Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Hydrology 278 (2003) 1–16
Journal of Hydrology
www.elsevier.com/locate/jhydrol

Decoupling of modeling and measuring interval in groundwater time series analysis based on response characteristics

W.L. Berendrecht[a,*], A.W. Heemink[a], F.C. van Geer[b], J.C. Gehrels[b,c]

[a]*Department of Applied Mathematical Analysis, Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, P.O. Box 5031, 2600 GA Delft, The Netherlands*
[b]*Netherlands Institute of Applied Geoscience TNO, National Geological Survey, P.O. Box 80015, 3508 TA Utrecht, The Netherlands*
[c]*Department of Watermanagement, Faculty of Civil Engineering and Geosciences, Delft University of Technology, P.O. Box 5048, 2600 GA Delft, The Netherlands*

Received 20 March 2002; accepted 28 February 2003

Abstract

A state–space representation of the transfer function–noise (TFN) model allows the choice of a modeling (input) interval that is smaller than the measuring interval of the output variable. Since in geohydrological applications the interval of the available input series (precipitation excess) is often smaller than the interval of the output series (groundwater head), the state–space model opens the way to a more detailed description of the system. This paper evaluates the influence of the reduction of the modeling interval on the performance of the state–space time series model while keeping the measuring interval fixed. In order to obtain general conclusions of the relation between the modeling interval and the model performance, a large number of groundwater time series are generated and modeled with the state–space time series model. The results show that a reduction of the modeling interval noticeably improves the model performance. The degree of improvement depends on aspects like the response time of the system, the length of the time series and the amount of noise. A case study illustrates the effect of reducing the modeling interval as well as that of adding high-frequency measurements to the time series.

Keywords: Time series analysis; Modeling interval; Transfer function; Groundwater; Kalman filter

1. Introduction

Time series models (Box and Jenkins, 1970), and especially transfer function–noise (TFN) models, have been applied to analyze hydrological systems for many years (Hipel and McLeod, 1994; Van Geer and Zuur, 1997; Young et al., 1997). In groundwater hydrology, the main applications of TFN modeling are decomposition of groundwater level fluctuations into natural and anthropogenic fluctuations (Van Geer and Defize, 1987; Gehrels et al., 1994) and prediction of the effects of interventions (Knotters and Bierkens, 2000).

Besides the TFN model form proposed by Box and Jenkins, there is an alternative way to describe a dynamic system, namely by using the state–space form (Schweppe, 1973; Maybeck, 1979). This formulation is a very powerful and flexible representation of a system and has been successfully

* Corresponding author.
E-mail address: w.berendrecht@nitg.tno.nl (W.L. Berendrecht).

0022-1694/03/$ - see front matter © 2003 Elsevier Science B.V. All rights reserved.
doi:10.1016/S0022-1694(03)00075-1

applied in many problem areas, such as economics (Harvey, 1990; Young, 1994), catchment hydrology (O'Connell, 1980), and groundwater hydrology (Van Geer and Zuur, 1997; Bierkens et al., 2001). Shumway and Stoffer (1982) and more recently Bierkens et al. (1999) show that the state–space form is particularly useful for filling gaps of irregularly or sparsely observed time series. In the context of TFN modeling this means that the modeling interval is decoupled from measuring interval (interval of the output series) and based on the interval of the input series.

A perhaps even more important reason for decoupling the modeling and measuring interval is that a reduction of the modeling interval may greatly improve the descriptive performance of the transfer model. Clearly, a reduction of the modeling interval implies a finer discretization of the transfer function, resulting in a better approximation to the underlying process. In the Netherlands, for example, groundwater head is measured bimonthly on the 14th and 28th of each month. This means that the measuring interval varies between 14 and 17 days. Applying the state–space form of the TFN model, the modeling interval can be set to 1 day (provided that daily observations of precipitation excess are available) and the groundwater time series can be modeled more accurately.

The objective of this paper is therefore to determine the influence of a reduction of the modeling interval on the performance of the state–space time series model. In this context, performance is defined as how well the transfer function can be estimated and hence the fluctuations caused by the input series can be filtered out of the output series. For this purpose a large number of representative time series were generated, using a range of predefined transfer functions. Also, a stochastic component was added to the series to make the time series similar to real groundwater time series. Evaluation of the calibrated time series models for different measuring and modeling intervals shows that a reduction of the modeling interval generally improves the model. The rate of improvement depends on several variables, such as the system noise and the response time of the system. A validation study confirms these results, but is not discussed in this paper. A real-world model of irregularly observed groundwater head data illustrates and confirms these results.

2. Theoretical background

This section shortly describes the modeling framework that is used in the experiments presented in Section 3. A state–space form of the TFN model is applied to describe the groundwater system. The state–space representation of a system is a very generic formulation and is widely applied in system and control theory. Moreover, it allows the use of the well-known Kalman filter. The unknown parameters can then be estimated by combining the Kalman filter with a maximum likelihood criterion.

2.1. State–space model

The state–space form (Schweppe, 1973) of a linear single-input single-output (SISO) system consists of two equations:

$$\mathbf{x}_t = \mathbf{A}_t\mathbf{x}_{t-1} + \mathbf{b}_t u_t + \mathbf{g}_t w_t, \quad t = 1,\ldots,T, \tag{1}$$

$$y_t = \mathbf{c}_t\mathbf{x}_t + y_r + v_t, \quad t = 1,\ldots,T, \tag{2}$$

where $\mathbf{x}_t$ is the unobserved, m-dimensional system state; the $\mathbf{A}_t$ is an $m \times m$ transition matrix; $\mathbf{b}_t$ is an $m \times 1$ vector; u_t is a scalar representing the system input at time t; $\mathbf{g}_t$ is an $m \times 1$ vector that relates the noise statistics to the system states; $\mathbf{c}_t$ is a $1 \times m$ vector; y_t is a measurement available at time t; y_r is a constant depending on the reference level of the time series; w_t and v_t are scalars representing the system noise and measurement noise, respectively, with the following properties:

$$w \simeq N(0,q), \qquad v \simeq N(0,r), \qquad E\{w_t v_t\} = 0. \tag{3}$$

Based on Eqs. (1) and (2) the state–space representation of the linear TFN model is written as

$$\begin{bmatrix}\mathbf{x}_d\\ \mathbf{x}_s\end{bmatrix}_t = \begin{bmatrix}\mathbf{A}_d & 0\\ 0 & \mathbf{A}_s\end{bmatrix}\begin{bmatrix}\mathbf{x}_d\\ \mathbf{x}_s\end{bmatrix}_{t-1} + \begin{bmatrix}\mathbf{b}_d\\ 0\end{bmatrix}u_t + \begin{bmatrix}0\\ \mathbf{g}_s\end{bmatrix}w_t, \tag{4}$$

$$y_t = [\mathbf{c}_d\ \mathbf{c}_s]\begin{bmatrix}\mathbf{x}_d\\ \mathbf{x}_s\end{bmatrix}_t + y_r + v_t, \tag{5}$$

with

$$\mathbf{A}_d = \begin{bmatrix} \delta_1 & 1 & 0 & \cdots & 0 \\ \delta_2 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \delta_{r-1} & \vdots & & \ddots & 1 \\ \delta_r & 0 & \cdots & \cdots & 0 \end{bmatrix}, \tag{6}$$

$$\mathbf{b}_d^{\mathrm{T}} = \begin{bmatrix} \omega_0 & \omega_1 & \cdots & \omega_{r-2} & \omega_{r-1} \end{bmatrix},$$

$$\mathbf{c}_d = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix},$$

$$\mathbf{A}_s = \begin{bmatrix} \phi_1 & 1 & 0 & \cdots & 0 \\ \phi_2 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \phi_{p-1} & \vdots & & \ddots & 1 \\ \phi_p & 0 & \cdots & \cdots & 0 \end{bmatrix}, \tag{7}$$

$$\mathbf{g}_s^{\mathrm{T}} = \begin{bmatrix} 1 & \theta_1 & \cdots & \theta_{p-2} & \theta_{p-1} \end{bmatrix},$$

$$\mathbf{c}_s = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}.$$

The elements of $\mathbf{x}_d$ describe the deterministic component, z, of the system (i.e. that part of the system that can be related to the model input): $x_{d,1} = z_t$, $x_{d,2} = z_{t-1}, \ldots, x_{d,r} = z_{t-r+1}$. In a similar way $\mathbf{x}_s$ describe the stochastic component, ξ, of the system (which cannot be related to the model input): $x_{s,1} = \xi_t$, $x_{s,2} = \xi_{t-1}, \ldots, x_{s,r} = \xi_{t-p+1}$. Besides the reference level y_r and measurement noise r, the observation y_t thus consists of z_t and ξ_t.

