JOURNAL OF HYDROMETEOROLOGY

VOLUME 7

Improvement of Modeled Soil Wetness Conditions and Turbulent Fluxes through the Assimilation of Observed Discharge

VALENTIJN R. N. PAUWELS AND GABRIËLLE J. M. DE LANNOY

Laboratory of Hydrology and Water Management, Ghent University, Ghent, Belgium

(Manuscript received 23 February 2005, in final form 8 August 2005)

ABSTRACT

The objective of this paper is to improve the performance of a hydrologic model through the assimilation of observed discharge. Since an observation of discharge at a certain time is always influenced by the catchment wetness conditions and meteorology in the past, the assimilation method will have to modify both the past and present soil wetness conditions. For this purpose, a bias-corrected retrospective ensemble Kalman filter has been used as the assimilation algorithm. The assimilation methodology takes into account bias in the forecast state variables for the calculation of the optimal estimates. A set of twin experiments has been developed, in which it is attempted to correct the model results obtained with erroneous initial conditions and strongly over- and underestimated precipitation data. The results suggest that the assimilation of observed discharge can correct for erroneous model initial conditions. When the precipitation used to force the model is underestimated, the assimilation of observed discharge can reduce the bias in the modeled turbulent fluxes by approximately 50%. This is due to a correction of the modeled soil moisture. In the case of an overestimation of the precipitation, an improvement in the modeled wetness conditions is also obtained after data assimilation, but this does not lead to a significant improvement in the modeled energy balance. The results in this paper indicate that there is potential to improve the estimation of hydrologic states and fluxes through the assimilation of observed discharge data.

---

1. Introduction

The study of the relationship between precipitation and catchment discharge has always been one of the major interests in hydrology. Models describing this relationship have been developed conceptually, meaning that the complex physical reality is simplified using specific hypotheses and assumptions (Beven 2000). The drawback of such models is that their range of applicability is limited to the conditions under which they have been developed. To better model the rainfall-runoff behavior of a catchment, physically based models have been developed. More specifically, the goal of this type of model is to describe the partitioning of the incoming solar and atmospheric radiation into latent, sensible, and ground heat fluxes, and the partitioning of the precipitation into surface runoff, infiltration, and evapotranspiration. The study of these processes has led to

the development of soil–vegetation–atmosphere transfer schemes (SVATS). A general overview of the overall performance of a number of these models can be found in Wood et al. (1998).

It is well known that results from SVATS are prone to errors due to a variety of reasons, which can be errors or oversimplifications in the formulation of the model physics, and/or errors in the meteorological forcing data. Other reasons include the lack of soil, vegetation, and topographic data at a sufficiently high resolution and/or errors in these datasets and the derived parameters. Remote sensing of surface state variables (e.g., soil temperature, soil moisture) or surface fluxes (e.g., latent heat fluxes) forms an excellent opportunity to update SVATS with external datasets (Troch et al. 2003). The update of the model state with externally measured variables is commonly referred to as data assimilation.

A wide variety of studies have put data assimilation into practice. The most frequently used methods to update the soil moisture state of land surface models are direct insertion (Heathman et al. 2003), Newtonian nudging (Houser et al. 1998; Pauwels et al. 2001; Pani-

---

Corresponding author address: Valentijn R. N. Pauwels, Laboratory of Hydrology and Water Management, Ghent University, Coupure links 653, B-9000 Ghent, Belgium.
E-mail: Valentijn.Pauwels@UGent.be

© 2006 American Meteorological Society

coni et al. 2002), optimal interpolation (Seuffert et al. 2004), Kalman filtering (Galantowicz et al. 1999; Hoeben and Troch 2000; Walker and Houser 2001; Walker et al. 2001; Reichle et al. 2002a,b; Margulis et al. 2002; Crow and Wood 2003), Kalman smoothing (Dunne and Entekhabi 2005), and variational data assimilation (Castelli et al. 1999; Reichle et al. 2001a,b; Boni et al. 2001; Caparrini et al. 2003, 2004; Margulis and Entekhabi 2003; Crow and Kustas 2005). Several of these studies have used remotely sensed data, including radar and radiometric data. Recently, remotely sensed latent heat fluxes have been assimilated into a hydrologic model in order to update the soil moisture state (Schuurmans et al. 2003). Other studies have focused on the assimilation of surface temperature (Boni et al. 2001; Lakshmi 2000), or have assimilated both the soil moisture content and surface temperature (Walker et al. 2001).

The application of hydrologic models at large spatial scales renders the use of in situ observations difficult for the update of the model state. In this case remote sensing data are better fitted for the observation of the system state. Although it has been shown that the assimilation of remotely sensed data can lead to an overall increase in model performance, the use of remotely sensed data suffers from some significant drawbacks. Remote sensing data usually have a larger uncertainty than in situ observations and are usually available at a coarse temporal (e.g., daily, weekly, or monthly products) and spatial (10 m to 50 km) resolution. Because of these drawbacks, an alternative to using remote sensing observations is the use of catchment discharge observations in order to update the model soil wetness state at the catchment scale. Discharge observations represent the integrated effect of point-scale runoff-generating processes and catchment topography, and are an excellent indication of the catchment wetness conditions. The major advantage of the use of discharge observations for the updating of hydrologic models is that these observations are generally available at a high temporal resolution. Further, catchment discharge is usually monitored with a reasonable degree of confidence (Boiten 2000). Discharge measurements are also available across different spatial scales, in contradiction to remotely sensed observations of, for example, soil moisture, and they can thus be used for the improvement of model results across different spatial scales. The use of measured river discharge values for the updating of the state of hydrologic models through Kalman filtering has been given a good deal of attention in the past. Kalman filtering and observed discharge data have been used to update the state variables and/or parameters of simple linear (mostly time series) (Todini

and Wallis 1978; Wood 1978; Grigoriu 1978; Katayama 1980; Bras 1980; Ganendra 1980) and nonlinear rainfall-runoff models (Logan et al. 1978; Katayama 1980; Bras 1980). Recently, the possibility to improve the forecasts of simple conceptual rainfall-runoff models has been indicated by Aubert et al. (2003). In their study, the results from a simple lumped rainfall-runoff model, consisting of two reservoirs (a soil and routing reservoir) for the entire catchment, and requiring only four input parameters, have been updated through the assimilation of observed soil moisture values and/or discharge records. The objective of all these studies was to improve runoff forecasts, not the modeled soil moisture state or energy balance terms.

The overall goal of this paper is to improve the performance of a fully process-based lumped water and energy balance model [the TOPMODEL-Based Land–Atmosphere Transfer Scheme (TOPLATS; Famiglietti and Wood 1994)] through the assimilation of observed discharge values. More specifically, the first objective of this study is to correct the modeled soil wetness conditions and turbulent fluxes if the initial conditions are wrong. Further, it is known that precipitation is the most important variable in the determination of the soil wetness conditions, which in turn are very important in the calculation of evapotranspiration rates. For this reason, the second objective of this paper is to assess whether the assimilation of observed discharge can correct modeled soil moisture values, and consequently improve the modeled energy balance, when the precipitation used to force the model is either over- or underestimated. All model applications are done over the Zwalm catchment in Belgium, using hourly observed forcing data from 1994 through 1998.

The paper is organized as follows. First, the assimilation algorithm is explained. Then, a short description of the hydrologic model, runoff routing algorithm, test site, and datasets is given. After this the results of the various model applications are analyzed. Finally, a summary of the obtained results and conclusions that can be drawn from this work are given.

2. The assimilation algorithm

In the ensemble Kalman filter (EnKF) (Evensen 1994), the following system is considered:

$$\begin{cases} \mathbf{X}_{k+1} = f_k(\mathbf{X}_k, \mathbf{u}_k) + \mathbf{w}_k \\ \mathbf{Y}_k = \mathbf{H}_k \mathbf{X}_k + \mathbf{v}_k \end{cases}, \qquad (1)$$

where $\mathbf{X}_k$ is the system state vector with n_s entries, $\mathbf{Y}_k$ is the system observation vector with n_o entries, $\mathbf{u}_k$ is the system forcing, $\mathbf{w}_k$ is the process noise, $\mathbf{v}_k$ is the

measurement noise, and k indicates the time step. In our case, $\mathbf{X}_k$ consists of the soil moisture content (of the upper and lower soil layers) and catchment-averaged water table level, and $\mathbf{Y}_k$ consists of the measured discharge; $\mathbf{H}_k$ relates the system state to the observation of this system state (in our case it thus relates the soil water content and water table level to the discharge) and is calculated numerically. The statistics of $\mathbf{w}_k$ and v_k can be summarized as follows:

$$
\begin{cases}
E(\mathbf{w}_k) = 0, & E(\boldsymbol{v}_k) = 0 \\
E(\mathbf{w}_k\mathbf{w}_k^{\mathrm{T}}) = Q_k \geq 0, & E(\boldsymbol{v}_k\boldsymbol{v}_k^{\mathrm{T}}) = R_k \geq 0 \\
E(\mathbf{w}_k\mathbf{w}_l^{\mathrm{T}}) = 0, & E(\boldsymbol{v}_k\boldsymbol{v}_l^{\mathrm{T}}) = 0, & \forall k \neq l \\
E(\mathbf{w}_k\boldsymbol{v}_l^{\mathrm{T}}) = 0, & & \forall k, l
\end{cases}
$$

(2)

where Q_k is the process noise covariance, R_k is the measurement noise covariance, and T stands for the transpose operator. Given the above definitions, Eq. (1) is an accurate description of the system for the purpose of assimilating data that are, at each measurement time step, determined by the system state at the same time step. However, runoff generated at a certain time and location in a catchment will only reach the catchment outlet after a certain time period. For this reason, a runoff observation at a certain time is determined by the state of a catchment at a number of previous time steps. Thus, if runoff time series are assimilated into a hydrologic model, not only the state of the catchment for a given location at the current time step, but also the state at a number of different locations and previous time steps, have to be updated and propagated forward to the current time step. For this reason, the system state vector $\mathbf{X}_k$ is augmented so that it contains not only the system state at time step k, but also the system state at time steps $k - 1, k - 2, \ldots, k - (n_c + n_y - 1)$. Here n_c is the number of time steps in the concentration time of the catchment (the number of time steps it takes for surface runoff generated at the point in the catchment farthest from the outlet to reach the outlet), and n_y is the number of time steps with observed discharge values in the observation vector. Using this representation, an observation at time step k will thus be used to update the model state at time steps k through $k - n_c$. If observations at time step k and $k - 1$ are used, the model state at time steps k through $k - n_c - 1$ will be updated. These model states at time steps $k - i$, with i ranging from zero to n_c, are in fact the initial conditions for time step $k - i + 1$. Using these initial conditions, the model is then reapplied for time steps k through $k - n_c + 1$, from which the modeled fluxes are calculated. The approach of distributing the system state
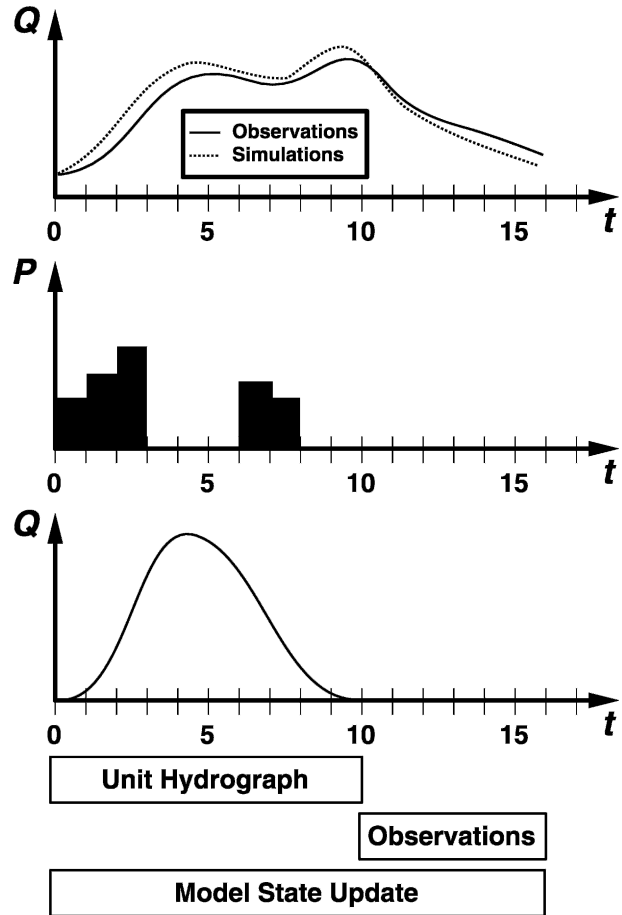

FIG. 1. Schematic of the data assimilation algorithm. (top) The comparison between the discharge simulations and observations. (middle) The precipitation time series that were used to obtain the discharge simulations. (bottom) The unit hydrograph of the catchment.

vector in time, and the use of observed data to update the model state at a number of time steps prior to the observation time steps, using the traditional EnKF equations, is referred to as retrospective ensemble Kalman filtering (REnKF).

