

## Correcting for forecast bias in soil moisture assimilation with the ensemble Kalman filter

Gabriëlle J. M. De Lannoy,<sup>1</sup> Rolf H. Reichle,<sup>2,3</sup> Paul R. Houser,<sup>4</sup> Valentijn R. N. Pauwels,<sup>1</sup> and Niko E. C. Verhoest<sup>1</sup>

Received 21 August 2006; revised 5 June 2007; accepted 21 June 2007; published 19 September 2007.

[1] Land surface models are usually biased in at least a subset of the simulated variables even after calibration. Bias estimation may therefore be needed for data assimilation. Here, in situ soil moisture profile observations in a small agricultural field were merged with Community Land Model (CLM2.0) simulations using different algorithms for state and forecast bias estimation with and without bias correction feedback. Simple state updating with the conventional ensemble Kalman filter (EnKF) allows for some implicit forecast bias correction. It is possible to estimate the soil moisture bias explicitly and derive superior soil moisture estimates with a generalized EnKF that uses a simple persistence model for the bias and assumes that the a priori bias error covariance is proportional to the a priori state error covariance. For the case of bi-weekly assimilation of the entire profile of soil moisture observations, bias estimation and correction typically reduces the *RMSE* in soil moisture (over the standard EnKF without bias correction) by around 60 percent. However, under the above assumptions, significant improvements are limited to state variables for which observations are available. Therefore, it is crucial to measure the state variables of interest. The best variant for state and bias estimation depends on the nature of the model bias and the output of interest to the user. In a model that is only biased for soil moisture, large and frequent increments for soil moisture updating may be required, which in turn may negatively impact the water balance and output fluxes. It is then better to post-process the soil moisture with the bias analysis without updating the model state.

**Citation:** De Lannoy, G. J. M., R. H. Reichle, P. R. Houser, V. R. N. Pauwels, and N. E. C. Verhoest (2007), Correcting for forecast bias in soil moisture assimilation with the ensemble Kalman filter, *Water Resour. Res.*, 43, W09410, doi:10.1029/2006WR005449.

### 1. Introduction

[2] Through a large number of synthetic experiments [e.g., *Hoeben and Troch*, 2000; *Reichle et al.*, 2001; *Walker and Houser*, 2004; *Pauwels et al.*, 2007], the hydrologic community has gained an increased understanding of different data assimilation techniques and requirements for observations and land surface models that may improve the estimation of hydrologic states and related water and energy fluxes. However, the assimilation of real data [*Houser et al.*, 1998; *Crow and Wood*, 2003; *Margulis et al.*, 2002; *Reichle and Koster*, 2005] presents difficulties that still need to be addressed far more comprehensively, including the characterization of a priori model and observation error statistics, and the presence of bias in models and observations in particular.

[3] A fundamental assumption of the standard Kalman filter is that the observations as well as the model are unbiased. Bias in observations typically reflects instrumental inaccuracies, representativeness errors, or, in the case of remote sensing observations, errors in the retrieval algorithm. After quality control [*Lorenz and Hammon*, 1988] one can often, but not always, assume that the observations are largely unbiased. Model forecasts, however, are hardly ever unbiased. Details of the forecast error depend on the characteristics of the model (including model structure, parameters, and discretization) and on the model initial conditions, which are often influenced by observations in a cycling data assimilation system. In general, the forecast error contains a random and a systematic component. The latter could be constant or varying with time.

[4] Many authors have addressed the discrepancy in soil moisture climatologies from observations and models with a focus on bias removal prior to (and outside of) data assimilation [*Reichle et al.*, 2004; *Reichle and Koster*, 2004; *Drusch et al.*, 2005; *Ni-Meister et al.*, 2005; *Crow et al.*, 2005]. Earlier, *Walker et al.* [2003] pointed to the need to deal with biased observations and the possibility of improper information propagation over a soil profile. *Walker et al.* [2001], for example, noticed difficulties in the estimation of soil moisture profiles because of bias caused by the model structure and parameterization. Forcing errors are

<sup>1</sup>Laboratory of Hydrology and Water Management, Ghent University, Ghent, Belgium.

<sup>2</sup>Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, Greenbelt, Maryland, USA.

<sup>3</sup>Goddard Earth Sciences and Technology Center, University of Maryland, Baltimore County, Baltimore, Maryland, USA.