The parameters in the system matrices are unknown and defined as follows: δ_i and ω_j represent, respectively, the ith autoregressive parameter and the jth moving average parameter of the transfer model; ϕ_k and θ_l represent, respectively, the kth autoregressive parameter and the lth moving average parameter of the noise model. This set of parameters together with the variance of the system noise q, and the reference level y_r, will be referred to as the parameter set α. The variance of the measurement noise, r, is assumed to be known.

2.2. State estimation

The state $\mathbf{x}_t$ is estimated from the observations $y_1, \ldots, y_t$ using the Kalman filter, which is a recursive procedure for computing the optimal estimator of the state vector at time t, based on the information available at time t. The Kalman filter algorithm for the state equation (4) consists of the following equations (Bierkens et al., 1999):

Initial conditions:

$$\hat{\mathbf{x}}_0 \text{ and } \mathbf{P}_0. \tag{8}$$

Time update:

$$\bar{\mathbf{x}}_t = \mathbf{A}_t \hat{\mathbf{x}}_{t-1} + \mathbf{b}_t u_t, \tag{9}$$

$$\mathbf{M}_t = \mathbf{A}_t \mathbf{P}_{t-1} \mathbf{A}_t^{\mathrm{T}} + \mathbf{g}_t q \mathbf{g}_t^{\mathrm{T}}. \tag{10}$$

Measurement update:

$$n_t = y_t - y_r - \mathbf{c}_t \bar{\mathbf{x}}_t, \tag{11}$$

$$f_t = \mathbf{c}_t \mathbf{M}_t \mathbf{c}_t^{\mathrm{T}} + r, \tag{12}$$

$$\mathbf{K}_t = \mathbf{M}_t \mathbf{c}_t^{\mathrm{T}} f_t^{-1}, \tag{13}$$

$$\hat{\mathbf{x}}_t = \bar{\mathbf{x}}_t + \mathbf{K}_t n_t, \tag{14}$$

$$\mathbf{P}_t = (\mathbf{I} - \mathbf{K}_t \mathbf{c}_t) \mathbf{M}_t, \tag{15}$$

where $\hat{\mathbf{x}}_t$ is the measurement update; $\bar{\mathbf{x}}_t$ is the time update; $\mathbf{P}_t$ is the covariance matrix of the error in the measurement update: $\mathrm{cov}(\mathbf{x}_t - \hat{\mathbf{x}}_t)$; $\mathbf{M}_t$ is the covariance matrix of the error in the time update: $\mathrm{cov}(\mathbf{x}_t - \bar{\mathbf{x}}_t)$; n_t is the innovation; f_t is the innovation variance; $\mathbf{K}_t$ is the Kalman gain; and $\mathbf{I}$ is the identity matrix. If at time step t no measurement is available,

$$\hat{\mathbf{x}}_t = \bar{\mathbf{x}}_t, \tag{16}$$

$$\mathbf{P}_t = \mathbf{M}_t. \tag{17}$$

2.3. Parameter estimation

A basic assumption in the filtering algorithm described in Section 2.2 is that the system matrices together with $\hat{\mathbf{x}}_0$ and $\mathbf{P}_0$ are known. However, the parameter vector α is unknown and

needs to be estimated by evaluating the log-likelihood function for the innovations of the Kalman filter:

$$\log L(N;\alpha) = -\frac{N}{2}\log 2\pi - \frac{1}{2}\sum_{t=1}^{N}\log f_t(\alpha) - \frac{1}{2}\sum_{t=1}^{N}\frac{n_t^2(\alpha)}{f_t(\alpha)}. \quad (18)$$

In order to reduce the number of parameters to be estimated, Harvey (1990) suggests multiplication of the variances in the model by a scaling factor. If the scaling variance is

$$\sigma_*^2 = \mathrm{var}(w_t), \quad (19)$$

then r is scaled to

$$r_* = \frac{\mathrm{var}(v_t)}{\sigma_*^2}, \quad (20)$$

and

$$q_* = \frac{\mathrm{var}(w_t)}{\sigma_*^2} = 1. \quad (21)$$

As a result, the dimension of α reduces by 1. Writing down the log-likelihood in terms of these newly defined parameters gives

$$\log L(N;\alpha) = -\frac{N}{2}\log 2\pi - \frac{N}{2}\log\sigma_*^2(\alpha) - \frac{1}{2}\sum_{t=1}^{N}\log f_t(\alpha) - \frac{1}{2\sigma_*^2(\alpha)}\sum_{t=1}^{N}\frac{n_t^2(\alpha)}{f_t(\alpha)}. \quad (22)$$

The advantage of this is that Eq. (22) can be maximized with respect to $\sigma_*^2(\alpha)$ by setting the derivative to zero, resulting in

$$\sigma_*^2(\alpha) = \frac{1}{N}\sum_{t=1}^{N}\frac{n_t^2(\alpha)}{f_t(\alpha)}. \quad (23)$$

If $\sigma_*^2(\alpha)$ is substituted in Eq. (22), the 'reduced' log-likelihood function is obtained:

$$\log L(N;\alpha) = -\frac{N}{2}(\log 2\pi + 1) - \frac{1}{2}\sum_{t=1}^{N}\log f_t(\alpha) - \frac{N}{2}\log\sigma_*^2(\alpha). \quad (24)$$

A sequential quadratic programming (SQP) method (Gill et al., 1981) is used to optimize Eq. (24). The required Jacobian of the parameter vector α is calculated by evaluating the derivatives of f_t and n_t analytically (through differentiating Eqs. (9)–(17), running in parallel with the Kalman filter).

In absence of any prior information on the initial state, $\mathbf{x}_0$ is set to zero and $\mathbf{P}_0$ to κ times the identity matrix, where κ is a large number (e.g. 10,000). This large covariance matrix indicates that little is known about the initial state. As $\mathbf{x}_0$ and $\mathbf{P}_0$ are independent of α, $\partial\mathbf{x}_0/\partial\alpha_i = 0$ and $\partial\mathbf{P}_0/\partial\alpha_i = 0$. In practice, the first few observations are used to estimate the starting values, so the first prediction errors and variances should be omitted from the likelihood function. Janacek and Swift (1993) suggest to use the first m observations to estimate the m-dimensional state so that N is replaced by $N - m$ in Eqs. (23) and (24).

The covariance matrix of parameter estimation errors is estimated using the Cramer–Rao lower bound (Schweppe, 1973):

$$R_C^{-1} = \frac{\partial^2 \log L(N;\alpha)}{\partial\alpha\partial\alpha^{\mathrm{T}}}, \quad (25)$$

where R_C is the Cramer–Rao lower bound.