Figure 1 explains this system representation in more detail. Assume a catchment has a unit hydrograph that lasts for 10 time steps. If one wants to assimilate a runoff observation at a certain time step, the system state up to 10 time steps before the runoff observation has to be updated. If, for example, one wants to assimilate six consecutive runoff observations (this could be any number; in this example the number six was randomly chosen), the system state starting 10 time steps before the first observation, up to the state at the time of the last observation, thus in total 16 time steps, has to be updated. The size of the assimilation window (the number of observations taken into account) can be cho-

sen based on computational efficiency, or an optimal value can be determined based on a comparison between the results of the assimilation procedure and observations. The objective of this paper is to examine the data assimilation algorithm with an assimilation window of one observation. An examination of the optimal size of the assimilation window is the purpose of a further study.

A number of other methods are available to update state variables at a certain time step using observations from different time steps and that potentially could be used for the assimilation of runoff observations, by making the observations $\mathbf{Y}$ at time step k in Eq. (1) dependent not only on the state $\mathbf{X}$ at time step k, but also at time steps $k - 1$, $k - 2$, and so on. For example, Cohn et al. (1994) developed a Kalman smoothing approach, in which the system state vector is not distributed in time, but in which the cross covariances between the errors at the current and the previous time steps have to be propagated in order to allow the calculation of the Kalman gain factors. In this method, a different gain factor is calculated for every time step at which the model state variables have to be updated. Todling and Cohn (1998) developed and analyzed a number of simplifications to this fixed-lag Kalman smoother. Another method to update the system state at a certain time before the observation was developed by Evensen and van Leeuwen (2000), in which a first guess for the analysis is the result of the application of the ensemble Kalman filter. This first guess is then propagated backward in time by using the ensemble error cross covariances. In other words, every time a new dataset becomes available during the forward integration, an analysis is first computed for that time step using the ensemble Kalman filter solutions. Since the covariances between the states at the observation time steps and the states at previous time steps are known, the update at the observation time step can be propagated backward in time for all time steps up to that time. The approach of distributing the system state vector in time was recently used by Dunne and Entekhabi (2005) for the assimilation of surface soil moisture in a reanalysis approach. In their study, soil moisture observations with a temporal resolution of 3 days,

and an assimilation window of three observations, were used. They concluded that for the purpose of assimilating soil moisture data their method can yield an improved state estimate as compared to traditional filtering. Because of these promising results, and the relative simplicity of implementation, the approach of distributing the system state vector in time is used in this study.

An additional difficulty in the assimilation of discharge records in this experiment is the representation of forecast bias. If the meteorological data used to force the model are biased, or if certain model parameters are wrong (e.g., the hydraulic conductivity), the modeled soil wetness, and consequently the modeled energy balance, will also be biased. If this forecast bias is not adequately represented in the assimilation algorithm, the resulting state estimates will not be optimal. In this paper, we use the online forecast bias estimation and correction with feedback approach described in Dee and Da Silva (1998). In this approach, forecast bias estimate and the state estimate are modeled separately, and observations of the system state are used to optimize both estimates.

The data assimilation algorithm works as follows. To apply the REnKF, an ensemble of N system state vectors has to be propagated in parallel. Each system state vector represents one single realization of the possible model trajectories and bias.

First, for each ensemble member i, the state and bias estimates are propagated:

$$\begin{cases} \hat{\mathbf{X}}_k^{i-} = f_{k-1}(\tilde{\mathbf{X}}_{k-1}^{i+}, \mathbf{u}_{k-1}^i) \\ \hat{\mathbf{B}}_k^{i-} = \hat{\mathbf{B}}_{k-1}^{i+} \\ \tilde{\mathbf{X}}_k^{i-} = \hat{\mathbf{X}}_k^{i-} - \hat{\mathbf{B}}_k^{i-} \end{cases}, \quad (3)$$

where $\mathbf{B}_k$ is the forecast bias, defined as the average modeled state variable minus the truth. The hat indicates an estimate of the variable, the tilde an unbiased state estimate, the superscript $^-$ the a priori estimate (forecast, before the assimilation of the observations), and the superscript $^+$ the a posteriori estimate (analysis, after the assimilation). The a priori system state vector $\tilde{\mathbf{X}}_k^{i-}$ of each ensemble member i is used to calculate the a priori state error covariance as follows:

$$\begin{cases} \mathbf{P}_k^- = \dfrac{1}{N-1} \mathbf{D}_k^- \mathbf{D}_k^{-\mathrm{T}} \\ \mathbf{D}_k^- = [\tilde{\mathbf{X}}_k^{1-} - \tilde{\mathbf{X}}_k^-, \quad \cdots, \quad \tilde{\mathbf{X}}_k^{i-} - \tilde{\mathbf{X}}_k^-, \quad \cdots, \quad \tilde{\mathbf{X}}_k^{N-} - \tilde{\mathbf{X}}_k^-], \\ \tilde{\mathbf{X}}_k^- = \dfrac{1}{N} \sum_{i=1}^{N} \tilde{\mathbf{X}}_k^{i-} \end{cases} \quad (4)$$

where $\mathbf{P}_k^-$ is the a priori error covariance matrix. The assimilation algorithm first updates the bias estimate, after which the state estimate is updated. As explained by Dee and Da Silva (1998), the forecast error covariance $\mathbf{S}_k^-$, and the bias prediction error covariance $\mathbf{S}_k^{b-}$, can be calculated as

$$\begin{cases} \mathbf{S}_k^- = (1 - \gamma)\mathbf{P}_k^- \\ \mathbf{S}_k^{b-} = \gamma\mathbf{P}_k^- \end{cases}, \tag{5}$$

where γ is estimated by comparing the root-mean-square error (rmse) between the modeled and observed state variables to the rmse between the modeled states variables with removal of the bias and the observations. Division of the second by the first rmse will provide a rough estimate of γ. This is further explained in section 6d. The Kalman gain factor for the bias estimates is then calculated:

$$\mathbf{L}_k = \mathbf{S}_k^{b-} \mathbf{H}_k^{\mathrm{T}}(\mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^{\mathrm{T}} + \mathbf{R}_k)^{-1}. \tag{6}$$

The estimate of the bias is updated as follows:

$$\hat{\mathbf{B}}_k^{i+} = \hat{\mathbf{B}}_k^{i-} - \mathbf{L}_k(\mathbf{Y}_k - \mathbf{H}_k \tilde{\mathbf{X}}_k^{i-} + \boldsymbol{v}_k^i), \tag{7}$$

where $\boldsymbol{v}_k^i$ is a random realization of the measurement error (Burgers et al. 1998). The a priori error covariance is then used to calculate the Kalman gain factor:

$$\mathbf{K}_k = \mathbf{S}_k^- \mathbf{H}_k^{\mathrm{T}}(\mathbf{H}_k \mathbf{S}_k^- \mathbf{H}_k^{\mathrm{T}} + \mathbf{R}_k)^{-1}. \tag{8}$$

The a posteriori state vector of each ensemble member is then calculated as follows:

$$\tilde{\mathbf{X}}_k^{i+} = \tilde{\mathbf{X}}_k^{i-} - \hat{\mathbf{B}}_k^{i+} + \mathbf{K}_k[\mathbf{Y}_k - \mathbf{H}_k(\tilde{\mathbf{X}}_k^{i-} - \hat{\mathbf{B}}_k^{i+}) + \boldsymbol{v}_k^i]. \tag{9}$$

If observations are not available, no update is performed. The system state at each time step is simply the average of the a posteriori system states of all ensemble members.

3. Model description

TOPLATS has as its foundation the concept that shallow groundwater gradients set up spatial patterns of soil moisture that influence infiltration and runoff during storm events, and evaporation and drainage between storm events. The assumption is made that these gradients can be estimated from local topography [through a soil–topographic index (Sivapalan et al. 1987)]. From this foundation, the model was expanded to include infiltration and resistance-based evaporation processes, a surface vegetation layer and a surface energy balance equation with an improved ground heat flux parameterization, and the effect of atmospheric stability on energy fluxes (Famiglietti and Wood 1994;

Peters-Lidard et al. 1997). The model was originally developed to simulate the surface water and energy balance for warm seasons (Famiglietti and Wood 1994; Peters-Lidard et al. 1997). More recently, winter processes (frozen ground and a snowpack), an improved water and energy balance scheme for open-water bodies, and a two-layer vegetation parameterization were added (Pauwels and Wood 1999a). Application to the Zwalm catchment (Pauwels et al. 2001, 2002) and to field experiments such as the First International Satellite Land-Surface Climatology Project (ISLSCP) Field Experiment (FIFE) (Peters-Lidard et al. 1997) and the Boreal Ecosystem Atmosphere Study (BOREAS) (Pauwels and Wood 1999b, 2000) has shown that the model can adequately simulate surface energy fluxes, soil temperatures, and soil moisture.

The soil column in divided into a thin upper layer (depth 5 cm) and a lower layer. Water for bare-soil evaporation is supplied by the thin upper layer, while water for vegetation transpiration can be supplied by either layer, depending on the distribution of the roots in the soil column. This distribution is specified for each land-cover class as a model parameter. Surface runoff is generated using both the saturation and infiltration excess mechanisms. When the rate of net precipitation exceeds the infiltration capacity of the soil [which depends on the soil moisture content (Philip 1957)], infiltration excess runoff is generated. When the capillary layer above the groundwater table reaches the surface, the net precipitation cannot infiltrate and is considered to be saturation excess runoff. Baseflow is calculated using the TOPMODEL approach (Beven and Kirkby 1979), which states that the baseflow varies exponentially with the depth to the water table. The total catchment discharge is then the sum of the baseflow and the saturation and infiltration excess runoff. For a more complete description of the model physics at the point scale we refer to Pauwels et al. (2002).

Horizontal transport of water is implicitly modeled through the TOPMODEL formulation. At the end of each time step the sum of the fluxes to and from the water table are calculated, after which the new average water table depth over the catchment is calculated. This catchment-average water table depth is then used to calculate the local water table depth using TOPMODEL (Sivapalan et al. 1987).

While the model can be run in two different modes (distributed and statistical), the statistical mode is used in this study. The statistical mode has been developed under the similarity concept that locations within the domain (or macroscale grid), with the same soil–topographic index and vegetation, respond similarly. Using the statistical distributions of these controlling

factors, rather than applying the model for each grid, considerably reduces the computational effort.