<sup>4</sup>George Mason University and Center for Research on Environment and Water, Calverton, Maryland, USA.

another factor which may contribute to bias in output from (typically non-linear) land surface models [Berg *et al.*, 2003]. Walker and Houser [2004] demonstrated that – even in a twin experiment – bias in model forcings or observations that is not accounted for in the assimilation scheme may cause adverse effects when near-surface soil moisture observations are assimilated. Therefore, Pauwels and De Lannoy [2006] included a bias estimation algorithm even in a synthetic discharge assimilation experiment.

[5] A typical characteristic of many data assimilation studies is that the forecast skill of the Kalman filter is limited to the time horizon where the effects of the corrected initial conditions are dampened. Often, the model quickly drifts back to a biased state after the update. This suggests that state updating alone is not an adequate solution to improve model results persistently. Madsen and Skotner [2005], for example, added an error forecast model to compensate for the decline in river forecast skill in time. In a groundwater study, Drécourt and Madsen [2002] calculated bias off-line as the difference between piezometric head values at the assimilation points and applied kriging for extrapolation to the rest of the domain. For off-line bias correction schemes, typically a time mean or time function of the bias is estimated in advance from the model and observation or analysis climatologies.

[6] Under the condition that observations are unbiased, different procedures for on-line forecast bias estimate updating have been proposed. A common practice has been to augment the state vector of the original state estimation problem by adding some uncertain parameters that are designated as bias terms, or model error terms in general [Jazwinski, 1970]. This was already tested with ensemble Kalman filter methods [Evensen, 2003], including studies with different types of specific bias models [Baek *et al.*, 2006], and more general model error processes as discussed by Zupanski and Zupanski [2006] and by Reichle *et al.* [2002] for soil moisture assimilation in particular. State augmentation has also been applied to include model error in variational assimilation approaches [Griffith and Nichols, 2000]. A practical problem arises when the number of model error terms is on the order of the number of state variables and the augmented state vector becomes so large that filter computations become excessive. To keep the size of the augmented control vector manageable and the system observable, often only the random, systematic, or time-correlated part is included [Vidard *et al.*, 2004].

[7] To circumvent these problems, Friedland [1969] presented a technique known as ‘separate-bias estimation’ or ‘two-stage estimation’. In Friedland’s method, the estimation of the bias is essentially decoupled from the computation of the bias-blind estimate of the state. The naming ‘bias-blind’ (instead of Friedland’s ‘bias-free’) is adopted from Dee and da Silva [1998] and refers to estimates that contain bias. Variants of Friedland’s method have been applied for example by Dee and Todling [2000] in an atmospheric model and by Keppenne *et al.* [2005] and Chepurin *et al.* [2005] in oceanography. Few hydrologic applications of Friedland’s method have been reported. Drécourt *et al.* [2006] applied an ensemble Kalman filter to a simple one-dimensional synthetic groundwater problem and compared the use of a colored noise model to the implementation of the separate bias filter. De Lannoy *et al.*

[2007] showed preliminary results of soil moisture profile estimates by a state filter and those of a separate bias filter by combining real observations with land model state estimates. Moreover, Bosilovich *et al.* [2006] assimilated remotely sensed land surface (skin) temperature, dynamically estimated model bias with a variational scheme based on Dee and da Silva [1998], and demonstrated improvements in 2 m air temperature and sensible heat flux estimates. None of these studies discussed any alternative methods to optimize the results for particular kinds of model bias and only Bosilovich *et al.* [2006] studied the effect of bias estimation for one variable on dependent variables or fluxes (surface energy budget). In the present study, several variants of Friedland’s state and bias estimation were tested for the first time for soil moisture estimation of individual profiles with a land surface model and real soil moisture observations. We investigated whether the biased or the bias-corrected estimates should feed back into the (biased) model and how the bias estimates could be used for correcting model forecasts. Moreover, the effect of soil moisture filtering on fluxes was studied.