Finally, several tests have to ensure that the fitted model adequately describes the time series under consideration. The main diagnostics are based on the innovations obtained by the Kalman filter. The whiteness of the innovations can be tested using the autocorrelation function of the innovations. Harvey (1990) gives an estimation of the autocorrelation function of innovations containing missing values:

$$r(k) = \left[\sum_{t=m+1}^{N^\dagger}\frac{\tilde{n}_t^\dagger\tilde{n}_{t-k}^\dagger}{N(k)}\right]\left[\sum_{t=m+1}^{N^\dagger}\frac{\tilde{n}_t^{\dagger 2}}{(N-m)}\right]^{-1}, \quad (26)$$

where k is the time lag; $N^\dagger$ is the length of the series generated by the underlying model; $N(k)$ is the number of non-zero cross-products of innovations in the numerator of the statistic; and $\tilde{n}_t^\dagger$ denotes the standardized innovation defined by

$$\tilde{n}_t^\dagger = \frac{n_t}{f_t^{1/2}}, \quad t = m+1,\ldots,N \quad (27)$$

if y_t is observed and is set to zero for all other values of t for $t = m+1,\ldots,N^\dagger$. The cross-correlation between input series and innovations as well as the cross-correlation between input series

and residuals (i.e. estimated stochastic component) provide useful information on the correctness of the model structure.

3. Description of experiment

The effect of reducing the modeling interval is evaluated using generated time series. The advantage of a generated series is that the relation (transfer function) between the input and output is exactly known, enabling an accurate evaluation of the model performance. This section describes how time series were generated. Also it describes the criteria that were used to evaluate the model results.

3.1. Generation of groundwater time series

Groundwater time series were generated by transferring a daily input series of precipitation excess, using a predefined transfer function described by the probability density function of a lognormal distribution:

$$\Psi(t) = \frac{c}{t\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\ln t - \mu}{\sigma} \right)^2 \right], \quad \sigma > 0, 0 < t < \infty, \tag{28}$$

where μ and σ are the geometric mean and standard deviation of the distribution, respectively, and c is a scaling constant. The main reasons for choosing this function are its flexibility (continuous in time) and the fact that the response of many hydrological systems to precipitation can be described by an exponential function. The deterministic component z_t of the generated time series can thus be written as

$$z_t = \sum_{\tau=1}^{t-1} \Psi_\tau P_{e,t-\tau}. \tag{29}$$

Here, P_e is the precipitation excess [LT^{-1}]:

$$P_e = P - f_p E_0, \tag{30}$$

where P is the precipitation [LT^{-1}]; E_0 is the Penman open water evaporation [LT^{-1}]; and f_p is a crop factor [−] which was set to a value of 0.8. Both precipitation and evaporation were obtained from daily observations at the main meteorological station of the Royal Netherlands Meteorological Institute at De Bilt in the period from July 1, 1957 to December 31, 1999.

In addition to the deterministic transfer function, a stochastic component was added to the system. This component represents the part of the system dynamics that is not related to the input signal. The stochastic component is assumed to be described by the following autoregressive model:

$$\xi_t = \phi \xi_{t-1} + a_t, \tag{31}$$

where a_t is normally distributed with zero mean and variance σ_a^2, which was set to 1% of the variance of the deterministic component. A number of 20 independent realizations of each stochastic component were generated to obtain a statistically correct experiment.

Using Eqs. (29) and (31) many different time series were generated, varying from fast responding systems (peak response within a couple of days) to very slow responding systems (peak response after one year). Fig. 1 gives some transfer functions used in this experiment. This paper only describes the results of the time series generated with the transfer curve in Fig. 1 depicted in bold, having a time of peak response $t_p = 43$ days. Conclusions from this experiment are similar to those of the other analyzed transfer functions.

Each deterministic component was combined with two different stochastic components. Table 1 gives the parameters of this transfer function as well as the parameters of the stochastic component. Here, the ratio S_N is defined as

$$S_N = \frac{\mathrm{var}(\xi_t)}{\mathrm{var}(z_t)}, \tag{32}$$

where $\mathrm{var}(\xi_t)$ is the variance of the stochastic component, and $\mathrm{var}(z_t)$ is the variance of the deterministic component. Table 1 shows that, as a result of the higher value of ϕ, the value of S_N of series S2 is higher.

Fig. 2 shows one realization of time series S2. The time series was finally split into a calibration period (1957–1989) and a validation period (1990–1999) to validate the results obtained by the calibration.

3.2. Resampling of time series

A thorough analysis of the relation between modeling interval and model performance not only

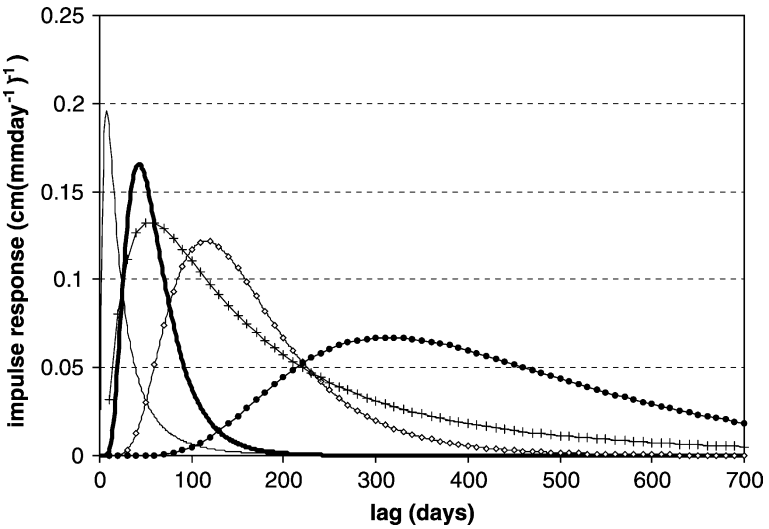

Fig. 1. Examples of transfer functions applied to generate time series. The bold curve is used in this paper.

requires a range of transfer functions (as described in Section 3.1) but also a number of other variables that influence the relation between modeling interval and model performance. Table 2 gives an overview of the variables that were taken into account and the range of variation that was analyzed in this paper. Basically, these time series were modeled with two different model forms: model 1 has a modeling interval equal to the measuring interval; model 2 has a modeling interval of 10 days.

3.3. Evaluation criteria

Before comparing models of different time series and with different modeling intervals, the number of parameters of each model needs to be selected. In practical applications, well-known criteria are the Akaike Information Criterion (AIC), Bayes Information Criterion (BIC) as well as the innovation variance of the calibrated and validated time series. In this experiment, with generated time series and a known deterministic component, the 'fit' of the calibrated deterministic component of the model to the real deterministic component can be calculated. A good measure of 'fit' (or 'error') is the Mean Absolute Error (MAE) [L] of the deterministic component, described as

$$\mathrm{MAE} = \frac{1}{N-m}\sum_{t=m+1}^{N} |\hat{z}_t - z_t|, \tag{33}$$

where $\hat{z}_t$ represents the estimated deterministic component. In other words, the MAE quantifies how well the deterministic component can be separated from the stochastic component. Based on the MAE, the optimal number of parameters was selected for each time series model. The MAE-criterion was also applied to compare the models of the time series given in Table 2. A second criterion was used to indicate the accuracy of the estimated transfer function. For this purpose, a useful parameter is the standard deviation of the gain of the transfer function s_G [$\mathrm{L}(\mathrm{LT}^{-1})^{-1}$], where the gain G [$\mathrm{L}(\mathrm{LT}^{-1})^{-1}$] represents the area