4. The runoff routing algorithm

The runoff routing algorithm was developed by Troch et al. (1994). The algorithm is based on a solution of the de Saint Venant equations, neglecting the acceleration terms. When the upstream boundary condition is given by the Dirac function, Brutsaert (1973) solved these equations as follows:

$$q(x, t) = \frac{x}{\sqrt{2\pi}bt^{3/2}} \exp\left[-\frac{(x - at)^2}{2b^2t}\right], \quad (10)$$

where q is the discharge per unit width of channel (s^{-1}), x is the distance along the channel (m), t is the time (s), a is the drift velocity $(m\,s^{-1})$, and b^2 is the diffusion coefficient $(m^2\,s^{-1})$; a and b are calculated as

$$\begin{cases} a = (1 + a_0)V \\ b^2 = \dfrac{V^3}{gS_0F^2}(1 - a_0^2F^2), \end{cases} \quad (11)$$

where V is the vertically averaged channel velocity $(m\,s^{-1})$, S_0 is the channel bed slope $(-)$, F is the dimensionless Froude number, g is the gravity constant $(9.81\,m\,s^{-1})$, and a_0 is equal to 2/3.

Equation (10) describes the response in the channel at a distance x from the origin of the instantaneous unit perturbation. Using the normalized width function, the response of the entire channel network can be expressed as (Mesa and Mifflin 1986)

$$f_c(t) = \int_0^\infty q_c(x, t)W(x)\,dx. \quad (12)$$

The normalized width function $W(x)$ (m^{-1}) represents the distribution of runoff entering the network at a distance x from the outlet and can be written as

$$W(x) = \frac{1}{L_T}N_c(x), \quad (13)$$

where $N_c(x)$ is the number of channel links at a given distance from the outlet and L_T is the total channel length (m).

Troch et al. (1994) expanded this method to calculate the channel flow to calculate the overland flow. Instead of calculating the distribution of runoff entering the network at a distance x, the distribution of runoff generated at a distance x from the channel network is calculated. This distribution is represented by the normalized hillslope function $H(x)$ (m^{-1}):

$$H(x) = \frac{I\epsilon\overline{N}_x}{P\epsilon^2}, \quad (14)$$

where ϵ is the resolution at which the normalized hillslope function is calculated (m), I is the number of cells in the channel, and P is the number of cells in the hillslope; $\overline{N}_x$ is calculated as

$$\overline{N}_x = \frac{1}{I}\sum_{i=1}^{I} N_x(i), \quad (15)$$

where $N_x(i)$ is the number of cells that drain directly to cell i at an overland flow distance x. Using Eq. (10) with adapted parameters a and b for the overland flow, and Eq. (12) with $H(x)$ instead of $W(x)$, the response of the hillslopes $f_h(t)$ (s^{-1}) to an instantaneous unit input of water can be calculated.

The response of the catchment to an instantaneous unit input of water $[f_b(t), s^{-1}]$ can then be calculated through convolution:

$$f_b(t)\int_0^\infty f_h(t - \tau)f_c(\tau)\,d\tau. \quad (16)$$

The channel and overland parameters were optimized through a study of the time lag between the simulated and the observed discharge. Figure 2 shows the normalized hillslope and width functions, and the resulting unit hydrograph for the Zwalm catchment.

5. Site and data description

Figure 3 shows the location of the Zwalm catchment. For a complete description of the Zwalm catchment we refer to Troch et al. (1993) as only a short overview will be given here. The total drainage area of the catchment is 114 km² and the total length of the perennial channels is 177 km. The maximum elevation difference is 150 m. The average year temperature is 10°C, with January the coldest month (mean temperature 3°C) and July the warmest month (mean temperature 18°C). The average yearly rainfall is 775 mm and is distributed evenly throughout the year. The annual actual evaporation is approximately 450 mm.

Meteorological forcing data with an hourly resolution (the model time step) from 1994 through 1998 are used in this study. These data were prepared based on daily observations of air temperature and humidity, solar radiation, wind speed, and precipitation, from the climatological station located in Kruishoutem, Belgium, approximately 5 km outside the catchment (KMI 1994–1998). Pauwels et al. (2002) give a detailed description of the processing of the meteorological forcings. Discharge observations at the outlet of the catchment, used for the validation of the model, were recorded at an hourly time step. Finally, a 30-m-resolution digital elevation model (DEM), a soil texture map from the Bel-
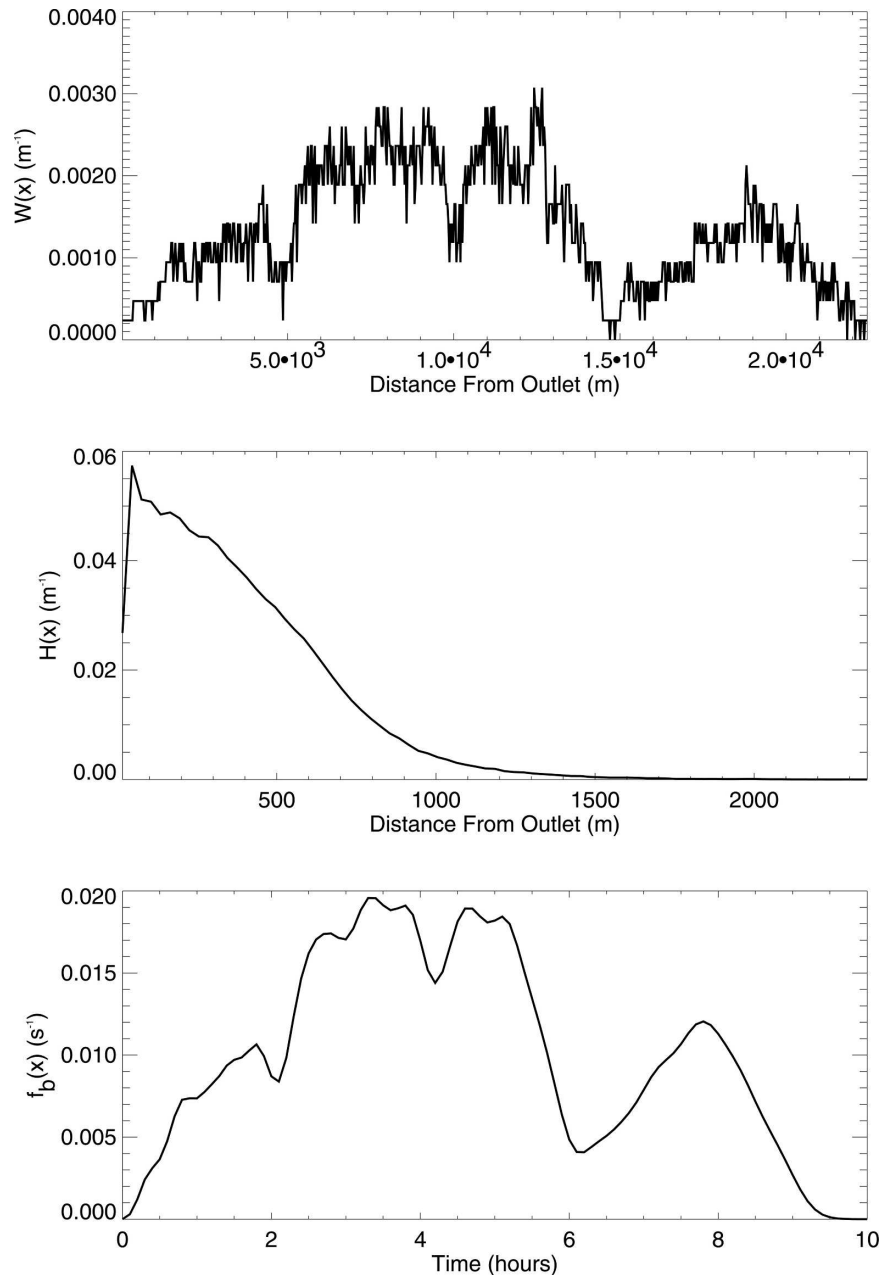
FIG. 2. The runoff routing functions for the Zwalm catchment. (top) The normalized width function. (middle) The normalized hillslope function. (bottom) The resulting unit hydrograph.

gian National Geographic Institute, and a Systeme pour l'Observation de la Terre (SPOT)-derived land cover classification map from 3 August 1998, were used to define model parameters. Soil parameters were derived using the relationship with soil texture given by Rawls et al. (1982), and the vegetation parameters were derived for each land-cover classification following Peters-Lidard et al. (1997). The catchment base flow parameters (the base flow at saturation and the paramater of the exponential decay in the saturated hydraulic con-

ductivity with depth) were taken from Troch et al. (1993).

6. Results of the model applications

a. Validation of the model

A model run is first established, in which the modeled discharge is compared to the observations. The model is run from 1994 through 1998, at an hourly time step. The distribution of the topographic indices con-
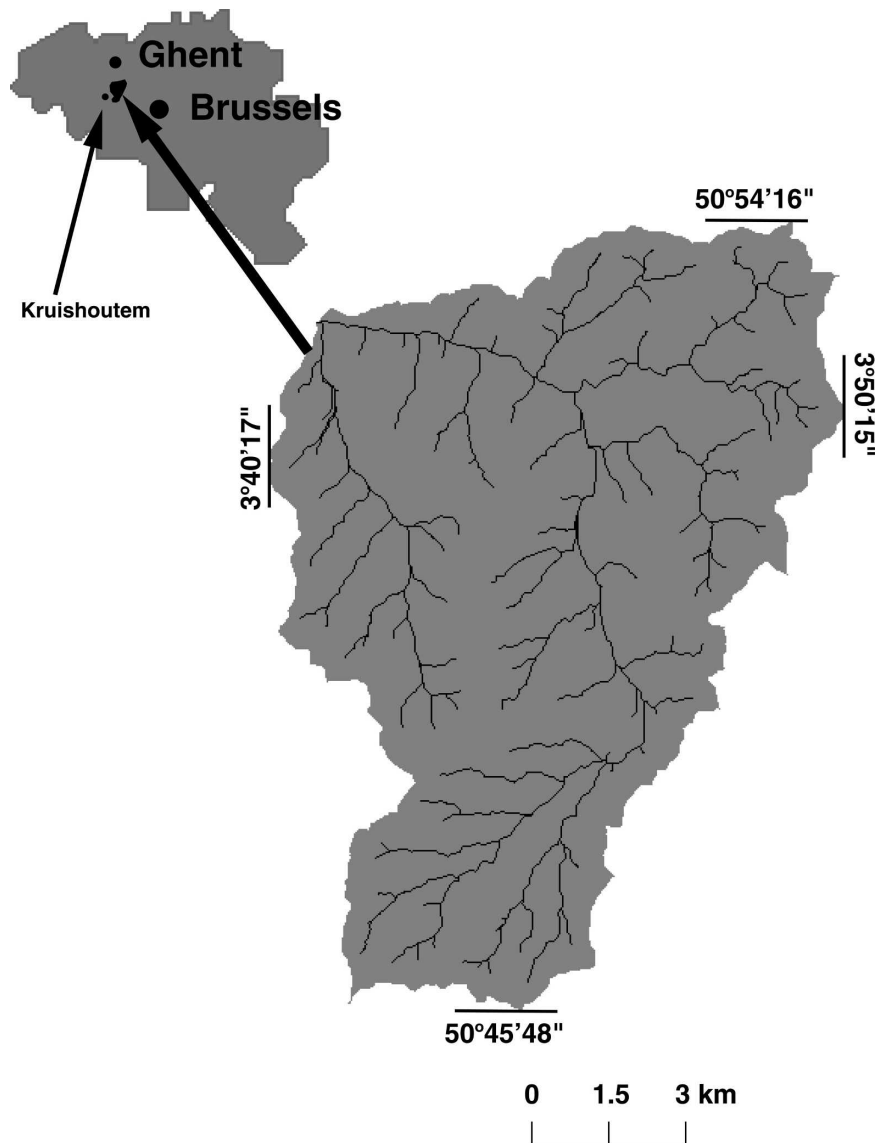
FIG. 3. The location of the Zwalm catchment.

sists of 13 intervals. Figure 4 shows the comparison between the measured and simulated discharge for the entire simulation period. The average modeled discharge is 1.47 m^3 s^{-1} as compared to 1.52 m^3 s^{-1} for the observations. The correlation coefficient between the observations and the simulations is 0.80. These statistics, together with the comparison in Fig. 4, indicate that the model is capable of representing the rainfall-runoff dynamics of the Zwalm catchment.

b. Description of the twin experiments

In the validation of the assimilation algorithm observations of the energy and water balance variables are needed. Although some observations of, for example, soil moisture content, evapotranspiration, and water

table levels were made during the study period, these observations have not been done with a sufficient frequency and spatial resolution to be useful for a 5-yr model validation. For this reason a twin experiment is developed, in which the routed discharge from the model run described in section 6a is used as observation of the catchment discharge. For the remainder of this paper, we refer to this discharge as the synthetically observed discharge. The modeled energy and water balance terms from this model application are used as the observations of these terms, and are for the remainder of this paper referred to as synthetical data.

