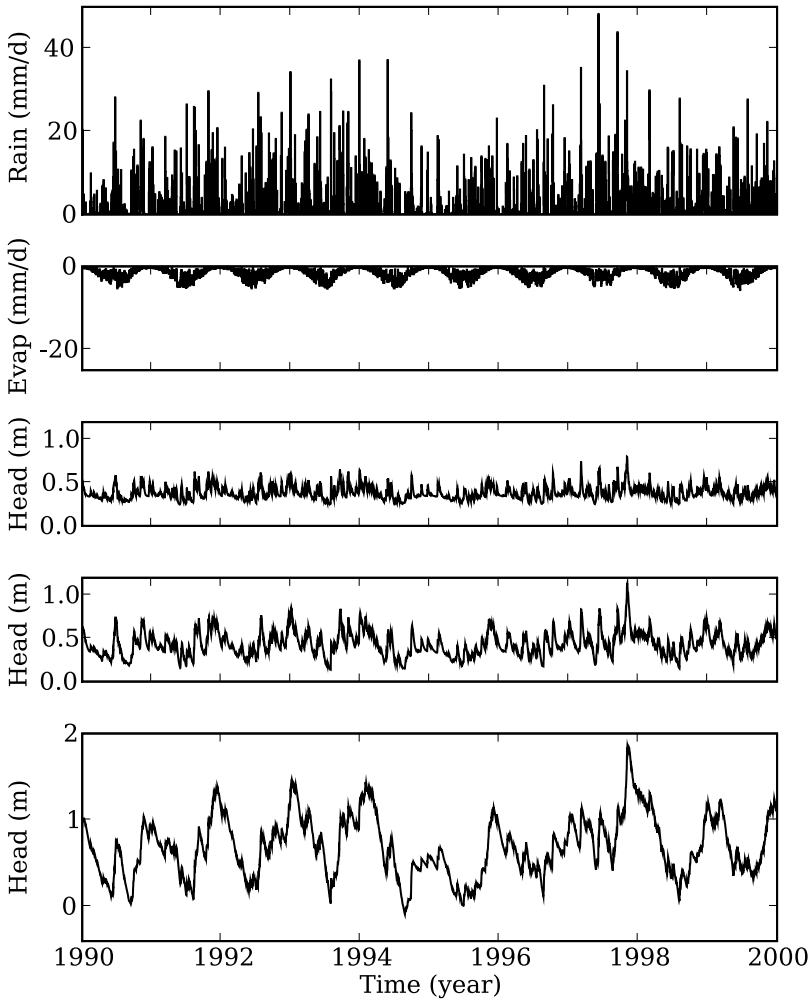

Figure 2. (top to bottom) Rainfall, reference evaporation, heads for short memory, heads for intermediate memory, and heads for long memory.

starting at time $t = 1$ d. We used a finite difference (FD) model with a uniform grid of 201 by 201 cells and a time step of 0.005τ (a further reduction of time step and cell size did not give a significant change in the outcome); use was made of MODFLOW2000 [*Harbaugh et al.*, 2000]. The cells along the edge of the model are specified as zero (see boundary conditions in Table 1). The time it takes for the head to rise to α% of its final position is called t_α and may be expressed in terms of the characteristic time τ. For example, at the center of the model area, $t_{50} = 0.16\tau$, $t_{90} = 0.49\tau$, $t_{95} = 0.64\tau$, and $t_{99} = 0.97\tau$.

[22] Measured time series of daily rainfall and reference evaporation are used from the weather station De Bilt in the Netherlands for the period 1988–2000. In the absence of a measured time series of daily heads for our hypothetical study area, a head series is generated at the center of the domain for the period 1990–2000 through convolution of the daily recharge data with the block response obtained with the FD model described above (Figure 2); only one head series is considered. Two additional years of recharge data are used to obtain accurate results starting on the first day of 1990. The recharge is computed as rainfall minus 80% of the reference evaporation; half the length of a side is set to $L = 100$ m. Three cases are considered with a memory varying from short ($\tau = 20$ d) to intermediate ($\tau = 50$ d) to long ($\tau = 200$ d). The aquifer properties for all three cases are given in Table 2. It is emphasized that we will pretend that we do not know these aquifer parameters but only know the head, rainfall, and reference evaporation series shown in Figure 2. Note that the head variation is smoother and the range (maximum to minimum) is larger when the memory of the system is longer. Systems with a longer memory have a larger storage coefficient, a smaller transmissivity, longer sides, or a combination thereof (see equation (16)).