Table 1
Parameters of predefined transfer function used to generate time series

Series	Deterministic component			Stochastic component	
	μ	σ	c	ϕ	S_N
S1	4.0	0.5	10	0	0.01
S2	4.0	0.5	10	0.99	0.37

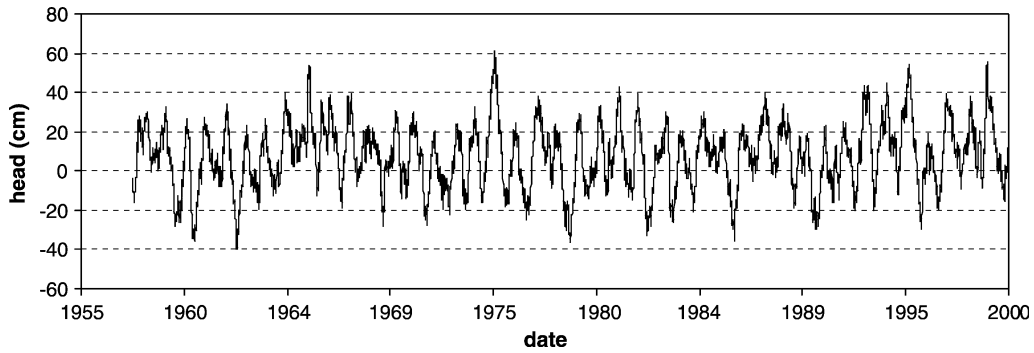

Fig. 2. Realization of time series generated with a predefined transfer function (log–normal probability density function) and an autoregressive stochastic component corresponding to S2 in Table 1.

under the impulse response curve:

$$G = \frac{\sum_i \omega_i}{1 - \sum_k \delta_k}, \tag{34}$$

with ω and δ as defined in Eq. (6).

4. Comparison of model results for varying modeling and measuring interval

This section discusses the results of the experiment. First, the influence of modeling interval, measuring interval, and length of time series on the fit of the deterministic component is evaluated. Next, the influence of these aspects on the parameter accuracy is analyzed. Finally, the effect of adding high-frequency measurements to an existing time series of low-frequency measurements is demonstrated.

4.1. Fit of deterministic component

Fig. 3 shows the fit of the deterministic component expressed in terms of MAE, for four different cases and seven measuring intervals. For the calculation of the MAE only the errors at measurement points were used. The four cases are interpreted as follows:

(1) S1—Model 1: $\phi = 0$; $\mathrm{d}t_{\mathrm{mod}} = \mathrm{d}t_{\mathrm{meas}}$. Fig. 3 shows that the performance of the model improves as the measuring interval decreases: from a mean absolute error of ca. 3.5 cm to an error of ca. 0.5 cm. A smaller measuring interval (and thus a smaller modeling interval) clearly allows for a finer discretization and hence a better approximation of the transfer function. The curve of MAE flattens from the moment the measuring interval becomes larger than the time of peak response t_{p}. From this point, a coarser discretization of the transfer function will hardly influence the model fit. This is directly related to the gradient of the transfer function. The rising limb of the transfer function is steep and requires a small interval, whereas the falling limb of the transfer function has a smaller gradient and can therefore be well approximated by a model with a large interval.

(2) S1—Model 2: $\phi = 0$; $\mathrm{d}t_{\mathrm{mod}} = 10$. In this case, the same time series S1 is modeled, but with the modeling interval fixed at 10 days. Fig. 3 shows that the performance of this model is much better and in fact practically insensitive to the measuring interval. This is easily understood by realizing again that a smaller modeling interval gives a finer discretization

Table 2
Variation of variables applied for resampling of the time series

Variable	Range of variation
Measuring interval	10, 20,…,70 days
Length of time series	10, 20,…,100% of maximum length (1957–1989)
Varying measuring interval	First fraction (0, 10, 20,…,100%) of time series has measuring interval of 70 days, last fraction a measuring interval of 10 days; modeling interval is 10 days

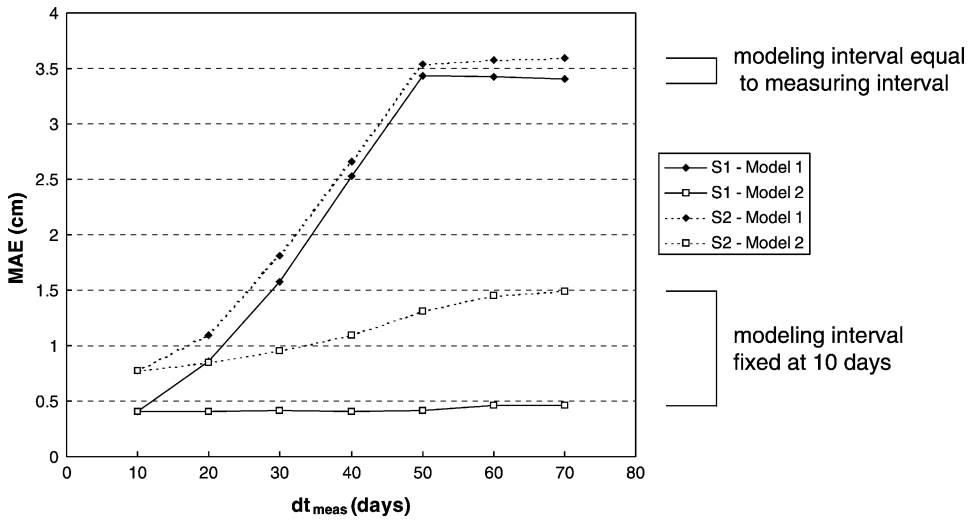

Fig. 3. Fit of deterministic component (mean absolute error, MAE) as a function of measuring interval, for series S1 using 20 realizations of a white-noise stochastic component ($\phi = 0$) and S2 using 20 realizations of an autoregressive stochastic component ($\phi = 0.99$). Model 1 has a modeling interval equal to the measuring interval, whereas Model 2 keeps the modeling interval fixed at 10 days.

of the transfer function, resulting in a better approximation.

(3) S2—Model 1: $\phi = 0.99$; $dt_{mod} = dt_{meas}$. The difference between series S1 and S2 is that the stochastic component of Series S2 has an autoregressive part, resulting in a larger value of S_N. Consequently, the fit of the deterministic component of series S2 is not as good as for S1. The difference, however, is limited compared to the effect of the measuring interval.

(4) S2—Model 2: $\phi = 0.99$; $dt_{mod} = 10$. Again, a reduction of the modeling interval greatly improves the performance of the model. However, the MAE slowly increases as the measuring interval increases, whereas the curve of MAE flattens from the moment the measuring interval becomes larger than t_p. It can therefore be concluded that the influence of the modeling interval on MAE increases as the correlation length (i.e. the value of ϕ) of the stochastic component increases.

Another important aspect that influences the fit of the deterministic component is the length of the time series. In Fig. 3 the length of the time series was fixed at 33 years. In the context of this paper an important question is whether a reduction of the modeling interval still improves the fit if the time series is short. Fig. 4 shows the relation between the MAE and the length of the time series. Only the results for series S2 are presented, because this series has a large stochastic component.

For a length of 100% the curve of the MAE is the same as the curve in Fig. 3. If the modeling interval is equal to the measuring interval (Fig. 4a), the MAE first hardly increases as the length decreases. Only if the length is reduced to less than 50% (16.5 years) the increase becomes noticeable. The pattern of the curve becomes a little irregular for combinations of large measuring interval and short time series. Much more realizations would be necessary to obtain a smooth curve. These combinations of measuring interval and length of time series, however, result in datasets with an unrealistic small number of measurements. For example, the worst combination (measuring interval is 70 days and length is 10% of the full length) consists of only 16 measurements. For all other combinations, a reduction of the modeling interval clearly improves the fit of the deterministic component.