The ensemble members are generated following the approach of Reichle et al. (2002a). The meteorological forcings are perturbed by adding Gaussian white noise
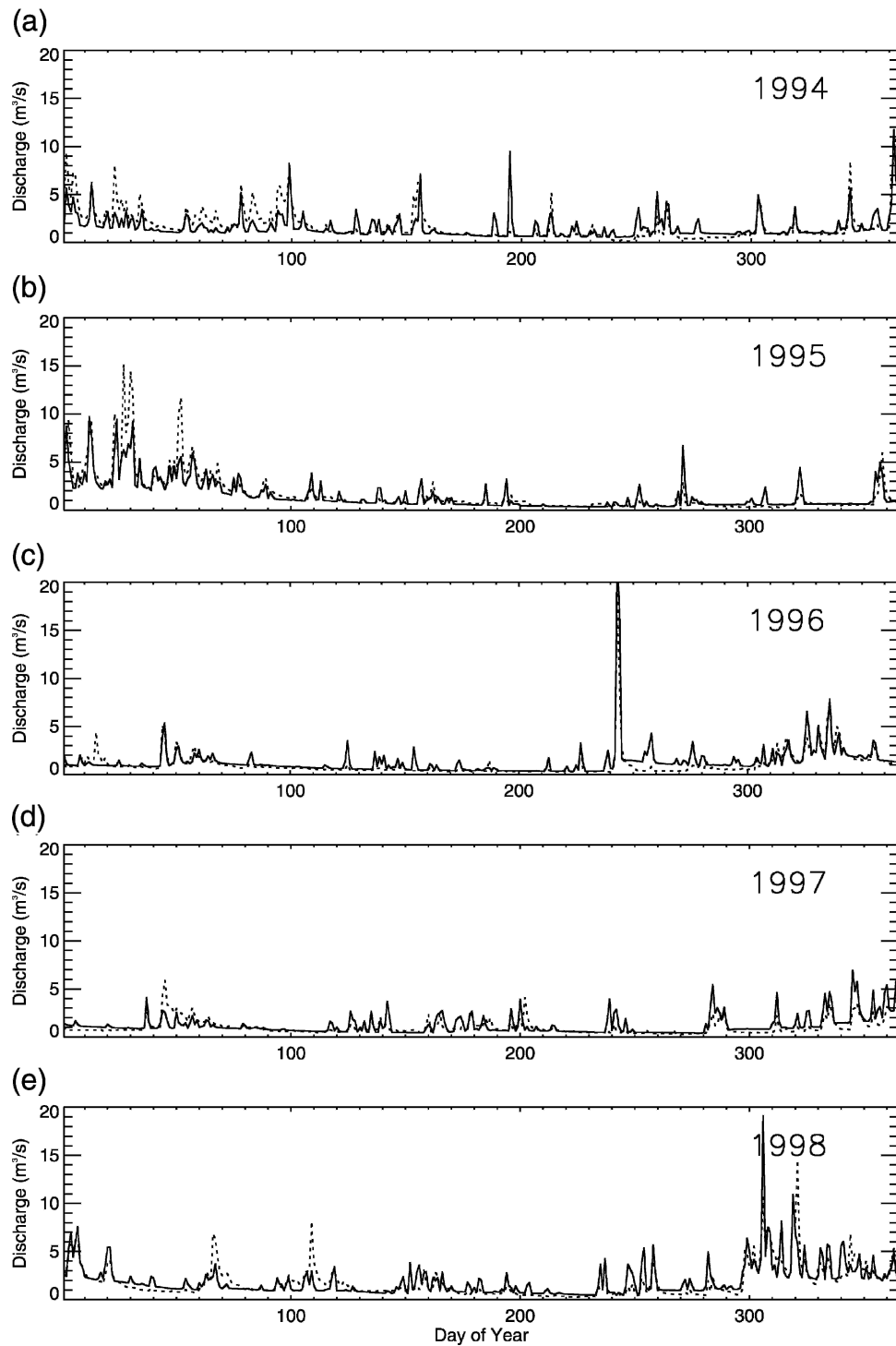
Fig. 4. Comparison between the modeled and simulated discharge for the entire study period. Observations are in solid lines and simulations are in dashed lines.

to the observed meteorological data used to force each ensemble member. The standard deviation of the noise is 5 K for the air and dewpoint temperatures, 1 m s^{-1} for the wind speed, 50 W m^{-2} for the shortwave radia-tion, 25 W m^{-2} for the longwave radiation, 10 mbar for the surface pressure, and 50% of the magnitude for the precipitation. Furthermore, the saturated hydraulic conductivity, the exponential decay parameter of the

hydraulic conductivity with depth, and the catchment-averaged base flow at saturation are also perturbed with a random number with mean zero and a standard deviation of 50% of their magnitude. The additional perturbation is based on De Lannoy et al. (2006) in which it is shown that a perturbation of only the meteorological forcings will lead to an unrealistic ensemble spread, and consequently an unrealistic error covariance. The initial conditions for the different ensemble members were not disturbed. The number of ensemble members is set to the number of entries in the system state vector, in this case 1570. This number is obtained by multiplying 6 (the number of land cover classes) by 13 (the number of intervals in the distribution of the soil–topographic index), by 2 (both the upper- and lower-layer soil water content), adding 1 (the catchment-averaged water table depth) to this number, and then multiplying this number by 10 (the number of time steps in the unit hydrograph). Only the catchment-averaged water table depth needs to be updated, since the TOPMODEL approach calculates the local water table depth using the catchment-averaged water table depth and the topographic indices. Therefore, an update of the average water table depth will be distributed to an update of the local water table depth through the TOPMODEL formulations.

Four different experiments are run. In the first two experiments, the capability of the data assimilation algorithm to correct for erroneous initial conditions is assessed. In the first experiment, the initial water table level is lowered by 0.3 m, and the synthetically observed discharge from section 6a is assimilated into the model. The impact of the assimilation on the modeled water table levels and soil moisture contents is assessed. In the second experiment, the initial water table is raised by 0.3 m, and again the synthetically observed discharge is assimilated. It should be noted that the change in the initial water table depth will automatically lead to a change in the initial soil moisture values, since these are calculated based on the assumption of equilibrium initial conditions and the water table depth.

In the third and fourth experiment, the possibility of the data assimilation algorithm to correct the model state under erroneous model forcings is assessed. For this purpose, the precipitation used to force each ensemble member is biased. In the third experiment, the observed precipitation is multiplied by a random number with mean 1.5 and standard deviation 0.5. In the fourth experiment, the mean and standard deviation of this random number are both 0.5. The impact of the data assimilation on the modeled soil water content and water table level, and the modeled surface turbulent fluxes, is assessed. In this case there are thus two per-turbations of the forcing data: the first perturbation consists of the multiplication of the observed precipitation by a random number with average 1.5 or 0.5, and the second perturbation consists of the addition of a random number to all the forcings for each ensemble member, as described in the second paragraph of this section.

The application of the filter equations requires knowledge about the statistics of the observation error, v_k. Since the objective of this study is to assess the potential improvement in hydrologic model results through the assimilation of observed discharge, as a first step the observations are assumed to be perfect, so a standard deviation of zero is used. All elements in R_k are thus zero. A sensitivity analysis is then done, in which the observation error is modified between 0.01 and 0.5 m^3 s^{-1}. The improvement in the modeled energy balance as a consequence of the assimilation of erroneous discharge data is then assessed. The measurement noise is assumed to be constant throughout the simulation.

c. Correction for erroneous initial conditions

If a hydrologic model is forced with correct meteorological data, and if the model formulations and parameters are correct, after a certain time period an error in the model initial conditions will disappear. Thus, under these conditions, one can expect the long-term model results to be unbiased. For this reason, the parameter γ and initial estimate of the bias in the assimilation algorithm are both set to zero, for the first and second twin experiments. Figure 5 shows the impact of the data assimilation on the modeled soil wetness conditions for the model applications with erroneous initial conditions. Both the initially lowered and raised water tables are corrected quickly (after approximately one day). The upper-layer soil moisture is, in both cases, also corrected after approximately one day. The modeled lower-layer soil moisture is strongly improved by the assimilation algorithm, but becomes equal to the synthetical observations only after approximately 50 days. Table 1 shows the impact of the corrected soil wetness conditions on the modeled discharge. In both cases the bias and rmse between the synthetical observations and the model results are basically eliminated. Together with the regression line closer to the 1:1 line and the increase in the correlation for the REnKF runs, this shows that, as can be expected, the assimilation procedure leads to a strong improvement in the modeled discharge. From these two twin experiments we can conclude that the assimilation of observed discharge can correct for erroneous initial conditions.
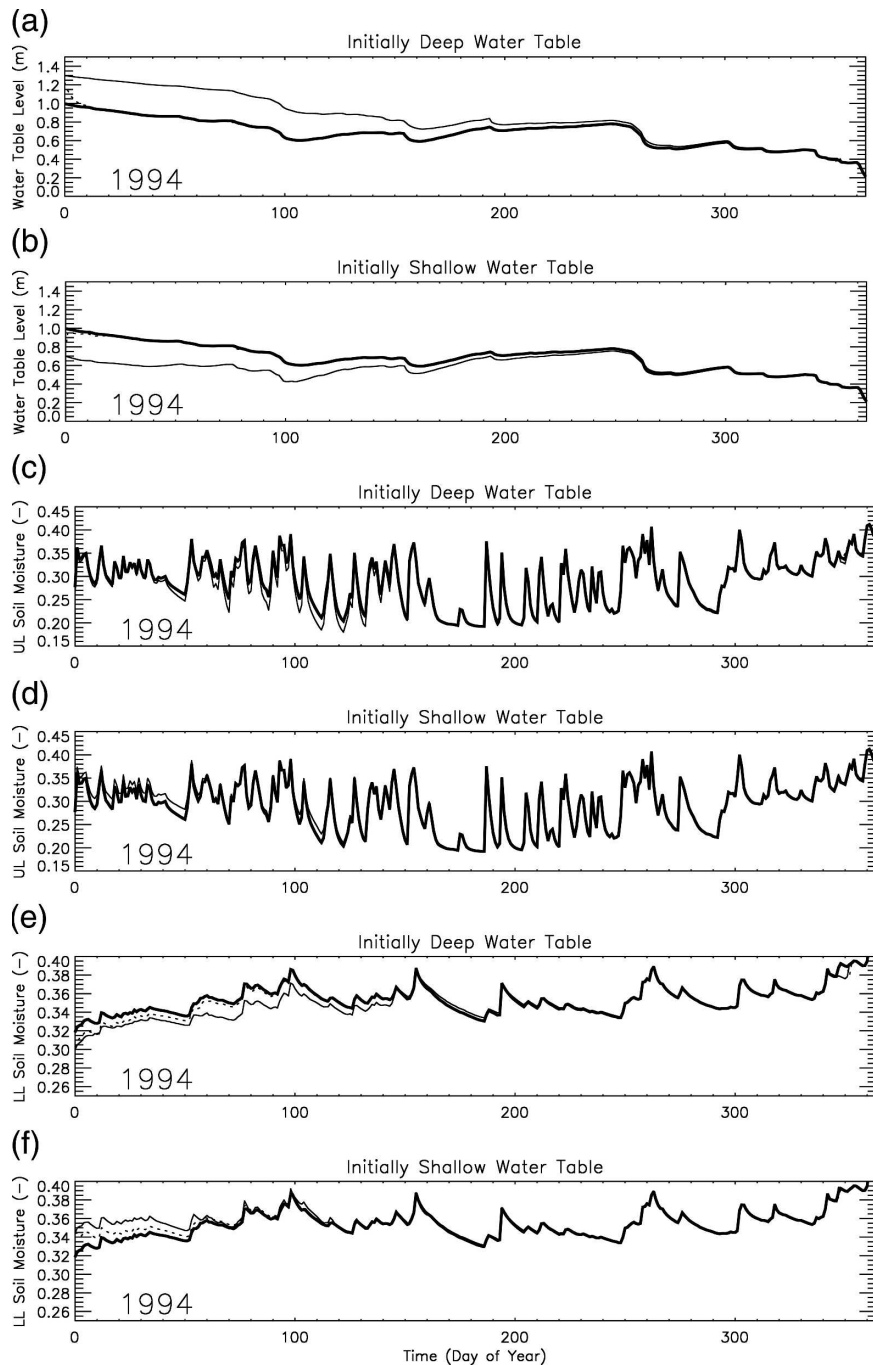
FIG. 5. Correction of erroneous initial conditions through the assimilation of synthetically observed discharge. The synthetical observations are in thick solid lines, the results of the baseline runs are in thin solid lines, and the results of the assimilation runs are in dotted lines. UL stands for upper layer and LL for lower layer. If invisible, the dotted lines coincide with the thick solid lines.