[23] We will illustrate the calibration process outlined in section 4 by using the rainfall and reference evaporation series as input and the head series as output. First, we

Table 2. Aquifer Properties for Three Cases

Case	T, m^2/d	S	τ, d
1	100	0.2	20
2	40	0.2	50
3	10	0.2	200

Table 3. Recharge Case

	Time Series Analysis					Calibration			
τ, d	M_0, d	M_1, d^2	d, m	f	ε	T, m^2/d	δ_T, %	S	δ_S, %
20	29.43	130.57	0.350	0.80	0.001	100.1	0.1	0.2013	0.7
50	73.76	823.67	0.350	0.80	0.002	39.95	0.1	0.2023	1.2
200	294.8	13222	0.350	0.80	0.004	9.996	0.04	0.2033	1.7

perform a time series analysis with the PIRFICT method. The head h_i at time t_i in equation (5) is written as

$$h_i = \sum_{j=0}^{i} \left(N_j - f E_j\right)\Theta_j(t_i) + d, \tag{17}$$

where N_j and E_j are the average daily rainfall and reference evaporation from t_{j-1} to t_j, respectively, and d is the steady component when all transient stresses are zero. The block response Θ_j is the same for rainfall and evaporation [*Von Asmuth et al.*, 2008], and the factor f is used to scale the potential evaporation to the actual evaporation. There are six parameters to be determined: the four parameters A, ν, a, and b of the block response, f, and d. The parameters are obtained using a least squares fit (as explained in section 2). The resulting zeroth and first moments, calculated using (15), and the values for f and d are presented in Table 3. The normalized root mean square error ε was computed as the root mean square error divided by the range of the head variation (maximum level minus minimum level) and varied from 0.004 for the case with the longest memory to less than 0.001 for the case with the shortest memory (Table 3). Another measure for the goodness of fit is the percentage of variance accounted for R^2_{adj}, defined as

$$R^2_{\text{adj}} = \frac{\sigma_p^2 - \sigma_\rho^2}{\sigma_p^2} 100\%, \tag{18}$$

where σ_p^2 and σ_ρ^2 are the variance of the measured head and of the residual ρ, respectively. The percentage of variance accounted for was 100.0% for all three cases. As an illustration, the block response used to generate the input series (obtained with the FD model) and the block response obtained from time series analysis (4) are shown for the case with intermediate memory as the dots and line, respectively, in Figure 3.

[24] Next, we construct a model for moments M_0 and M_1 of the impulse response function for each of the three cases. The model for M_0 fulfills (12) and thus represents a steady model with a unit recharge over the entire model area and a fixed value of zero along the boundary; the FD method is applied to solve the model. We use the value of M_0 at the center of the plot as obtained from time series analysis (Table 3) to calibrate for the transmissivity. As there is only one unknown parameter, calibration is straightforward. The results for all three cases are shown in Table 3. The difference δ_T between the calibrated value and the true value is less than 0.2% for all cases.

[25] Next, we construct a model for M_1, which is a steady model of which the recharge is equal to SM_0 (see (14)), and again there is a fixed value of zero along the boundary. We calibrate this model for the storage coefficient S by using the value of M_1 at the center of the plot (obtained from time series analysis). This is again a straightforward calibration. The difference δ_S between the true and calibrated values is less than 1.8% for all cases.

[26] In summary, we solved only two steady models to calibrate each transient case, and the calibrated values for the transmissivity and storage coefficient were very close to the true values for all cases. The next question is how the calibration approach performs when the head and recharge series contain errors.

6. Error in Input and Output Series

[27] In section 5, we demonstrated our calibration approach by using a "measured" head series that was generated with a high-resolution finite difference model. This series was the exact response to the recharge series (at least within the numerical accuracy of the model). Time series analysis gave an almost perfect fit between the head series and the recharge series. In reality, a perfect fit will never be obtained. There are many reasons why a time series model (or a regular groundwater model for that matter) cannot give a perfect match. For example, measurements of heads, meteorological data, and hydrological data contain errors. Furthermore, weather stations are rarely located at the same location as the observation well, and thus the weather data differ from the actual rain and evaporation at the observation well. In addition, there may be unknown or unmeasured stresses influencing the head at the well. We will investigate the performance of our proposed calibration method with two types of errors: a random error in the head and a random error in the recharge.