Bierkens et al. (1999) found as well (for a time series with $t_p = 1$ day) that similar parameter values can be obtained for different measuring intervals as long as the measuring interval is smaller than the characteristic response time (i.e. time for which the response is only 5% of the maximum response).

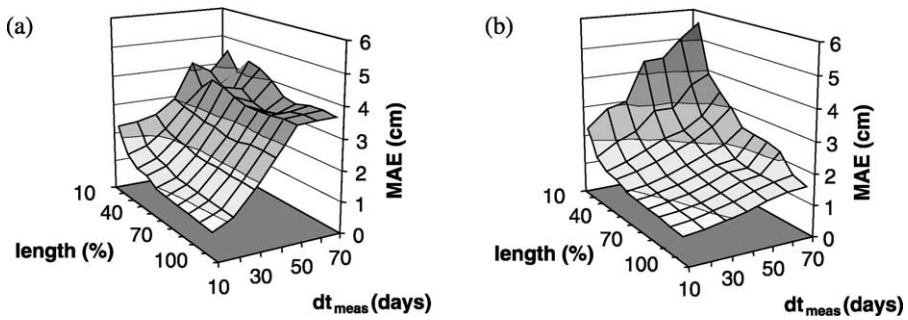

Fig. 4. Fit of deterministic component (mean absolute error, MAE) as a function of measuring interval and length of the time series (as a percentage of the total length, i.e. 33 years), for series S2 and (a) Model 1 having a modeling interval equal to the measuring interval; and (b) Model 2 having a modeling interval of 10 days. The results are based on 20 realizations of the stochastic component.

On the basis of the present study more general statements can be made, summarized as follows:

(1) A reduction of the modeling interval results in a better fit of the deterministic component, regardless the measuring interval;
(2) The effect of a reduction of the modeling interval increases for larger measuring intervals;
(3) For a small stochastic component, the fit is practically insensitive to the measuring interval if a small modeling interval (i.e. small with respect to the time of peak response) is used;
(4) For a large stochastic component, the fit slightly decreases with increasing measuring interval if a small modeling interval is used.

These results are of great practical importance, because the models of existing groundwater time series can simply be improved by only reducing the modeling interval.

4.2. Parameter accuracy

The overall accuracy of the estimated parameters, expressed by the standard deviation of the gain (s_G), is evaluated in relation to the measuring interval and the length of the time series. Fig. 5 shows this relation for the same four cases as in Section 4.1. In this context, it is important to note that the estimated gain is about 10 cm(mmday^{-1})$^{-1}$ (= c in Eq. (28)). The figures are interpreted as follows:

(1) Fig. 5a. S1—Model 1: $\phi = 0$; $\mathrm{d}t_{\mathrm{mod}} = \mathrm{d}t_{\mathrm{meas}}$. The figure shows that the standard deviation of the gain decreases considerably with decreasing measuring interval. This is not surprising because more measurements result in a more accurate estimation of the parameters. In addition, the curve of s_G has the same pattern as found for the MAE: a rather fast increase for measuring intervals smaller than t_{p} and a flattening of the curve for measuring intervals larger than t_{p}. Finally, the influence of the measuring interval increases with decreasing length of the time series. High standard deviations occur for short time series with large measuring interval.

(2) Fig. 5b. S1—Model 2: $\phi = 0$; $\mathrm{d}t_{\mathrm{mod}} = 10$. Similar to the results presented for the MAE, a fixed small modeling interval results in better models for all measuring intervals in the sense that the transfer function is estimated more accurately. However, a small increase of s_G can still be observed as the measuring interval increases. This is directly related to the fact that the number of measurements reduces as the measuring interval increases.

(3) Fig. 5c and d. S2 ($\phi = 0.99$)—Model 1 ($\mathrm{d}t_{\mathrm{mod}} = \mathrm{d}t_{\mathrm{meas}}$) and Model 2 ($\mathrm{d}t_{\mathrm{mod}} = 10$). The pattern of s_G is similar to the pattern of series S1. However, the influence of the measuring interval on s_G is small with respect to the influence of the stochastic component on s_G.

Summarizing, a reduction of the modeling interval has a positive effect on the accuracy of the estimated transfer function. The relative improvement (improvement with respect to its original value) depends on the contribution of the stochastic component.

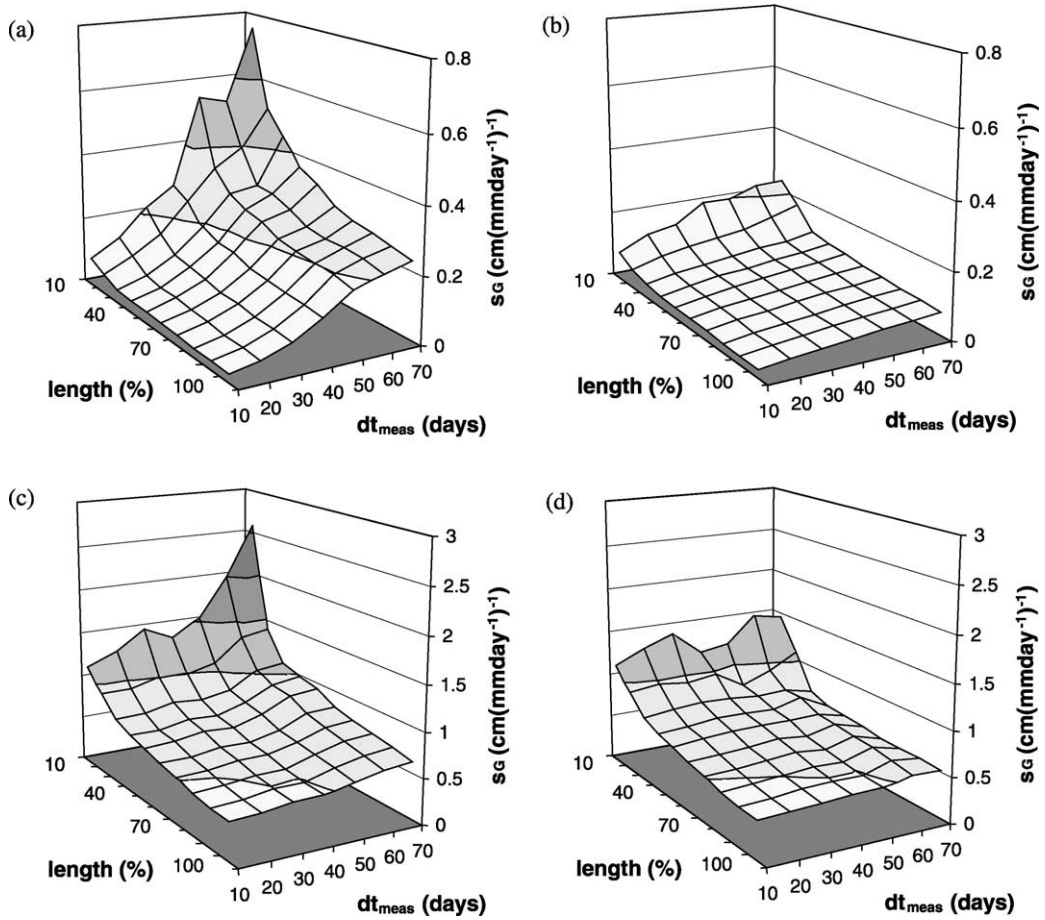

Fig. 5. Relation between parameter accuracy, measuring interval and length of time series (relative to the original length of the time series (33 years) for (a) series S1, Model 1; (b) series S1, Model 2; (c) series S2, Model 1; and (d) series S2, Model 2. The results are based on 20 realizations of the stochastic component.