## 2. State and Bias Estimation

### 2.1. Definition of Forecast Bias

[8] A sequential data assimilation system successively alternates between a model propagation step and an update step. The nonlinear land surface model used in our system is expressed in the generic form

$$\hat{\mathbf{x}}_i^- = \mathbf{f}_{i,i-1}(\hat{\mathbf{x}}_{i-1}^+, \mathbf{u}_i), \quad (1)$$

where  $\mathbf{f}_{i,i-1}$  denotes the non-linear discrete transition function and  $\mathbf{u}_i$  the external forcings. The model propagates the land surface state from time  $i - 1$  to provide a forecast (or a priori) state estimate  $\hat{\mathbf{x}}_i^-$  at time  $i$ . The initial condition for the land surface model at time  $i - 1$  is the a posteriori state estimate (or analysis)  $\hat{\mathbf{x}}_{i-1}^+$ . This a posteriori state estimate is obtained from updating the a priori state estimate  $\hat{\mathbf{x}}_{i-1}^-$  at time  $i - 1$  (discussed below). If no update was performed, then formally  $\hat{\mathbf{x}}_{i-1}^+ = \hat{\mathbf{x}}_{i-1}^-$ .

[9] The forecast bias  $\mathbf{b}_i$  at time step  $i$  is defined as the expectation of the forecast error, which is given by the difference between the true state  $\mathbf{x}_i$  and the forecasted a priori state estimate  $\hat{\mathbf{x}}_i^-$ , that is  $\mathbf{b}_i = E[\mathbf{x}_i - \hat{\mathbf{x}}_i^-]$ . In the ensemble filtering framework discussed below, the expectation is computed as the ensemble mean. The forecast bias vector  $\mathbf{b}_i$  has thus the same dimension as the state vector. The symbol  $\tilde{\cdot}$  is used for bias-blind quantities that are not corrected with an estimate of the bias. In general, the forecast bias can be time dependent and will depend on the state and the parameters of the land surface model.

[10] For the bias, we assume a persistence model to propagate the bias estimate in time between successive assimilation updates, that is

$$\hat{\mathbf{b}}_i^- = \hat{\mathbf{b}}_{i-1}^+. \quad (2)$$

The bias estimate is sequentially updated through filtering (see below). This persistent bias assumption is appropriate if the bias evolves more slowly than the forecast [Dee and da

*Silva*, 1998; *Keppenne et al.*, 2005]. In our study, the model quickly drifted to its climatology after re-initialization, resulting in a rapid bias increase until a stable bias value was reached and our assumption was valid. An alternative approach was used by *Chepurin et al.* [2005], where the bias was a predefined function of time and bias updates were calculated to correct the state estimate, but not propagated in the bias model.

## 2.2. Separate Bias Estimation

[11] In this paper we estimate the bias on-line, that is within the cycling assimilation system. In the algorithm that was originally proposed by *Friedland* [1969], a conventional Kalman filter is used to derive a bias-blind state estimate (stage 1), and a second Kalman filter is used to estimate the bias (stage 2). Unlike in state augmentation, the separate bias estimation algorithm assumes a zero cross-correlation between the a priori state and bias estimation errors. *Friedland's* original algorithm does not modify the (biased) forecast model equations, but rather corrects the output.

[12] For an ensemble Kalman filter, the procedure can be summarized as follows. An ensemble of biased forecasts  $\hat{\mathbf{x}}_{j,i}^-$  ( $j = 1, \dots, N$ ) is generated by a biased non-linear model to obtain an estimate of the state's pdf:

$$\hat{\mathbf{x}}_{j,i}^- = \mathbf{f}_{i,i-1}(\hat{\mathbf{x}}_{j,i-1}^+, \mathbf{u}_i, \mathbf{w}_{j,i-1}), \quad (3)$$

with  $\mathbf{w}_{j,i}$  a realization of the (zero mean assumed) model error that represents the complete effect of perturbations to forcings, parameters and state variables. The a priori estimate  $\hat{\mathbf{x}}_{j,i}^-$  of the state is given by the ensemble mean. Through the non-linear model, the zero mean estimation error and the bias in  $\hat{\mathbf{x}}_{j,i-1}^+$  as well as the zero mean  $\mathbf{w}_{j,i-1}$  result in both random and systematic error in  $\hat{\mathbf{x}}_{j,i}^-$ . When an observation  $\mathbf{y}_i$  is available, each ensemble member  $j$  is updated individually to obtain the bias-blind a posteriori state estimate as follows (stage 1):

$$\hat{\mathbf{x}}_{j,i}^+ = \hat{\mathbf{x}}_{j,i}^- + \tilde{\mathbf{K}}_{x,i} [\mathbf{y}_{j,i} - \mathbf{H}_i \hat{\mathbf{x}}_{j,i}^-]. \quad (4)$$