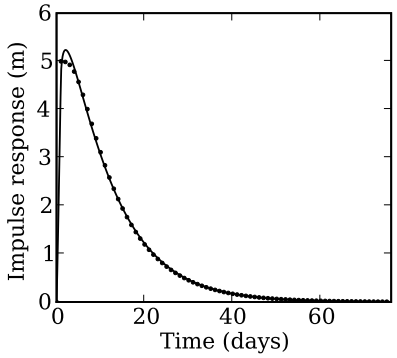

Figure 3. Block response of time series analysis (line) and finite difference (FD) model used to generate head series (dots) for case with intermediate memory.

Table 4. Case of Recharge With Errors in Series

	Time Series Analysis						Calibration			
	M_0, d	σ_0, d	M_1, d^2	σ_1, d^2	ε	R^2_{adj}, %	T, m^2/d	δ_T, %	S	δ_S, %
Head	73.25	0.675	796.3	16.6	0.027	95.9	40.22	0.55	0.1982	0.9
Recharge	74.35	0.534	865.5	14.3	0.022	97.6	39.63	0.93	0.2091	4.6
Both	73.72	0.875	832.8	22.4	0.035	93.5	39.97	0.075	0.2047	2.4

[28] We will continue our investigation for only one system with an intermediate memory (τ = 50 d); the aquifer parameters are T = 40 m^2/d and S = 0.2. A random error in the head is created by adding a random error to our head series. The random error has a uniform distribution from −5% to +5% of the head range. Time series analysis on this modified head series gives a good match: the root mean square error is 2.7% of the total variation in head, and the percentage of variance accounted for is 95.9% (row labeled "Head" in Table 4). The fit is illustrated for one of the 10 years of data in Figure 4 (middle); Figure 4 (top) shows the fit when no error is introduced. Our calibration approach still gives accurate results for the transmissivity and storage coefficient (Table 4).

[29] Next, we modify the recharge series by adding a random error with a uniform distribution from −5% to +5% of the recharge range. We pretend that this modified recharge series is the true recharge series at the observation well, while the rainfall and reference evaporation series are the true input series measured at the weather station. The modified recharge series is used to generate a head series at the observation well. Time series is performed on this head series, using the unmodified series of the weather station as the input series; the goodness of fit is illustrated in Figure 4 (bottom). The root mean square error and the explained variance are better than the previous case, but the zeroth and first moments have a larger error. Results of the calibration process are still good, with deviations in the transmissivity of 0.93% and storage coefficients of 4.6% (Table 4).

[30] The third case is the same as the second case, except that we added an additional random error to the head series. As expected, time series analysis on this head series gives the largest root mean square error (still only 3.5% of the total head range), but the fit remains good, and the calibration process gives accurate results for the transmissivity (deviation of 0.075%) and storage coefficient (deviation of 2.4%) (Table 4).

[31] Time series analysis also provides a confidence interval for the computed moments. The computed standard deviations of M_0 and M_1 are presented in Table 4 as σ_0 and σ_1, respectively. The standard deviation of M_0 is on the order of 1%, while the standard deviation of M_1 is on the order of 2.5%. These standard deviations may be used to compute standard deviations of the aquifer parameters T and S. Alternatively, when a model is calibrated using multiple