d. Correction of results under erroneous precipitation

It is undisputed that precipitation is the most important variable in the determination of discharge and soil wetness conditions, and thus also the modeled evapotranspiration. On the other hand, it is also widely known that this variable is very difficult to observe accurately over large spatial scales, because of its high intermittency and strong spatial variability. Because of

TABLE 1. Results of the data assimilation on the modeled discharge. The synthetic observations are the independent variable (x axis), and the model applications are the dependent variable (y axis). Units are in $m^3\ s^{-1}$. Application 1 is the model application with an initially too deep water table, application 2 is the model application with an initially too shallow water table, application 3 is the model application with overestimated precipitation, and application 4 is the model application with underestimated precipitation. The results for applications 1 and 2 are for 1994 only; the results for the other model applications are for 1994 through 1998.

Application	Run	$\overline{X}$	$\overline{Y}$	Slope	Intercept	Rmse	R
1	Base	1.18	0.96	0.90	−0.10	0.41	0.96
1	REnKF	1.18	1.18	1.00	−0.00	0.03	1.00
2	Base	1.18	1.37	0.97	0.22	0.29	0.98
2	REnKF	1.18	1.18	1.00	0.00	0.02	1.00
3	Base	1.53	2.55	1.91	−0.36	2.31	0.88
3	REnKF	1.53	1.68	1.30	−0.31	1.14	0.89
4	Base	1.53	0.61	0.27	0.20	1.47	0.79
4	REnKF	1.53	1.53	0.75	0.39	0.61	0.92

these two reasons, estimates of precipitation over large spatial scales will often be biased. One of the objectives of this paper is to check whether the assimilation of observed discharge can correct erroneous model results obtained under biased precipitation.

The parameter γ in Eq. (5) is calculated by dividing the rmse between the modeled and synthetically observed state variables with removal of the bias (thus subtracting the bias from the modeled state) by the rmse between the model results and synthetical observations without bias removal. The average value for the three types of state variables (water table level, and upper- and lower-layer water content) and both twin experiments is approximately 0.6. This value was found to be similar when it was calculated for 1994 only or for the entire simulation period. For the remainder of this section, a value of 0.6 for γ is used. In section 6f the sensitivity of the model results to the value for γ is examined.

Figure 6 shows the impact of the assimilation procedure on the modeled water table levels, for the model applications in which the precipitation was over- and underestimated. Results are shown for 1994 only; for the other years similar results were obtained. One can immediately see that the assimilation leads to a larger temporal variability in the catchment-averaged water table level. These peaks can be explained by the way the precipitation data are perturbed. Since the precipitation is multiplied by a random number (with mean 1.5 or 0.5), the precipitation and synthetically observed discharge will, during periods with nonzero precipitation, be inconsistent with each other. When the precipitation is overestimated, the assimilation algorithm will lower the water table, thus reducing the amount of base flow and the amount of saturated areas (thus reducing the amount of surface runoff). The modeled soil moisture will also be decreased, in order to increase the infiltration and reduce the amount of surface runoff. The lower amount of base flow and surface runoff will then

assure that the modeled discharge is equal to the synthetical observations. In case the precipitation is underestimated, the opposite will occur. The correction of the modeled discharge, however, is limited, since, in the case of underestimated precipitation, the amount of base flow the model can generate is limited to the base flow at saturation, and the amount of surface runoff is always limited to the amount of precipitation. On the other hand, if the precipitation is overestimated, the model can only increase the amount of infiltration to the maximum allowed for the soil type and lower the water table to a level where the base flow becomes insignificant.

During periods without precipitation, the assimilation algorithm will set the water table level equal to the synthetic data, correcting the base flow, and thus the total amount of discharge, since under these conditions no surface runoff occurs.

Table 1 shows that, for both model applications, the bias in the modeled discharge is almost eliminated. The rmse between the synthetically observed and modeled discharge is strongly reduced, and the correlation between the synthetically observed and modeled discharge is increased.

Figure 6 shows the impact of the assimilation procedure on the modeled upper-layer soil moisture. Again, results are shown for 1994 only given that similar results were obtained for the other years. In both cases the errors in the modeled soil moisture are basically eliminated, although the impact of the errors in the precipitation, and the consequent improvement in the modeled soil moisture, is stronger when the precipitation is underestimated. The stronger impact of an underestimation of the precipitation on the modeled soil moisture can be explained by the combined effect of evapotranspiration and infiltration. Table 2 explains this effect. Using the correct precipitation, approximately 500 mm of water infiltrates into the soil per year. When the precipitation is overestimated, the infiltration is in-
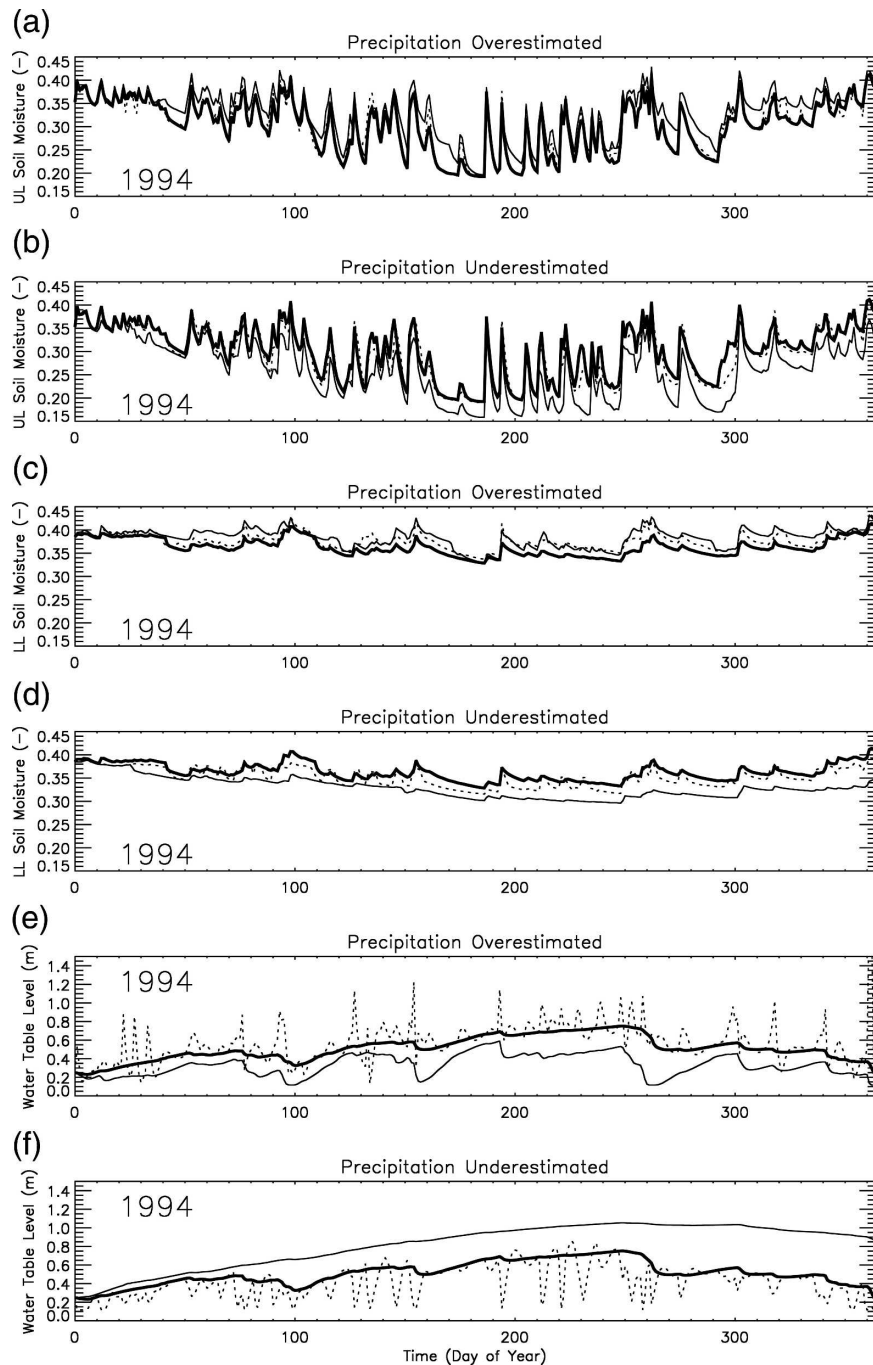
FIG. 6. Impact of the data assimilation on the modeled water balance variables in case of erroneous meteorological forcings. The synthetical observations are in thick solid lines, the results of the baseline runs are in thin solid lines, and the results of the assimilation runs are in dotted lines. UL stands for upper layer and LL for lower layer.

creased to approximately 700 mm yr^{-1}, and a reduction in the precipitation leads to a reduction of the infiltration to approximately 285 mm yr^{-1}. If the amount of soil water that is discharged as base flow is subtracted from the infiltration, approximately 215 mm yr^{-1} re-

mains available for transpiration when the precipitation is correct. When the precipitation is overestimated, approximately 300 mm yr^{-1} remains available, and when the precipitation is underestimated, approximately 150 mm yr^{-1} remains. When the precipitation is overesti-

TABLE 2. Partitioning of the precipitation for the different model applications (mm yr^{-1}).