[13] The observations are perturbed to ensure sufficient spread, i.e.,  $\mathbf{y}_{j,i} = \mathbf{y}_i + \mathbf{v}_{j,i}$ , with  $\mathbf{v}_{j,i}$  the imposed zero-mean Gaussian perturbation *Burgers et al.* [1998]. In this study, the observation operator  $\mathbf{H}_i$  linearly relates the observations  $\mathbf{y}_i$  to the state  $\mathbf{x}_i$ . It contains only 1 and 0, because the observations are direct measurements of the state variables. The innovations  $[\mathbf{y}_{j,i} - \mathbf{H}_i \hat{\mathbf{x}}_{j,i}^-]$  of the stage 1 filter contrast the observations with the corresponding model forecasts and provide important information for validating the performance of the filter (section 4). The bias-blind gain  $\tilde{\mathbf{K}}_{x,i}$  is identical for all ensemble members and given by:

$$\tilde{\mathbf{K}}_{x,i} = \tilde{\mathbf{P}}_{x,i}^- \mathbf{H}_i^T [\mathbf{H}_i \tilde{\mathbf{P}}_{x,i}^- \mathbf{H}_i^T + \mathbf{R}_i]^{-1}, \quad (5)$$

with  $\tilde{\mathbf{P}}_{x,i}^-$  the a priori bias-blind state error covariance (obtained from the sample covariance) and  $\mathbf{R}_i$  the observation error covariance, respectively.

[14] The bias forecast  $\hat{\mathbf{b}}_i^-$  follows the persistence model (equation (2)) (after initializing  $\hat{\mathbf{b}}_0^- = \mathbf{0}$  at time zero). Unlike

the state forecast, the bias is not perturbed (no ensemble members), and its uncertainty will be derived empirically (see below). The updated (a posteriori) bias estimate  $\hat{\mathbf{b}}_i^+$  is obtained as a linear combination of the a priori bias estimate  $\hat{\mathbf{b}}_i^-$  and the difference between the observations and the a priori bias-corrected state estimate (stage 2):

$$\hat{\mathbf{b}}_i^+ = \hat{\mathbf{b}}_i^- + \mathbf{K}_{b,i} [\mathbf{y}_i - \mathbf{H}_i (\hat{\mathbf{x}}_i^- + \hat{\mathbf{b}}_i^-)]. \quad (6)$$

[15] The gain for the bias update  $\mathbf{K}_{b,i}$  is given by [*Friedland*, 1969], *Dee and da Silva* [1998]:

$$\mathbf{K}_{b,i} = \mathbf{P}_{b,i}^- \mathbf{H}_i^T [\mathbf{H}_i \mathbf{P}_{b,i}^- \mathbf{H}_i^T + \mathbf{H}_i \tilde{\mathbf{P}}_{x,i}^- \mathbf{H}_i^T + \mathbf{R}_i]^{-1}, \quad (7)$$

with  $\mathbf{P}_{b,i}^-$  the a priori bias error covariance. If observations are not available at any given time step, no bias update is computed.

[16] A key assumption is now made to estimate  $\mathbf{P}_{b,i}^-$ . Typically, a lot of effort is expended on the careful estimation of the a priori state error covariance matrix  $\tilde{\mathbf{P}}_{x,i}^-$ . By contrast, little or no information is available about the bias error covariance  $\mathbf{P}_{b,i}^-$ . We follow *Dee and da Silva* [1998] and express  $\mathbf{P}_{b,i}^-$  as a fraction of  $\tilde{\mathbf{P}}_{x,i}^-$ :

$$\mathbf{P}_{b,i}^- = \frac{\gamma}{1-\gamma} \tilde{\mathbf{P}}_{x,i}^-. \quad (8)$$