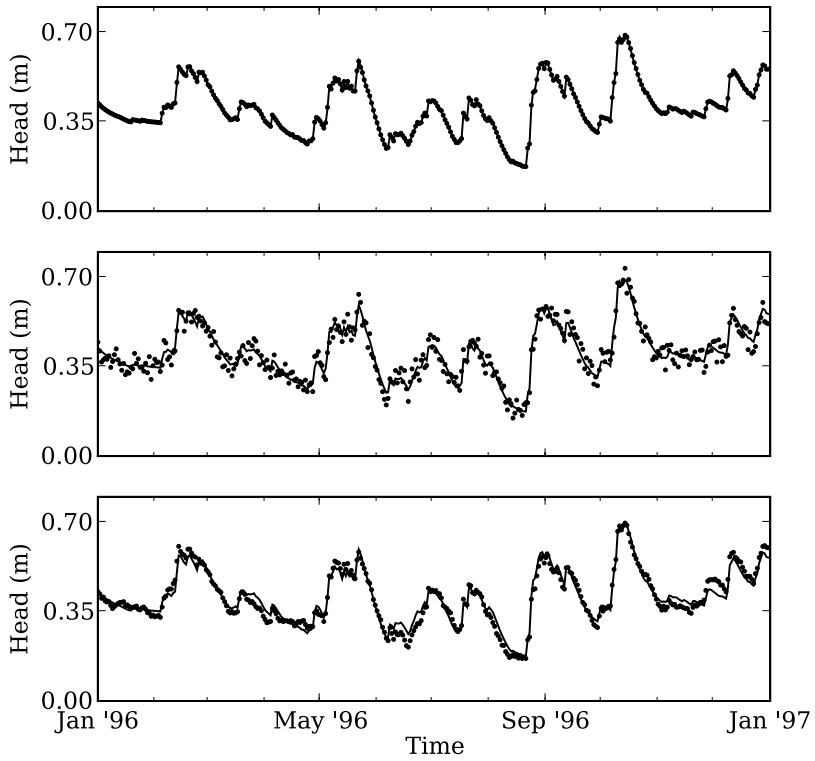

Figure 4. Fit of time series model (line) to head series (dots) for 1996 for (top) no error, (middle) error in head, and (bottom) error in recharge.

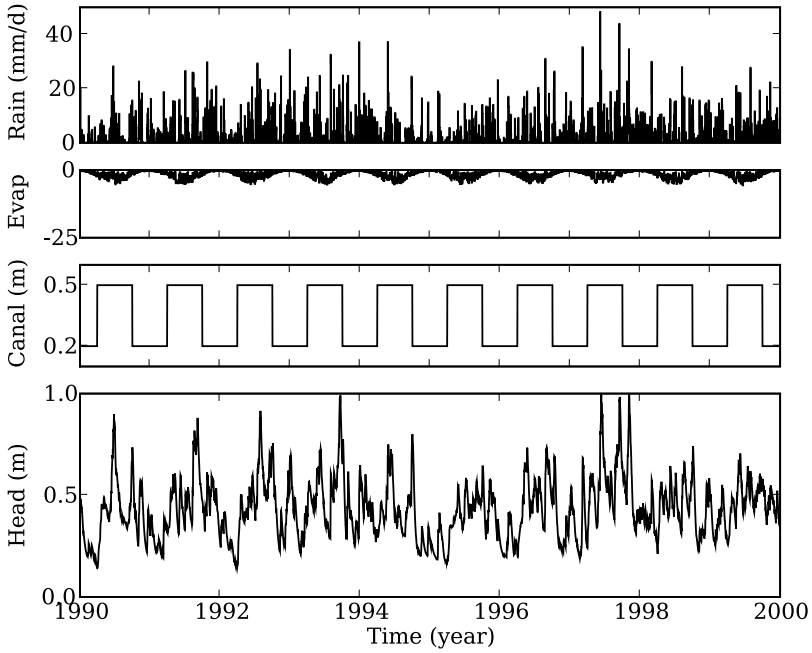

Figure 5. (top to bottom) Rain, evaporation, canal stage, and head.

observation wells, the inverse of the standard deviation may be used to weigh the moments in the objective function.

[32] It is noted that the result of the calibration approach is not very sensitive to the location of the observation well. In the current example the observation well is located at the center of the agricultural plot $(x, y) = (0,0)$. The procedure may be repeated for an observation well at $(x, y) = (50, 50)$, which results in deviations of 2.4% for T and 1.7% for S for the case with errors introduced in both the recharge and head. We even moved the observation well to 1 m from the canal along the edge, $(x, y) = (0, 99)$. Even though this is an extremely poor location for measuring groundwater dynamics, the aquifer parameters T and S were still obtained with deviations of only 0.35% for T and 9.7% for S for the case with errors introduced in both the recharge and head.

7. Varying Recharge and Canal Stage

[33] We will expand the example of sections 5 and 6 to three stresses. In addition to the rainfall and evaporation, the canals along the four sides of the model area have a seasonal value of 0.5 m from April through September and 0.2 m from October through March. We will use the aquifer parameters for a system with an intermediate memory used in section 6; random errors are not introduced. A head series at the center of the model area is again generated with an FD model; fluctuation of the recharge, canal stage, and head at the center of the model area are shown in Figure 5. As before, we will pretend that we do not know the aquifer parameters but only the time series of rainfall, reference evaporation, canal levels, and heads at the observation well.