4.3. Adding high-frequency measurements

In practice, a large data set of groundwater measurements is often available. The previous sections have shown that a reduction of the modeling interval can improve the model performance considerably. Another way to improve the model performance is to extend the time series with a set of high-frequency measurements obtained with automatic data loggers. This section evaluates the influence of such an extension on the model performance. MAE and s_G are used again as evaluation criteria.

Fig. 6 shows the effect of adding high-frequency measurements to the time series S1 and S2 used earlier. The horizontal axis represents the fraction of the time series that has a measuring interval of 10 days (instead of 70 days). Hence, if $t_{10}/t_{tot} = 0$ the whole time series has a measuring interval of 70 days, whereas $t_{10}/t_{tot} = 0.1$ means that the first 90% of the time series has a measuring interval of 70 days and the final 10% a measuring interval of 10 days. The MAE is calculated at each model time step to obtain a valid comparison between the models.

Fig. 6a shows that the first high-frequency measurements (between 0% and 20% of the total

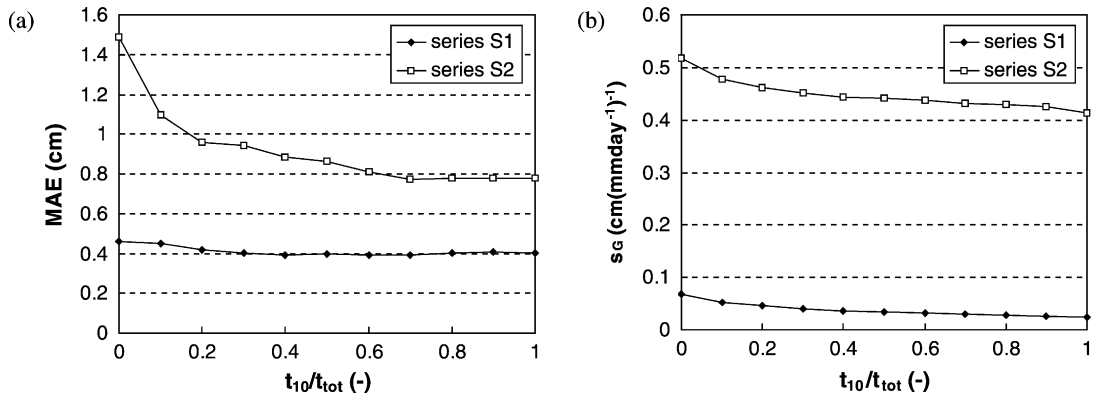

Fig. 6. Influence of reducing the measuring interval from 70 to 10 days in the final t_{10}/t_{tot} fraction of time series S1 and S2 on (a) the fit of the deterministic component (MAE); and (b) the parameter accuracy (s_G). The modeling interval is 10 days and the length of the series is 33 years. The results are based on 20 realizations of the stochastic component.

length of the series) are the most effective in terms of an increase in the model performance, especially for series S2. This is very attractive, because it means that relatively few extra measurements are needed to reduce the MAE considerably. The curves of s_G (Fig. 6b) are different: instead of a quick drop during the first 20%, the curves display a steady decrease of s_G. One of the reasons is that s_G depends more on the total number of measurements and length of the time series than on the measuring interval. The same experiment was applied to shorter time series, showing the same results.

Summarizing, the model performance can be further improved by extending (low-frequency) time series with high-frequency measurements. The effect of this extension becomes larger for time series with large stochastic components. This section only evaluated the influence of high-frequency measurements at the ending of a period, because there are already many time series with low-frequency measurements available. If one starts monitoring, however, it could be advantageous to start with high-frequency measurements. This topic will not be discussed in detail, but several calculations have shown that high-frequency measurements at the beginning of the period have more effect than at the end of the period. The reason for this is that the Kalman filter is a forward scheme.

5. Case study

5.1. Description of the data set

The groundwater head data (1990–2000) were obtained from an observation well (phreatic aquifer) in the east of the Netherlands. The measuring frequency of this series is 24 observations per year. In 1999, the measuring frequency has been increased to one observation per day during a period of three months (which is about 2.5% of the total length of the time series). This time series is therefore a good example to examine the practical significance of the results reported in the previous sections. The series was split into a calibration series (from 1990 to June 1999, which is at the end of the daily measurements) and a validation series (from July 1999 to October 2000). The daily meteorological input data were obtained from two meteorological stations: the precipitation from a nearby station at Eerbeek (5 km distance) and the potential evapotranspiration from a station at De Bilt (65 km distance).

From this time series three different samples were selected:

(1) Measuring interval = 14 days, modeling interval = 14 days,
(2) Measuring interval = 14 days, modeling interval = 1 day,

(3) Measuring interval = 14 days and 1 day, modeling interval = 1 day.

5.2. Modeling results

Identification of the system resulted in the following model form for all three samples:

$$\begin{bmatrix} \mathbf{x}_{d,1} \\ \mathbf{x}_{d,2} \\ \mathbf{x}_{s,1} \end{bmatrix}_t = \begin{bmatrix} \delta_1 & 1 & 0 \\ \delta_2 & 0 & 0 \\ 0 & 0 & \phi_1 \end{bmatrix} \begin{bmatrix} \mathbf{x}_{d,1} \\ \mathbf{x}_{d,2} \\ \mathbf{x}_{s,1} \end{bmatrix}_{t-1} + \begin{bmatrix} \omega_0 \\ \omega_1 \\ 0 \end{bmatrix} u_t + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} w_t,$$

$$y_t = [1\ 0\ 1] \begin{bmatrix} \mathbf{x}_{d,1} \\ \mathbf{x}_{d,2} \\ \mathbf{x}_{s,1} \end{bmatrix}_t + y_r + v_t,$$

with elements as defined in Eqs. (4)–(7).

Table 3 lists the estimated parameters of the three models. The graphical output is given in Figs. 7 and 8. Since the difference between the estimated deterministic parameters (δ_1, δ_2, ω_0, ω_1) of series 2 and series 3 is not significant (see Table 3), the result of series 2 is omitted. Note that due to the extra measurements in series 3 the standard deviations of the parameters (except y_r) decrease.

The results in both figures allow for some general statements. A small modeling interval enables a much better approximation to the peak response (compare Figs. 7d and 8d). As a result, the extreme values of the time series are modeled clearly better (compare the peaks in 1994 and 1999 in Figs. 7a and 8a). Also, Fig. 8c shows that the daily fluctuations of groundwater head are modeled rather well. Only a slight local trend is observed. A possible explanation of this trend is that the assumption of linearity is incorrect.

The difference in model performance has to be quantified by criteria such as the innovation variance of the calibrated series, var($n_{c,t}$), the innovation variance of the validated series, var($n_{v,t}$), and the standard deviation of the gain, s_G. As different modeling intervals are used, comparison of the variance of the system noise $q = \mathrm{var}(w_t)$ is not useful. Instead, the variance of the stochastic component is compared:

$$\mathrm{var}(\xi_t) = \frac{q}{1-\phi_1^2}. \tag{35}$$

Table 4 evaluates the model performance, using these criteria. Before comparing the results it is important to remark that the coefficient of determination, R_T^2, is based on all measurements. Hence, the value of R_T^2 of series 3 is based on more measurements and thus not directly comparable to R_T^2 of series 1 and 2. If the daily measurements of series 3 would be ignored in the calculation of R_T^2, then $R_T^2 = 0.86$.