Variable	Synthetic data	Baseline run overestimated precipitation	Assimilation run overestimated precipitation	Baseline run underestimated precipitation	Assimilation run underestimated precipitation
Precipitation	795.72	1182.47	1182.47	430.62	430.62
Infiltration	2.66	13.13	17.64	0.25	0.22
Excess runoff					
Saturation	131.29	265.63	162.38	37.71	84.05
Excess runoff					
Surface runoff	133.95	278.76	180.02	37.96	84.27
Baseflow	286.95	407.82	273.08	136.41	353.22
Total discharge	420.90	686.58	453.10	174.37	437.49
Infiltration	501.88	707.41	825.89	286.16	230.43

mated, the amount of soil water available for root up-take is thus increased by approximately 85 mm, and when the precipitation is underestimated, this amount is reduced by approximately 150 mm yr^{-1}. Since in all cases the evapotranspiration remains below the potential evapotranspiration, the effect on the modeled soil moisture will thus be more pronounced when the precipitation is underestimated.

Figure 6 shows the impact of the assimilation procedure on the modeled lower layer soil moisture, again for 1994 only. The same conclusions, as can be expected, can be drawn for the lower-layer soil moisture as for the upper-layer soil moisture: in both cases the error in the modeled soil moisture is reduced, but the effect of the errors in the precipitation, and the consequent improvement through data assimilation, is more pronounced when the precipitation is underestimated. Figure 6 shows that after the data assimilation a higher error in the modeled lower-layer soil moisture remains than for the modeled upper-layer soil moisture.

Table 2 shows that, when the precipitation is overestimated, the reduction in the upper-layer soil moisture and the increase in the water table depth lead to an increase in the infiltration, a reduction in the surface runoff, and a reduction in the base flow. For the case of underestimated precipitation, the opposite occurs. The improvement in the modeled discharge is thus due to an improvement in both the modeled base flow and surface runoff.

Figure 7 shows the bias and the standard deviation of the noise of the estimated water table levels. For both model applications the bias tends to be rather low in magnitude, rarely exceeding 5 cm. An overestimation of the precipitation leads to a slightly negative bias, thus indicating a small underestimation of the water table depth. An underestimation of the precipitation leads to the opposite effect. Figure 7 further shows that the error in the estimation of the water level is relatively low (rarely exceeding 10 cm), and that an overestimation of the precipitation leads to a larger uncertainty in the

modeled water table level than an underestimation. A seasonal cycle in the error can also be observed, which can be attributed to the seasonal cycle in the evapotranspiration. As evapotranspiration is very dependent on the amount of precipitation, an error in the precipitation will thus have its highest effect during the summer, because the evapotranspiration rates are highest in the summer. Similar results were obtained for the bias and noise of the upper- and lower-layer soil moisture contents.

Table 3 shows the results of the assimilation procedure on the modeled energy balance terms. When the precipitation is overestimated, the latent heat flux is slightly increased by the increase in precipitation, and the data assimilation has a negligible impact on the modeled latent heat flux (or evapotranspiration), as well as on the other energy balance terms. Together with Table 1, Fig. 6 leads to the conclusion that the increase in precipitation is partitioned mostly into an increase in the modeled discharge and not the modeled evapotranspiration, and that this can be explained by the limited sensitivity of the modeled soil moisture to the increase in precipitation.

When the precipitation is underestimated, an improvement in the modeled latent heat flux (or evapotranspiration) and sensible heat flux can be observed, while the net radiation and ground heat flux are essentially insensitive to the data assimilation procedure. The improvement in the modeled turbulent fluxes can be explained by the improvement in the modeled soil moisture, as shown in Fig. 6. The underestimation of the precipitation is thus partitioned not only into an underestimation of the discharge, but also an underestimation of the evapotranspiration. Because the data assimilation procedure does not fully correct the underestimation of the lower-layer soil moisture content, the modeled latent heat fluxes will still be underestimated when discharge data are assimilated, but a reduction in the bias of more than 50% is nevertheless obtained.

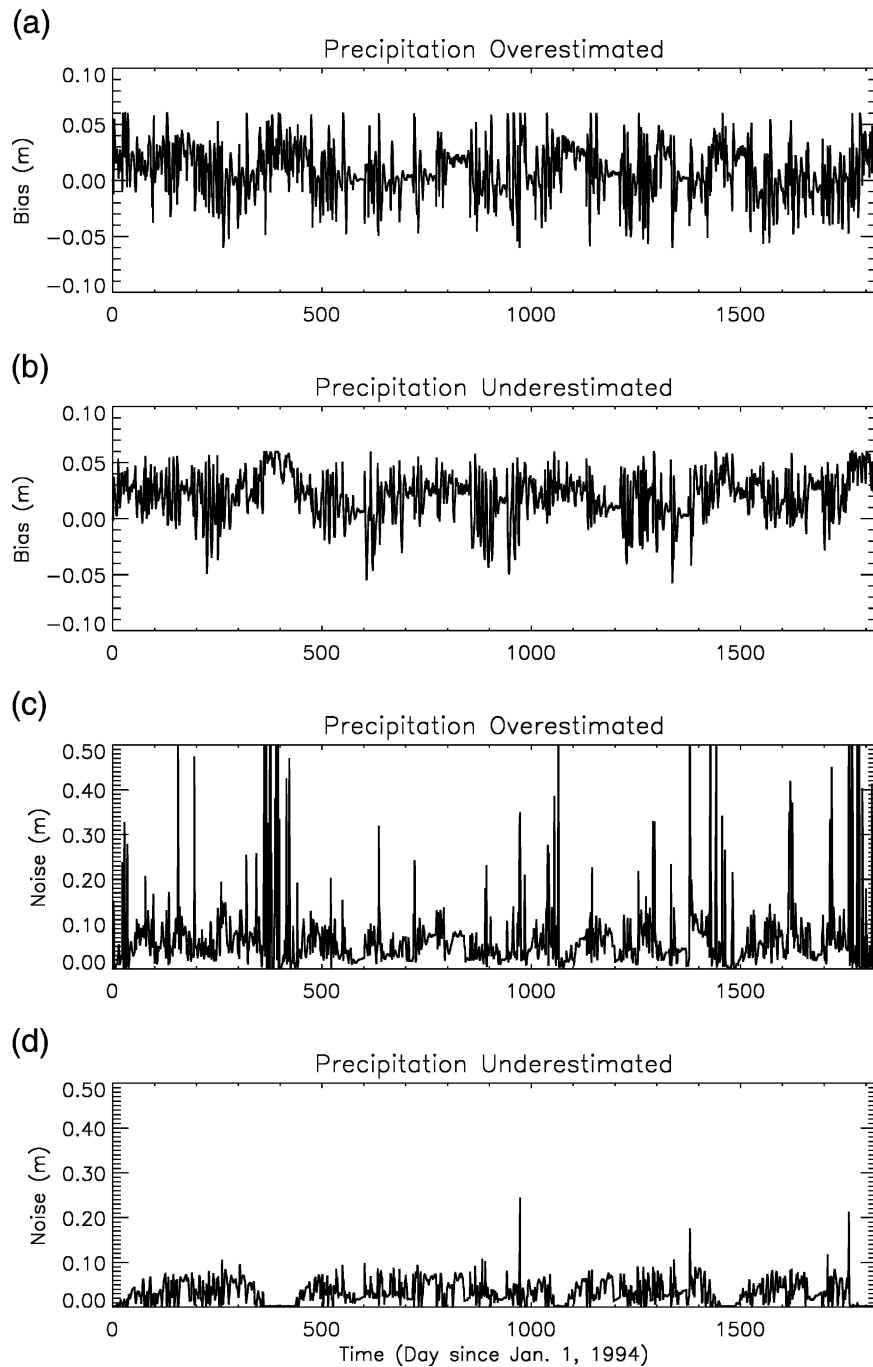Figure 8 shows the impact of the assimilation of the

FIG. 7. Bias and std dev of the error in the modeled water table levels in case of erroneous meteorological forcings.

synthetic discharge data on the monthly average diurnal cycles of the modeled latent heat fluxes. Here it can be seen that the largest improvement occurs when the precipitation is underestimated, and that this improvement is the weakest during the winter. During the other seasons the improvement is relatively constant through-

out the year. Figure 9 shows these diurnal cycles for the modeled sensible heat flux. The same conclusions can be drawn as for the latent heat flux.

As a summary, the results of the four twin experiments suggest that there is potential to improve the results of hydrologic models, but that this improvement

TABLE 3. Results of the data assimilation on the modeled energy balance parameters for the model run with overestimated precipitation. Here, R_n is the modeled net radiation; LE, H, and G are the modeled latent, sensible, and ground heat fluxes, respectively; ET is the modeled evapotranspiration; and Q is the modeled discharge. The synthetic observations are the independent variable (x axis); the model applications are the dependent variable (y axis). Units are in W m^{-2}, except for the evapotranspiration, which is in mm per hour, and for the discharge, which is in m^3 s^{-1}. The averages for the evapotranspiration are recalculated to mm per year. The top part of the table is for overestimated precipitation, and the bottom part for underestimated precipitation.

	$\overline{\text{X}}$	Baseline run		Error 0.00 m^3 s^{-1}		Error 0.01 m^3 s^{-1}		Error 0.10 m^3 s^{-1}		Error 0.50 m^3 s^{-1}		γ 0	
		$\overline{\text{Y}}$	rmse	$\overline{\text{Y}}$	rmse	$\overline{\text{Y}}$	rmse	$\overline{\text{Y}}$	rmse	$\overline{\text{Y}}$	rmse	$\overline{\text{Y}}$	rmse
R_n	96.88	97.26	1.00	97.11	0.69	97.11	0.69	97.11	0.68	97.09	0.71	97.09	0.65
LE	45.08	48.05	6.20	47.17	4.77	47.16	4.75	47.12	4.62	47.06	4.61	47.04	4.50
ET	336.73	374.98	0.0093	363.84	0.0071	363.71	0.0070	363.11	0.0069	362.32	0.0068	362.13	0.0067
H	47.12	44.33	6.80	45.15	4.71	45.16	4.69	45.20	4.58	45.27	4.63	45.28	4.40
G	4.68	4.88	2.58	4.79	1.65	4.79	1.64	4.79	1.65	4.76	1.85	4.77	1.58
Q	1.53	2.55	2.30	1.68	1.14	1.71	1.19	1.84	1.63	1.92	1.39	1.66	1.35
R_n	96.88	96.04	1.98	96.44	1.32	96.44	1.32	96.47	1.34	96.68	1.47	96.44	1.34
LE	45.08	38.36	13.30	41.67	9.28	41.68	9.28	41.88	9.32	43.21	9.40	41.58	9.48
ET	336.73	250.27	0.0197	293.10	0.0136	293.14	0.0136	295.83	0.0137	313.06	0.0138	291.92	0.0139
H	47.12	53.39	13.74	50.33	8.42	50.33	8.42	50.13	8.49	48.86	8.81	50.43	8.64
G	4.68	4.29	4.91	4.43	3.32	4.43	3.31	4.45	3.39	4.61	4.01	4.43	3.31
Q	1.53	0.61	1.47	1.53	0.61	1.53	0.62	1.64	0.73	2.09	1.14	1.48	0.60

is dependent on the meteorological conditions during the model application.

e. Sensitivity to observation error

To more realistically assess the effect of the assimilation of observed discharge, a Gaussianly distributed error was added to the synthetical discharge observations. An error of 0.01, 0.1, and 0.5 m^3 s^{-1} was assumed. The effect of this model error on the modeled energy balance was assessed.