[34] Following our calibration approach, we first perform a time series analysis, where h_i in equation (5) is now

$$h_i = \sum_{j=0}^{i} \left[(N_j - fE_j)\Theta_j(t_i) + P_j\hat{\Theta}_j(t_i) \right] + d, \qquad (19)$$

where P_j is the average canal stage between time t_{j-1} and t_j and $\hat{\Theta}_j$ is the block response function for the varying canal stage. The time series analysis gives a good fit with a root mean square error of only 0.2 cm. The two impulse response functions each have their own moments. For the recharge series, the moments obtained from time series analysis are labeled with an index r, $M_{r0} = 73.91$ d and $M_{r1} = 824.99$ d^2; the evaporation factor is $f = 0.80$. For the canal stage, the moments obtained from time series analysis are labeled with an index c, $M_{c0} = 1.0$ and $M_{c1} = 14.44$ d (dimensions of moments vary between stresses). Note that for this case M_{c0} is one. After all, the zeroth moment is the late time result of a unit step response, and a unit increase in the canal level will eventually result in a unit increase in the head.

[35] To calibrate for the transmissivity, we may construct a model for M_{r0} and calibrate against the value obtained from time series analysis, like we did in sections 5 and 6. This results in a transmissivity of $T = 39.87$ m^2/d, 0.32% different from the true value. It is also possible to construct a model of M_{c0}, but we cannot use it to calibrate for the transmissivity. After all, the model of M_{c0} fulfills (13), with fixed values of one along the boundaries. The solution of this model is a uniform value of one everywhere, independent of the choice of the transmissivity.

[36] Next, we want to calibrate for the storage coefficient. We can either calibrate a model of M_{r1} or we can calibrate a model of M_{c1}. The model of M_{r1} fulfills (14) where M_{r0} is used in the right-hand side. Calibration of this model gives $S = 0.2018$, 0.9% higher than the true value. The model of M_{c1} fulfills (14) as well, where $M_{r0} = 1$ everywhere. Calibration of this model gives $S = 0.1954$, 2.3% lower than the true value.

[37] Alternatively, it is possible to construct a model of a linear combination M_{s0} of M_{r0} and M_{c0}:

$$M_{s0} = \alpha M_{r0} + \beta M_{c0}, \qquad (20)$$

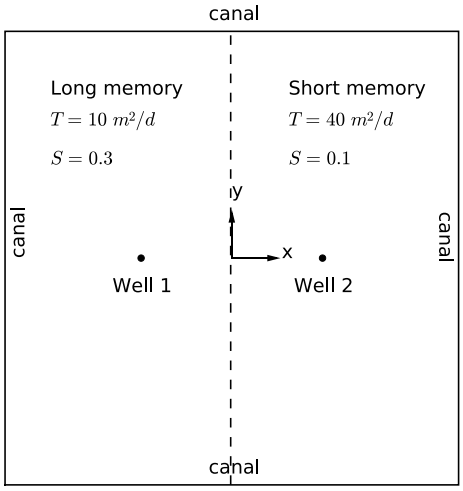

Figure 6. Setup for case with two sets of aquifer properties.

where α and β are weighing factors that may be chosen to emphasize calibration against the moments of the recharge impulse response or the canal impulse response. The differential equation of M_{s0} is obtained as the same linear combination of the differential equations for M_{r0} (12) and M_{c0} (13):

$$\nabla \cdot (T \nabla M_{s0}) = -\alpha. \tag{21}$$

The boundary conditions are also obtained as the same linear combination of the boundary conditions of M_{r0} and M_{c0}. For this case that means that the value in the canal is specified as β. The model of M_{s0} may be calibrated for the transmissivity against the value $M_{s0} = \alpha M_{r0} + \beta M_{c0}$ at the center of the model area.

[38] Calibration for S works in a similar fashion. The linear combination M_{s1} is defined as

$$M_{s1} = \alpha M_{r1} + \beta M_{c1}, \tag{22}$$

were M_{s1} fulfills

$$\nabla \cdot (T \nabla M_{s1}) = -S M_{s0}. \tag{23}$$

The value of M_{s1} is zero along the canal. The model of M_{s1} may be calibrated against the value $M_{s1} = \alpha M_{r1} + \beta M_{c1}$ at the center of the model area.

8. Variable Aquifer Parameters

[39] The final case concerns the same model area, but now the transmissivity and storage coefficient in the left half are different from the values used in the right half (Figure 6); this case will evaluate whether our calibration approach can handle systems with variable aquifer properties. To create a left half with a long memory, the transmissivity is set to T = 10 m^2/d, and the storage coefficient is raised to S = 0.3; the right half has a short memory with a higher transmissivity T = 40 m^2/d and a lower storage coefficient S = 0.1. (Since aquifer parameters are specified per cell in the FD model and heads are computed at cell centers, the left half is 99.5 m wide, and the right half is 100.5 m wide.) Rainfall and evaporation are the only transient stresses on the system. The FD model is used to compute the impulse response for recharge at two points: (x, y) = (−40, 0) and (x, y) = (40, 0). A head series for a period of 10 years is computed using the rainfall and reference evaporation series of Figure 2 and is shown for 1996 for both wells in Figure 7 (dots). We will again pretend we do not know the aquifer parameters and perform time series analysis on these two series with the rainfall and reference evaporation series as input series.