Table 4 shows the following:

(1) Reducing the modeling interval from 14 days to 1 day results in a reduction of var(ξ_t) of about 21%. The variance reduces another 3% if the daily measurements are added to the time series.
(2) The gain G increases slightly if the modeling interval is reduced to 1 day, which means that the area under the impulse response function increases. This is also expressed by an increase of the coefficient of determination.
(3) The accuracy of the estimated transfer function improves from series 1 to 2: s_G reduces by 22%. Adding daily measurements, however,

Table 3
Estimated parameters of the calibrated models. The standard deviation of the parameter estimation error is given in parentheses

Series	dt_{meas} (day)	dt_{mod} (day)	δ_1	δ_2	ϕ_1	ω_0 (cm(mmday^{-1})$^{-1}$)	ω_1 (cm(mmday^{-1})$^{-1}$)	y_r (cm)
1	14	14	1.0438(0.12)	−0.1698(0.092)	0.3906(0.044)	7.377(0.020)	−4.066(0.056)	−138.6(1.9)
2	14	1	1.8594(0.015)	−0.8609(0.015)	0.9417(0.0068)	0.8724(0.035)	−0.8317(0.034)	−140.1(1.6)
3	Mixed	1	1.8729(0.011)	−0.8742(0.010)	0.9578(0.0044)	0.8264(0.028)	−0.7905(0.027)	−139.8(1.8)

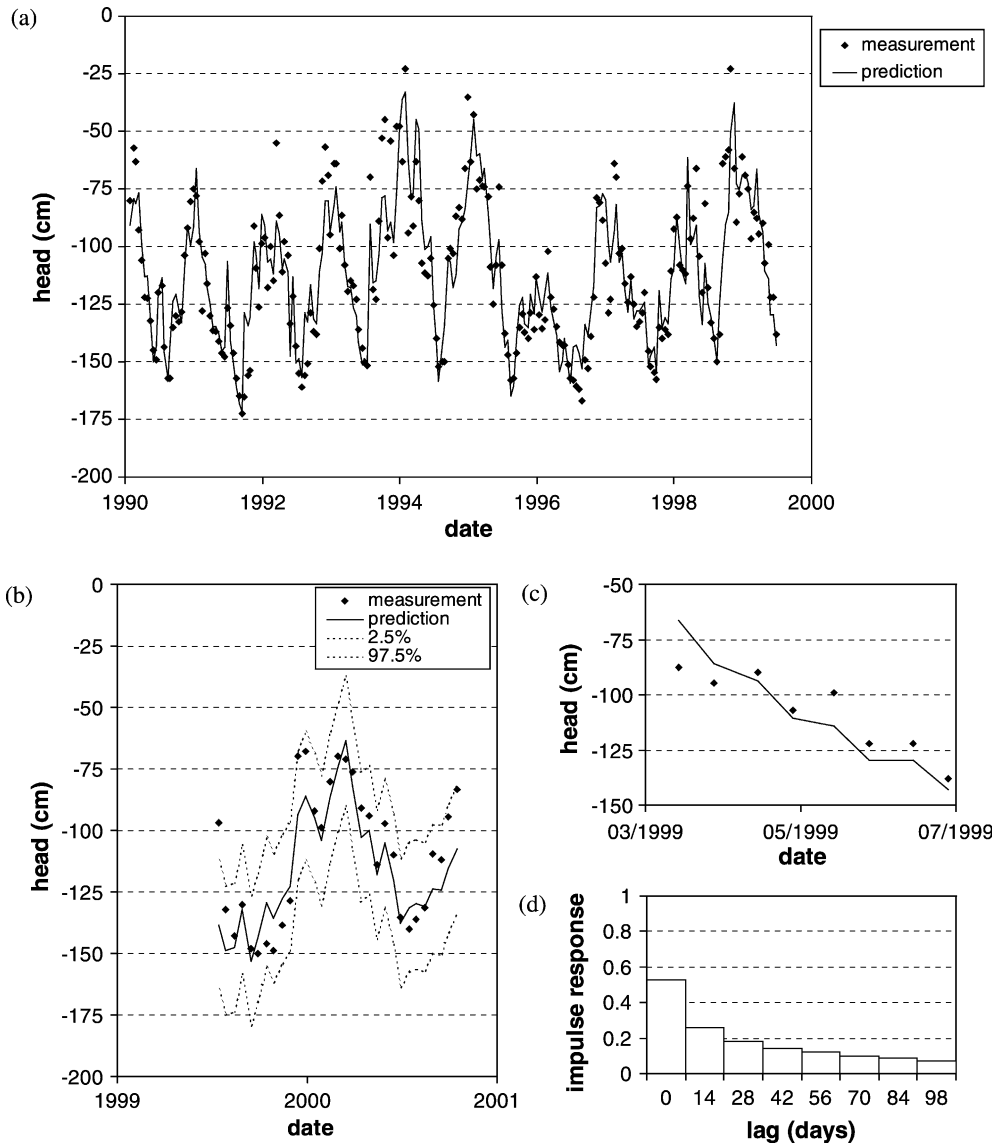

Fig. 7. Results of series 1; (a) measurements and predictions of the model for the calibration period; (b) measurements combined with the predictions and 95% prediction interval for the validation period; (c) measurements and predictions for the period of daily measurements (for comparison with Fig. 8); and (d) estimated impulse response function.

does not result in a reduction of s_G. The reason for this is the relatively large stochastic component and short series of high-frequency measurements.

This case study confirms that adjusting the modeling interval to the response time of the system, improves the model significantly. The model especially describes the extremes in the time series much better. As a result, the system noise reduces considerably. The noise could be reduced even further by adding high-frequency measurements. The improvements due to the relatively short series of high-frequency measurements is, however, small. A longer high-frequency time series will probably lead to a further improvement.