[17] Given this assumption, the Kalman gain for the bias  $\mathbf{K}_{b,i}$  can then be determined by:

$$\mathbf{K}_{b,i} = \gamma \tilde{\mathbf{P}}_{x,i}^- \mathbf{H}_i^T [\mathbf{H}_i \tilde{\mathbf{P}}_{x,i}^- \mathbf{H}_i^T + (1-\gamma)\mathbf{R}_i]^{-1}, \quad (9)$$

with  $\gamma$  a parameter that may be tuned to optimize the filtering system *Keppenne et al.* [2005]. Moreover, for  $\gamma \ll 1$ , the gain for the bias update is simply a fraction of the gain for the state update, that is  $\mathbf{K}_{b,i} \approx \gamma \tilde{\mathbf{K}}_{x,i}$ , and the typically expensive computation of the gain matrix need only be performed once instead of twice per update step. In this study, we could not make use of this simplification, because we chose  $\gamma = 0.5$ , representing that the a priori bias estimate and the a priori state estimate were equally uncertain. However, it will be shown in the results that this value was too large to be optimal. Increasing  $\gamma$  has the effect of decreasing the memory of the bias estimator, and results in noisy bias estimates. Equation (8) is a very rough approximation in our case of flow-dependent error covariance matrices: it implies that  $\mathbf{P}_{b,i}^-$  is propagated by the land model instead of the bias model and there is no guarantee that the covariance structure of the state errors is similar to that of the bias errors.

[18] Finally, the bias-corrected a posteriori state estimate  $\hat{\mathbf{x}}_i^+$  is given by [*Friedland*, 1969; *Dee and da Silva*, 1998]:

$$\hat{\mathbf{x}}_i^+ = \hat{\mathbf{x}}_i^- + [\mathbf{I} - \tilde{\mathbf{K}}_{x,i} \mathbf{H}_i] \hat{\mathbf{b}}_i^+ \quad (10)$$

$$\hat{\mathbf{x}}_i^+ = \frac{1}{N} \sum_{j=1}^N \hat{\mathbf{x}}_{j,i}^+, \quad (11)$$

**Table 1.** Overview of Assimilation Algorithms<sup>a</sup>

| Method               | Model Re-initialized<br>with ( $l = 0$ and $\tilde{\mathbf{K}}_i = \mathbf{0}$ if obs. not available)  | Output State  |   | State Update | Bias Update | Feedback of<br>Bias Estimate |
|----------------------|--|---|---|--------------|-------------|------------------------------|
|                      |  | Analysis  | Forecast  |              |             |                              |
| EnKF                 | $\hat{\mathbf{x}}_{j,i}^+ = \hat{\mathbf{x}}_{j,i}^- + \tilde{\mathbf{K}}_i[\mathbf{y}_{j,i} - \mathbf{H}_i\hat{\mathbf{x}}_{j,i}^-]$  | $\hat{\mathbf{x}}_{j,i}^+$  | $\hat{\mathbf{x}}_{j,i}^-$                        | yes          | no          | n/a                          |
| EnBKF_0              | $\hat{\mathbf{x}}_{j,i}^-$   | $\hat{\mathbf{x}}_{j,i}^- + \hat{\mathbf{b}}_i^+$   | $\hat{\mathbf{x}}_{j,i}^- + \hat{\mathbf{b}}_i^-$ | no           | yes         | no                           |
| EnBKF_1              | (same as EnKF)   | $\hat{\mathbf{x}}_{j,i}^+ + [\mathbf{I} - \tilde{\mathbf{K}}_i\mathbf{H}_i] \hat{\mathbf{b}}_i^+$ | $\hat{\mathbf{x}}_{j,i}^- + \hat{\mathbf{b}}_i^-$ | yes          | yes         | no                           |
| EnBKF_2              | $\hat{\mathbf{x}}_{j,i}^{l+} = \hat{\mathbf{x}}_{j,i}^- + \tilde{\mathbf{K}}_i[\mathbf{y}_{j,i} - \mathbf{H}_i(\hat{\mathbf{x}}_{j,i}^- + \hat{\mathbf{b}}_i^+)]$  | $\hat{\mathbf{x}}_{j,i}^+ + \hat{\mathbf{b}}_i^+$   | $\hat{\mathbf{x}}_{j,i}^- + \hat{\mathbf{b}}_i^-$ | yes          | yes         | no                           |
| EnBKF_3              | $\hat{\mathbf{x}}_{j,i}^+ = \hat{\mathbf{x}}_{j,i}^- + \tilde{\mathbf{K}}_i[\mathbf{y}_{j,i} - \mathbf{H}_i\hat{\mathbf{x}}_{j,i}^-] + l[\mathbf{I} - \tilde{\mathbf{K}}_i\mathbf{H}_i]\hat{\mathbf{b}}_i^+$ | $\hat{\mathbf{x}}_{j,i}^+$  | $\hat{\mathbf{x}}_{j,i}^-$                        | yes          | yes         | partial                      |
| EnBKF_3 <sup>+</sup> | (same as EnBKF_3)  | $\hat{\mathbf{x}}_{j,i}^+$  | $\hat{\mathbf{x}}_{j,i}^- + \hat{\mathbf{b}}_i^-$ | yes          | yes         | complete                     |