[40] The time series model is also shown in Figure 7 (lines). A comparison between the block response function of the FD model used to generate the head series (dots) and the fitted function of the time series model (lines) is shown in Figure 8. Note that the shapes of the block response functions are different from the shape shown in Figure 2. Reaction in well 1 (long memory system) is much slower

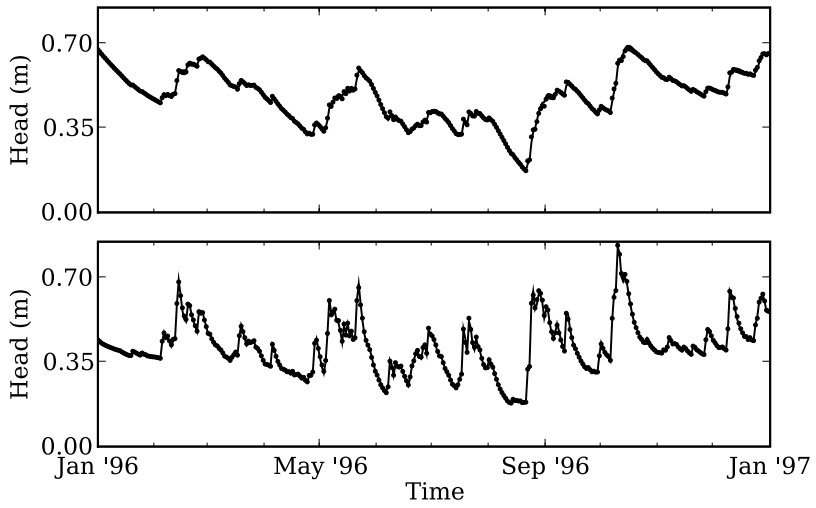

Figure 7. Head series for 1996 for case with two sets of aquifer properties: (top) well 1 (long memory) and (bottom) well 2 (short memory). Head series are indicated by dots, and results of time series analysis are indicated by lines.

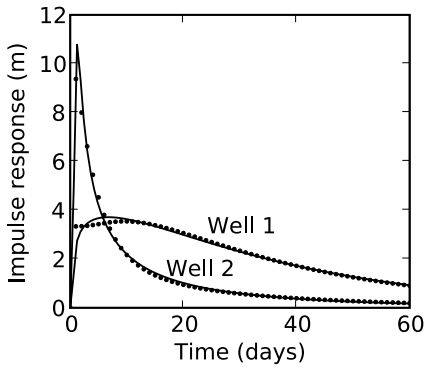

Figure 8. Block response functions for case with two sets of aquifer properties. Results of FD model are used to generate head series (dots), and results of time series analysis (lines) are shown. Well 1 is in long memory side, and well 2 is in short memory side.

than in well 2 (short memory system). In addition, we see a distinct effect of the two systems on each other. Well 1 in the long memory system shows a small second rise in the head after the initial rise caused by the block input. This second rise is due to the quick reaction of the short memory system on the right. The alteration in the shape of the block response in well 2 is more subtle. Here we see a quick decrease in head, as is expected for a short memory system, but then we see a much longer tail, which is caused by the slow reaction of the long memory system on the left. The impulse response function in well 1 shows some inaccuracy at early time, but this does not have a significant effect on the fluctuations in the well because of the long memory (Figure 7).

[41] Results of the time series analysis are still good and are summarized in Table 5. Our new calibration approach is again applied to find the values of T and S. There are now two T values and two known values of M_0. Similarly, there are two values of S and two known values of M_1. Calibration is still straightforward and leads to accurate results (Table 5). The deviation in the calibrated transmissivity is less than 1%, while the deviation in the calibrated storage coefficient is 7.4% in the short memory side and 1.3% in the long memory side.