Table 3 shows the effect of this observation error. It can clearly be seen, as can be expected, that the increase in the observation error leads to a decrease in the accuracy of the modeled discharge. However, this does not have a strong impact on the modeled energy balance. When the precipitation is overestimated, even an error of 0.5 m^3 s^{-1} leads to a similar improvement as perfect observations. When the precipitation is underestimated, the accuracy in the modeled energy balance slightly increases as the observation error decreases from 0.5 to 0.1 m^3 s^{-1}, but a further decrease in the observation error does not lead to a further improvement in the model results. From this sensitivity analysis we conclude that even under considerable errors in the synthetic observations (for the Zwalm catchment an error of 0.1 m^3 s^{-1} is a relatively high error), there remains a potential to improve the modeled water and energy balance terms through the assimilation of discharge data.

f. Sensitivity to the estimate of γ

In many cases it will be impossible to make an accurate estimate of γ, more specifically if no observations of soil moisture and/or water table levels are available. The objective of this section is to assess whether the assimilation algorithm will also provide acceptable results if the bias correction algorithm is omitted, in other words if γ is set to zero. For these model applications the observation error was also set to zero. Table 3 shows the results of these model applications, for the cases in which the precipitation was over- and underestimated. When the precipitation is overestimated, the bias between the model results and the synthetic observations is slightly larger as compared to the model runs where γ is set to 0.6, but the rmse is slightly lower. When the precipitation is underestimated the opposite behavior can be observed. Since the differences between the model results with γ zero and 0.6 are so low, it can be concluded that, when it is impossible to estimate γ, a value of 0 will lead to results that will be similar to the results with an accurate estimate γ, even though these results will not be accurate estimates.

7. Summary and conclusions

A method has been developed to update the past state of a land surface model using discharge observations in a retrospective ensemble Kalman filter framework. The method allows the correction of bias in the model forecasts. For each model state update, one discharge observation is used, although it is possible to use any number of discharge observations for the updating of the model. In a twin experiment it is shown that erroneous initial conditions can be corrected for through the assimilation of synthetically observed discharge. A second set of twin experiments is then devel-
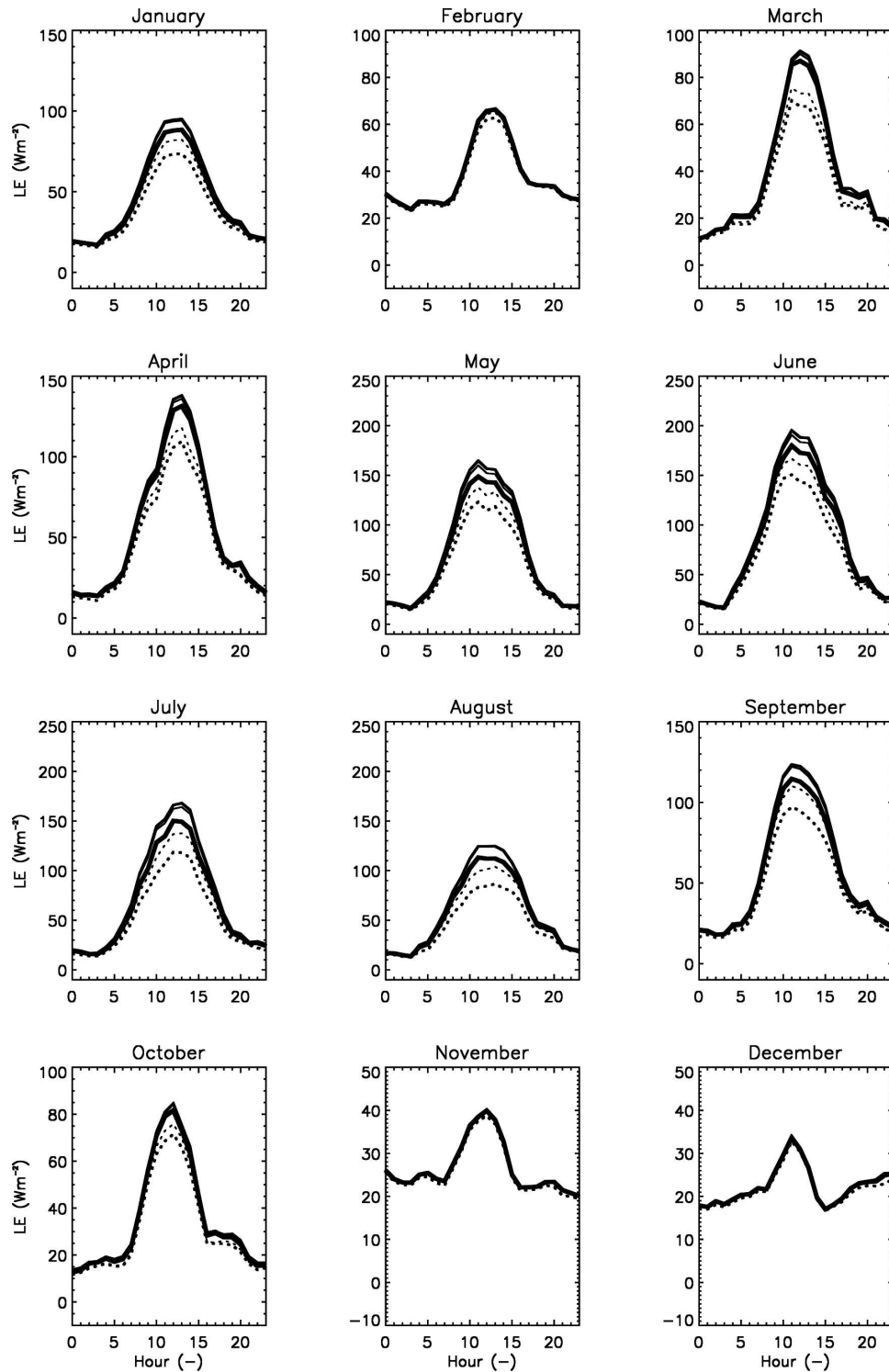
FIG. 8. Impact of the data assimilation on the monthly average diurnal cycle of the modeled latent heat fluxes (LE). The monthly averages are calculated for the 5-yr simulation period. The double-thick solid lines are the synthetic observations, the thick solid lines are the results of the baseline run with underestimated precipitation, and the thick dotted lines are the results of the baseline run with overestimated precipitation. The thin solid lines are the results of the baseline run with underestimated precipitation, and the thin dotted lines are the results of the baseline run with overestimated precipitation.
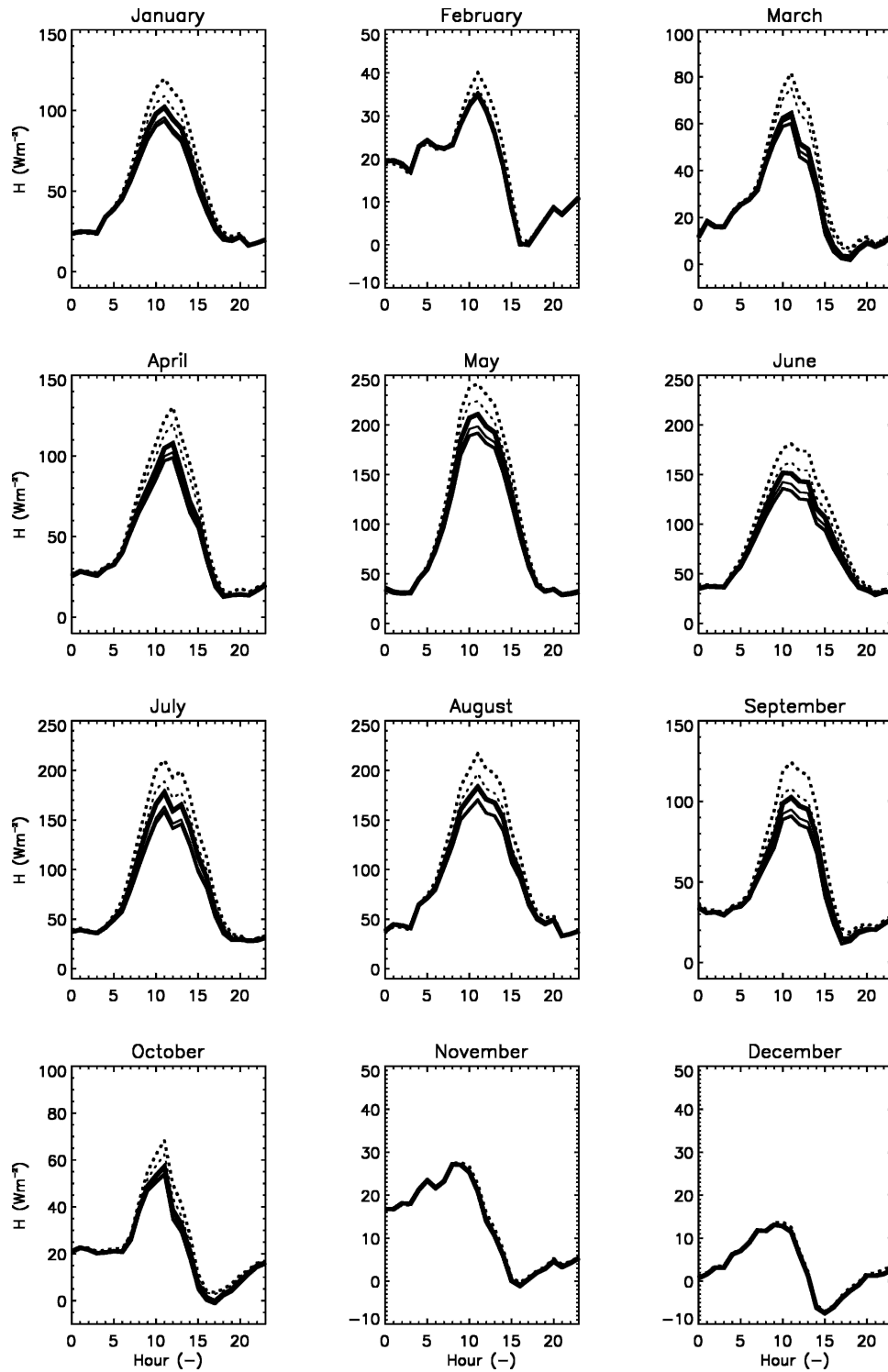
FIG. 9. Impact of the data assimilation on the monthly average diurnal cycle of the modeled sensible heat fluxes (H). The monthly averages are calculated for the 5-yr simulation period. The double-thick solid lines are the synthetic observations, the thick solid lines are the results of the baseline run with underestimated precipitation, and the thick dotted lines are the results of the baseline run with over-estimated precipitation. The thin solid lines are the results of the baseline run with underestimated precipitation, and the thin dotted lines are the results of the baseline run with overestimated precipitation.

oped, in which the precipitation is first strongly over-estimated, and then strongly underestimated. A reduction in the bias between the synthetically observed and modeled soil wetness conditions and fluxes is obtained through the assimilation procedure. When the precipitation is underestimated, these improved soil wetness conditions lead to an improvement in the modeled energy balance. The results have been found to be insensitive to the error in the observed discharge data, and to the estimate of the bias parameter γ. The results indicate that it is possible to improve the performance of land surface models through the assimilation of observed discharge records. Further research will focus on the determination of the optimal size of the assimilation window and on the application of the assimilation method for forecasting purposes.

REFERENCES

Aubert, D., C. Loumagne, and L. Oudin, 2003: Sequential assimilation of soil moisture and streamflow data in a conceptual rainfall-runoff model. *J. Hydrol.,* **280,** 145–161.

Beven, K. J., 2000: *Rainfall-Runoff Modelling: The Primer.* John Wiley and Sons, 360 pp.

——, and M. J. Kirkby, 1979: A physically based, variable contributing area model of basin hydrology. *Hydrol. Sci. Bull.,* **24,** 43–69.

Boiten, W., 2000: *Hydrometry.* A. A. Balkema, 246 pp.

Boni, G., D. Entekhabi, and F. Castelli, 2001: Land data assimilation with satellite measurements for the estimation of surface energy balance components and surface control on evaporation. *Water Resour. Res.,* **37,** 1713–1722.