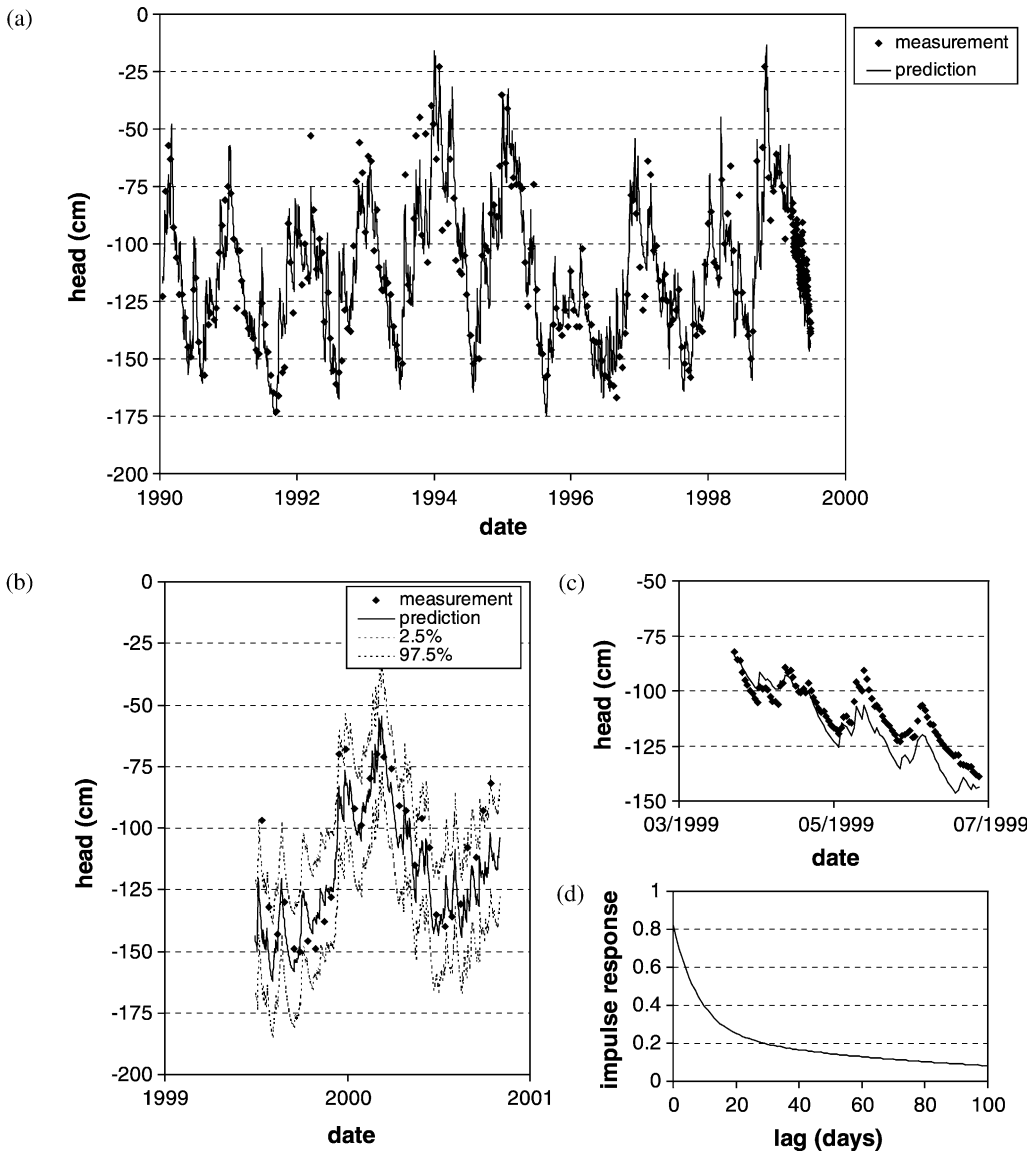

Fig. 8. Results of series 3; (a) measurements and predictions of the model for the calibration period; (b) measurements combined with the predictions and 95% prediction interval for the validation period; (c) measurements and predictions for the period of daily measurements; and (d) estimated impulse response function.

Table 4
Comparison of the criteria for the evaluation of the performance of the three different models

Series	dt_{meas} (day)	dt_{mod} (day)	$var(n_{c,t})$ (cm^2)	$var(n_{v,t})$ (cm^2)	$var(\xi_t)$ (cm^2)	G $(cm(mmday^{-1})^{-1}$	s_G $(cm(mmday^{-1})^{-1})$	R_T^2
1	14	14	152.7	118.2	180.5	26.29	1.8	0.83
2	14	1	120.7	116.7	142.0	26.74	1.4	0.87
3	mixed	1	89.2	114.8	136.7	26.50	1.4	0.84

6. Discussion and conclusions

The objective of this paper was to determine the influence of a reduction of the modeling interval on the performance (i.e. 'fit' and accuracy of estimated transfer function) of the state–space time series model. The fit can only be measured if the real transfer function is known. For this reason a large range of groundwater time series were generated. Calculations on several samples (with varying measuring and modeling intervals) of different time series show that the performance of transfer models can be increased by simply reducing the modeling interval. The degree of model improvement depends on several aspects.

First, the modeling interval itself relative to the time of peak response of the system is important. If the modeling interval is large with respect to the time of peak response, a reduction of the interval will greatly improve the performance. On the other hand, if the modeling interval is already small with respect to the time of peak response, a further reduction will not be very effective.

Second, the relative effect of a reduction of the modeling interval will be less if the stochastic component of the system (the part of the system dynamics that is not related to the input signal) is large.

Third, the effect of reducing the modeling interval becomes larger as the length of the time series increases. This reduction is especially observed for the fit of the deterministic component.

In addition to reducing the modeling interval, a time series can be extended with easily obtainable high-frequency measurements (i.e. a reduction of the *measuring* interval). The effect of such an extra set of high-frequency measurements again strongly depends on the stochastic component: high-frequency measurements are much more effective if the stochastic component is large. Moreover, the first high-frequency measurements have the greatest influence on the model performance. It is therefore attractive to add a small time period of high-frequency measurements to the existing time series of low-frequency measurements.

The state–space time series model has been used to model a real-world test case of which the measuring interval was reduced to one day for the last three months. The results of this case study support the results of the generated time series: a reduction of the modeling interval significantly improves the performance of the time series model. Due to the short period of high-frequency measurements in the test case (only three months), the improvement resulting from these high-frequency measurements was relatively moderate.

Further improvement of groundwater time series models can be achieved by including physical knowledge. The state–space approach allows for extension of the model with unobserved states. Current research focuses on incorporation of unsaturated processes in the time series model. This physically-based model has already been tested on time series of relatively deep groundwater levels and results show that this model gives much better predictions than the linear TFN model. Besides, the physical basis of the model parameters enables comparison of the calibrated parameters with physical knowledge of the system.

Acknowledgements

The authors wish to thank M. Knotters and an anonymous reviewer for their comments.

References

Bierkens, M.F.P., Knotters, M., Van Geer, F.C., 1999. Calibration of transfer function–noise models to sparsely or irregularly observed time series. Water Resour. Res. 35 (6), 1741–1750.

Bierkens, M.F.P., Knotters, M., Hoogland, T., 2001. Space–time modeling of water table depth using a regionalized time series model and the Kalman filter. Water Resour. Res. 37 (5), 1277–1290.

Box, G.E.P., Jenkins, G.M., 1970. Time Series Analysis, Forecasting and Control, Holden-Day, San Fransisco.

Gehrels, J.C., van Geer, F.C., de Vries, J.J., 1994. Decomposition of groundwater level fluctuations using transfer modelling in an area with shallow to deep unsaturated zones. J. Hydrol. 157, 105–138.

Gill, P., Murray, W., Wright, M.H., 1981. Practical Optimization, Sixth ed., Academic Press, New York.

Harvey, A.C., 1990. Forecasting, Structural Time Series Models and the Kalman Filter, Cambridge University Press, Cambridge.

Hipel, K.W., McLeod, A.I., 1994. Time series modelling of water resources and environmental systems, Developments in water science, vol. 45. Elsevier, New York.

Janacek, G., Swift, L., 1993. Time Series Forecasting, Simulation, Applications, Ellis Horwood, London.

Knotters, M., Bierkens, M.F.P., 2000. Physical basis of time series models for water table depths. Water Resour. Res. 36 (1), 181–188.

Maybeck, P.S., 1979. Stochastic Models, Estimation and Control, vol. 1, 141-1 in Mathematics in science and engineering. Ac. Press, New York.

O'Connell, P.E., 1980. Real-time hydrological forecasting and control, Proceedings of First International Workshop, July (1977), Institute of Hydrology, Oxon.

Schweppe, F.C., 1973. Uncertain Dynamic Systems, Prentice-Hall, New Jersey.

Shumway, R.H., Stoffer, D.S., 1982. An approach to time series smoothing and forecasting using the EM algorithm. J. Time Ser. Anal. 3 (4), 253–264.

Van Geer, F.C., Defize, P.R., 1987. Detection of natural and artificial causes of groundwater fluctuations. The influence of climate change and climatic variability on the hydrologic regime and water resources. IAHS Publ. No. 168, pp. 597–606

Van Geer, F.C., Zuur, A.F., 1997. An extension of Box–Jenkins transfer/noise models for spatial interpolation of groundwater head series. J. Hydrol. 192, 65–80.

Young, P.C., 1994. Time-variable parameter and trend estimation in non-stationary economic time series. J. Forecast. 13, 179–210.

Young, P.C., Jakeman, A.J., Post, D.A., 1997. Recent advances in the data-based modelling and analysis of hydrological systems. Water Sci. Technol. 36 (5), 99–116.