9. Discussion and Conclusions

[42] We presented a new, efficient, and comprehensive calibration approach for transient groundwater models. The approach consists of two steps. In the first step, time series analysis is performed on the measured head series in observation wells, using measured stresses such as recharge, surface water levels, and pumping as input series. The time series analysis results in approximations of the first two temporal moments of the impulse response function of each stress at the observation well. In the second step, spatial models are made of the these two moments. The zeroth moment M_0 fulfills the steady differential equation (12) or (13); the model of M_0 is calibrated against the values of M_0 at the observation wells to determine the transmissivity of the aquifer. The first moment M_1 fulfills the steady differential equation (14) and is calibrated against the values of M_1 at the observation wells to determine the storativity of the aquifer. The proposed approach performed well for all presented examples and only requires the calibration of a couple of steady state models.

[43] In our examples, the recharge was computed as rainfall minus evaporation and may be adjusted for runoff, for example, using the curve number method [e.g., *McCuen*, 2004]. The aquifer parameters may vary spatially. We presented an example of two zones with different but homogeneous aquifer parameters, but our approach may be applied to calibrate continuously varying aquifer parameters using, for example, pilot points [*Doherty*, 2003].

[44] We highly recommend performing time series analysis as a first step in the calibration process of transient groundwater models, regardless of which calibration strategy is used. Time series analysis gives an indication of the best possible fit between the measured stresses and the measured heads using a linear relationship [*Von Asmuth et al.*, 2002; *Von Asmuth and Knotters*, 2004]. If the measured stresses cannot explain the measured heads, it will be most unlikely that a linear groundwater model with time-invariant parameters and with the measured stresses as input can simulate heads that mimic the observed heads. Time series models are used to estimate values of the moments of the impulse response functions but also the standard deviations of the estimates and thus the confidence intervals. The inverse of the standard deviation may be used to weigh moments according to their estimation accuracy during the calibration process.

[45] Records of both heads and stresses are needed to perform time series analysis. When one or more stresses are not recorded, it is impossible to construct an accurate transient groundwater model, and similarly, it will be difficult, but not always impossible, to construct an accurate time series model. It may be possible to obtain an accurate time series model when the missing stress is uncorrelated to the measured stresses. In that case there will be significant differences between the time series model and the measured heads, but the impulse response functions are expected to be reasonable. When the missing stress is correlated to one of the measured stresses, e.g., low values in the summer and high values in the winter, the impulse response function of the given stress will be inaccurate as it tries to include the effect of the missing stress. In areas of groundwater mining due to pumping, it may not be possible to obtain accurate impulse response functions for the pumping wells since the

Table 5. Case of Variable Aquifer Properties

	Time Series Analysis				Calibration			
Well Location (x, y)	M_0, d	M_1, d^2	ε	R^2_{adj}, %	T, m^2/d	δ_T, %	S	δ_S, %
(−40,0)	166.6	5741	0.0043	99.9	9.95	0.5	0.3038	1.3
(40,0)	84.32	1447	0.0048	99.9	40.33	0.8	0.0926	7.4

period of record is too short (the groundwater heads are still declining). It is expected, however, that time series analysis will be able to filter out the effect of groundwater mining effectively by including a step trend or a linear trend so that reasonable impulse response functions may be estimated for the other stresses on the system.

[46] The proposed approach is suited for systems that are sufficiently linear. The effect of nonlinear features, such as drains that carry water only part of the year or perennial streams, form part of current research. Furthermore, the areal recharge (rainfall minus evaporation) is assumed to reach the groundwater table instantaneously (this is a common assumption for groundwater models that do not include an explicit unsaturated zone). The main effect of the unsaturated zone is retardation and dispersion of recharge [*Besbes and de Marsily*, 1984; *Parlange et al.*, 1992]. This may be taken into account by adding a transfer model to represent the transformation of the recharge series through the unsaturated zone in the time series analysis step [e.g., *Kruithof*, 2001]; further research on this issue is ongoing.

[47] As was demonstrated in section 7, stresses with different impulse response functions present different options for calibration. For example, models may be calibrated on any linear combination of the moments of the stresses. Alternatively, different aquifer parameters may be determined for different stresses; that is, aquifers may have different transmissivities and storativities depending on the type of stress. This is a new paradigm that deserves further attention.

[48] **Acknowledgment.** This research was funded in part by the Joint Water Research Program of the Dutch Water Supply Companies.

References

Bakker, M., K. Maas, F. Schaars, and J. R. von Asmuth (2007), Analytic modeling of groundwater dynamics with an approximate impulse response function for areal recharge, *Adv. Water Resour.*, *30*(3), 493–504.

Bear, J., (1972), *Dynamics of Fluids in Porous Media*, Dover, New York.