Bras, R., 1980: Short-term forecasting of rainfall and runoff. *Real-Time Forecasting/Control of Water Resource System: Selected Papers from an IIASA Workshop, October 18–21, 1976,* E. F. Wood and A. Szöllösi-Nagy, Eds., Pergamon, 125–137.

Brutsaert, W., 1973: Review of Green's functions for linear open channels. *J. Eng. Mech. Div. ASCE,* **99,** 1247–1257.

Burgers, G., P. J. van Leeuwen, and G. Evensen, 1998: Analysis scheme in the ensemble Kalman filter. *Mon. Wea. Rev.,* **126,** 1719–1724.

Caparrini, F., F. Castelli, and D. Entekhabi, 2003: Mapping of land–atmosphere heat fluxes and surface parameters with remote sensing data. *Bound.-Layer Meteor.,* **107,** 605–633.

——, ——, and ——, 2004: Variational estimation of soil and vegetation turbulent transfer and heat flux parameters from

sequences of multisensor imagery. *Water Resour. Res.,* **40,** W12515, doi:10.1029/2004WR003358.

Castelli, F., D. Entekhabi, and E. Caporali, 1999: Estimation of surface heat flux and an index of soil moisture using adjoint-state surface energy balance. *Water Resour. Res.,* **35,** 3115–3125.

Cohn, S. E., N. S. Sivakumaran, and R. Todling, 1994: A fixed-lag Kalman smoother for retrospective data assimilation. *Mon. Wea. Rev.,* **122,** 2838–2867.

Crow, W. T., and E. F. Wood, 2003: The assimilation of remotely sensed soil brightness temperature imagery into a land surface model using ensemble Kalman filtering: A case study based on ESTAR measurements during SGP97. *Adv. Water Resour.,* **26,** 137–149.

——, and W. P. Kustas, 2005: Utility of assimilating surface radiometric temperature observations for evaporative fraction and heat transfer coefficient retrieval. *Bound.-Layer Meteor.,* **115,** 105–130.

Dee, D. P., and A. M. Da Silva, 1998: Data assimilation in the presence of forecast bias. *Quart. J. Roy. Meteor. Soc.,* **124,** 269–295.

De Lannoy, G. J. M., P. R. Houser, V. R. N. Pauwels, and N. E. C. Verhoest, 2005: Assessment of model uncertainty for soil moisture through ensemble verification. *J. Geophys. Res.,* in press.

Dunne, S., and D. Entekhabi, 2005: An ensemble-based reanalysis approach to land data assimilation. *Water Resour. Res.,* **41,** W02013, doi:10.1029/2004WR003449.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.,* **99** (C5), 10 143–10 162.

——, and P. J. van Leeuwen, 2000: An ensemble Kalman smoother for nonlinear dynamics. *Mon. Wea. Rev.,* **128,** 1852–1867.

Famiglietti, J. S., and E. F. Wood, 1994: Multiscale modeling of spatially variable water and energy balance processes. *Water Resour. Res.,* **30,** 3061–3078.

Galantowicz, J. F., D. Entekhabi, and E. Njoku, 1999: Test of sequential data assimilation for retrieving profile soil moisture and temperature from observed *l*-band radio brightness. *IEEE Trans. Geosci. Remote Sens.,* **37,** 1860–1870.

Ganendra, T., 1980: A self-tuning predictor applied to river flow forecasting. *Real-Time Forecasting/Control of Water Resource System: Selected Papers from an IIASA Workshop, October 18–21, 1976,* E. F. Wood and A. Szöllösi-Nagy, Eds., Pergamon, 139–149.

Grigoriu, M., 1978: Prediction and simulation with autoregressive models in hydrology. *Applications of Kalman Filter to Hydrology, Hydraulics, and Water Resources: Proceedings of AGU Chapman Conference Held at the University of Pittsburgh, May 22–24, 1978,* C.-L. Chiu, Ed., University of Pittsburgh, 409–437.

Heathman, G. C., P. J. Starks, L. R. Ahuja, and T. J. Jackson, 2003: Assimilation of surface soil moisture to estimate profile soil water content. *J. Hydrol.,* **279,** 1–17.

Hoeben, R., and P. A. Troch, 2000: Assimilation of active microwave observation data for soil moisture profile estimation. *Water Resour. Res.,* **36,** 2805–2819.

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

Katayama, T., 1980: Application of maximum likelihood identification to river flow prediction. *Real-Time Forecasting/*

Control of Water Resource System: Selected Papers from an IIASA Workshop, October 18–21, 1976, E. F. Wood and A. Szöllösi-Nagy, Eds., Pergamon, 63–82.

KMI, 1994–1998: Maandbericht klimatologische waarnemingen (Monthly climatological observations). Part 1. Koninklijk Meteorologisch Instituut van Belgie, 46 pp.

Lakshmi, V., 2000: A simple surface temperature assimilation scheme for use in land surface models. *Water Resour. Res.,* **36,** 3687–3700.

Logan, W. C., W. N. Lennox, and T. E. Unny, 1978: A time-varying non-linear hydrologic response model for flood hydrography estimation in a noisy environment. *Applications of Kalman Filter to Hydrology, Hydraulics, and Water Resources: Proceedings of AGU Chapman Conference Held at the University of Pittsburg, May 22–24, 1978,* C.-L. Chiu, Ed., University of Pittsburg, 279–293.

Margulis, S. A., and D. Entekhabi, 2003: Variational assimilation of radiometric surface temperature and reference-level micrometeorology into a model of the atmospheric boundary layer and land surface. *Mon. Wea. Rev.,* **131,** 1272–1288.

——, D. McLauglin, D. Entekhabi, and S. Dunne, 2002: Land data assimilation and estimation of soil moisture using measurements from the Southern Great Plains 1997 Field Experiment. *Water Resour. Res.,* **38,** 1299, doi:10.1029/2001WR001114.

Mesa, O. J., and E. R. Mifflin, 1986: On the relative role of hillslope and network geometry in hydrologic response. *Scale Problems in Hydrology,* V. K. Gupta, I. Rodriguez-Iturbe, and E. F. Wood, Eds., D. Reidel, 1–17.

Paniconi, C., M. Maroccu, M. Putti, and M. Verbunt, 2002: Newtonian nudging for a Richards equation-based distributed hydrological model. *Adv. Water Resour.,* **26,** 161–178.

Pauwels, V. R. N., and E. F. Wood, 1999a: A soil–vegetation–atmosphere transfer scheme for the modeling of water and energy balance processes in high latitudes. 1. Model improvements. *J. Geophys. Res.,* **104** (D22), 27 811–27 822.

——, and ——, 1999b: A soil–vegetation–atmosphere transfer scheme for the modeling of water and energy balance processes in high latitudes. 2. Application and validation. *J. Geophys. Res.,* **104** (D22), 27 823–27 839.

——, and ——, 2000: The importance of misclassifications and spatial resolution of land cover data in the uncertainty in model results over boreal ecosystems. *J. Hydrometeor.,* **1,** 255–266.

——, R. Hoeben, N. E. C. Verhoest, and F. P. De Troch, 2001: The importance of the spatial patterns of remotely sensed soil moisture in the improvement of discharge predictions for small-scale basins through data assimilation. *J. Hydrol.,* **251,** 88–102.

——, ——, ——, ——, and P. A. Troch, 2002: Improvement of TOPLATS-based discharge predictions through assimilation of ERS-based remotely sensed soil moisture values. *Hydrol. Processes,* **16,** 995–1013.

Peters-Lidard, C. D., M. S. Zion, and E. F. Wood, 1997: A soil–vegetation–atmosphere transfer scheme for modeling spatially variable water and energy balance processes. *J. Geophys. Res.,* **102** (D4), 4303–4324.

Philip, J. R., 1957: The theory of infiltration. 4. Sorptivity and algebraic equations. *Soil Sci.,* **84,** 257–264.

Rawls, W. J., D. L. Brakensiek, and K. E. Saxton, 1982: Estimation of soil water properties. *Trans. ASAE,* **25,** 1316–1320.

Reichle, R. H., D. Entekhabi, and D. B. McLaughlin, 2001a: Downscaling of radio brightness measurements for soil moisture estimation: A four-dimensional variational data assimilation approach. *Water Resour. Res.,* **37,** 2353–2364.

——, D. B. McLaughlin, and D. Entekhabi, 2001b: Variational data assimilation of microwave radiobrightness observations for land surface hydrology applications. *IEEE Trans. Geosci. Remote Sens.,* **39,** 1708–1718.

——, ——, and ——, 2002a: Hydrologic data assimilation with the ensemble Kalman filter. *Mon. Wea. Rev.,* **130,** 103–114.

——, J. P. Walker, R. D. Koster, and P. R. Houser, 2002b: Extended versus ensemble Kalman filtering for land data assimilation. *J. Hydrometeor.,* **3,** 728–740.

Schuurmans, J. M., P. A. Troch, A. A. Veldhuizen, W. G. M. Bastiaanssen, and M. F. P. Bierkens, 2003: Assimilation of remotely sensed latent heat fluxes in a distributed hydrological model. *Adv. Water Resour.,* **26,** 151–159.

Seuffert, G., H. Wilker, P. Viterbo, M. Drusch, and J. F. Mahfouf, 2004: The usage of screen-level parameters and microwave brightness temperature for soil moisture analysis. *J. Hydrometeor.,* **5,** 516–531.

Sivapalan, M., K. Beven, and E. F. Wood, 1987: On hydrologic similarity. 2. A scaled model for runoff prediction. *Water Resour. Res.,* **23,** 2266–2278.

Todini, E., and J. R. Wallis, 1978: A real time rainfall-runoff model for an on line flood warning system. *Applications of Kalman Filter to Hydrology, Hydraulics, and Water Resources: Proceedings of AGU Chapman Conference Held at the University of Pittsburgh, May 22–24, 1978,* C.-L. Chiu, Ed., University of Pittsburgh, 355–367.

Todling, R., and S. E. Cohn, 1998: Suboptimal schemes for retrospective data assimilation based on the fixed-lag Kalman filter. *Mon. Wea. Rev.,* **126,** 2274–2286.

Troch, P. A., F. P. De Troch, and W. Brutsaert, 1993: Effective water table depth to describe initial conditions prior to storm rainfall in humid regions. *Water Resour. Res.,* **29,** 427–434.

——, J. A. Smith, E. F. Wood, and F. P. De Troch, 1994: Hydrologic controls of large floods in a small basin: Central Appalachian case study. *J. Hydrol.,* **156,** 285–309.

——, C. Paniconi, and D. McLaughlin, 2003: Catchment-scale hydrological modeling and data assimilation. *Adv. Water Resour.,* **26,** 131–135.

Walker, J. P., and P. R. Houser, 2001: A methodology for initializing soil moisture in a global climate model: Assimilation of near-surface soil moisture observations. *J. Geophys. Res.,* **106** (D11), 11 761–11 774.

——, G. R. Willgoose, and J. D. Kalma, 2001: One-dimensional soil moisture profile retrieval by assimilation of near-surface observations: A comparison of retrieval algorithms. *Adv. Water Resour.,* **24,** 631–650.

Wood, E. F., 1978: An application of Kalman filtering to river flow forecasting. *Applications of Kalman Filter to Hydrology, Hydraulics, and Water Resources: Proceedings of AGU Chapman Conference Held at the University of Pittsburgh, May 22–24, 1978,* C.-L. Chiu, Ed., University of Pittsburgh, 385–407.

——, and Coauthors, 1998: The Project for Intercomparison of Land-surface Parameterization Schemes (PILPS) phase 2(c) Red–Arkansas River basin experiment: 1. Experiment description and summary intercomparisons. *Global Planet. Change,* **19,** 115–135.