Besbes, M., and G. de Marsily (1984), From infiltration to recharge: Use of a parametric transfer function, *J. Hydrol.*, *74*, 271–293.

Box, G. E. P., and G. M. Jenkins (1970), *Time Series Analysis, Forecasting and Control*, Holden-Day, San Francisco, Calif.

Chaudhry, M. A., N. M. Temme, and E. J. M. Veling (1996), Asymptotics and closed form of a generalized incomplete gamma function, *J. Comput. Appl. Math.*, *67*, 371–379.

Cirpka, O. A., M. N. Fienen, M. Hofer, E. Hoehen, A. Tessarini, R. Kipfer, and P. K. Kitanidis (2007), Analyzing bank filtration by deconvolution time series of electric conductivity, *Ground Water*, *45*(3), 318–328.

Doherty, J. (2003), Ground water model calibration using pilot points and regularization, *Ground Water*, *41*(2), 170–177.

Fienen, M. N., J. Luo, and P. K. Kitanidis (2006), A Bayesian geostatistical transfer function approach to tracer test analysis, *Water Resour. Res.*, *42*, W07426, doi:10.1029/2005WR004576.

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*, 121 pp.

Kruithof, A. J. (2001), The impulse-response function of the groundwater table; separation of the influence of the unsaturated and saturated zones (in Dutch), M.Sc. thesis, Delft Univ. of Technol., Delft, Netherlands.

Lebbink, J. J., and H. L. M. Rolf (2003), Generation of time-independent "structural groundwater heads" to be used as observations in the calibration of steady-state groundwater models, paper presented at MODFLOW and More 2003: Understanding Through Modeling, Int. Ground Water Cent., Golden, Colo.

Li, W., W. Nowak, and O. A. Cirpka (2005), Geostatistical inverse modeling of transient pumping tests using temporal moments of drawdown, *Water Resour. Res.*, *41*, W08403, doi:10.1029/2004WR003874.

Li, W., A. Englert, O. A. Cirpka, J. Vanderborght, and H. Vereecken (2007), Two-dimensional characterization of hydraulic heterogeneity by multiple pumping tests, *Water Resour. Res.*, *43*, W04433, doi:10.1029/2006WR005333.

Luo, J., O. A. Cirpka, and P. K. Kitanidis (2006), Temporal-moment matching for truncated breakthrough curves for step or step-pulse injection, *Adv. Water Resour.*, *29*(9), 1306–1313.

McCuen, R. H. (2004), *Hydrologic Analysis and Design*, 3rd ed., Prentice-Hall, Upper Saddle River, N.J.

Olsthoorn, T. N. (2008), Do a bit more with convolution, *Ground Water*, *46*(1), 13–22.

Parlange, M. B., G. G. Katul, R. H. Cuenca, M. L. Kavvas, D. R. Nielsen, and M. Mata (1992), Physical basis for a time series model of soil water content, *Water Resour. Res.*, *28*(9), 2437–2446.

Von Asmuth, J. R., and M. F. P. Bierkens (2005), Modeling irregularly spaced residual series as a continuous stochastic process, *Water Resour. Res.*, *41*, W12404, doi:10.1029/2004WR003726.

Von Asmuth, J. R., and M. Knotters (2004), Characterising groundwater dynamics based on a system identification approach, *J. Hydrol.*, *296*, 118–134.

Von Asmuth, J. R., and C. Maas (2001), The method of impulse response moments: A new method integrating time series-, groundwater- and eco-hydrological modelling, in *Impact of Human Activity on Groundwater Dynamics, IAHS Publ.*, *269*, 51–58.

Von Asmuth, J. R., M. F. P. Bierkens, and C. Maas (2002), Transfer function-noise modeling in continuous time using predefined impulse response functions, *Water Resour. Res.*, *38*(12), 1287, doi:10.1029/2001WR001136.

Von Asmuth, J. R., C. Maas, M. Bakker, and J. Petersen (2008), Modeling time series of ground water head fluctuations subjected to multiple stresses, *Ground Water*, *46*(1), 30–40.

Yu, C., A. W. Warrick, and M. H. Conklin (1999), A moment method for analyzing breakthrough curves of step inputs, *Water Resour. Res.*, *35*(11), 3567–3572.

Zwillinger, D. I. (1997), *Handbook of Differential Equations*, 3rd ed., Academic, Orlando, Fla.

M. Bakker, K. Maas, and J. R. Von Asmuth, Water Resources Section, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Stevinweg 1, 2628 CN Delft, Netherlands. (mark.bakker@tudelft.nl)