<sup>a</sup>Bias estimates are always obtained as in equations (2) and (6). For clarity,  $\tilde{\mathbf{K}}_i \equiv \tilde{\mathbf{K}}_{x,i}$  and  $l = 1$ , if observations are available. If observations are not available, then formally  $\tilde{\mathbf{K}}_i = \mathbf{0}$  and  $l = 0$ .

with  $\mathbf{I}$  the identity matrix. Intuitively, the terms  $\hat{\mathbf{b}}_i^+ - \tilde{\mathbf{K}}_{x,i}\mathbf{H}_i\hat{\mathbf{b}}_i^+$  can be interpreted as  $(+\hat{\mathbf{b}}_i^+)$  a bias correction of the bias-blind a priori state estimate  $\hat{\mathbf{x}}_{j,i}^-$  and  $(-\tilde{\mathbf{K}}_{x,i}\mathbf{H}_i\hat{\mathbf{b}}_i^+)$  a bias correction of the innovations in equation (4). In Friedland's original algorithm, the bias-corrected analysis (equation (10)) is only used as output and is not used to re-initialize the model at the next propagation step. *Dee and Todling* [2000] clearly stated that for the analysis to be unbiased, the observation error must be zero mean and  $\hat{\mathbf{b}}_i^-$  must be an unbiased prediction of the forecast bias. The latter condition depends on the bias model and data coverage in space and time for updating.

### 2.3. Variant Algorithms for Separate State and Bias Estimation

[19] In this section we describe the variants of the above state and bias estimation algorithm that we explored. In the following, an important distinction is made between (i) the output analysis and forecasts, and (ii) the analysis that feeds back into the model as initial condition for the next propagation step. For the output states, the bias-corrected estimates (if available) are generally of most interest. For the model re-initialization step, on the other hand, it is not clear whether feeding back bias-blind, bias-corrected, or fully biased estimates is most beneficial. The best choice depends not only on the quality of the state estimates, but also on model-produced variables such as runoff and evaporation that depend on the state estimates. Land surface models are typically tuned to predict fluxes like runoff and evapotranspiration with a model-specific level of absolute soil moisture. If the model's soil moisture is generally biased, re-initializing the model with a bias-corrected soil moisture analysis may drive the model away from its soil moisture climatology and thus result in unrealistic output of dependent fluxes.

[20] Besides conventional state updating without explicit bias estimation (EnKF), separate state and bias estimation was tested (i) without feedback of the bias estimate on the analysis for model re-initialization (original *Friedland* [1969] algorithm, here labeled EnBKF\_1), (ii) with partial feedback of the bias estimate on the analysis through bias-corrected innovations (EnBKF\_2), and (iii) with complete feedback of the bias estimate on the analysis (EnBKF\_3). The possibility of bias estimation without state update was also investigated (EnBKF\_0). Details are described below, and Table 1 gives a summary. It should be noted that for a linear model and in the presence of bias, only EnBKF\_1 is optimal. Bias feedback destroys the optimality of the two-

stage algorithm, and the EnKF (without bias estimation) is anyway suboptimal in the presence of bias. Since the land model used here is non-linear, it is not clear a priori which algorithm performs best.

[21] The first algorithm that we tested, labeled EnBKF\_0, relies only on bias estimation and does not estimate the random errors via the state update. In this case the output analyses and forecasts are bias-corrected with the appropriate bias estimates, while the state update equation is ignored and the model integration proceeds without any feedback from the bias estimation (Table 1). In other words, the observations have no impact on the model integration whatsoever and affect only the output states through a post-processing bias analysis.

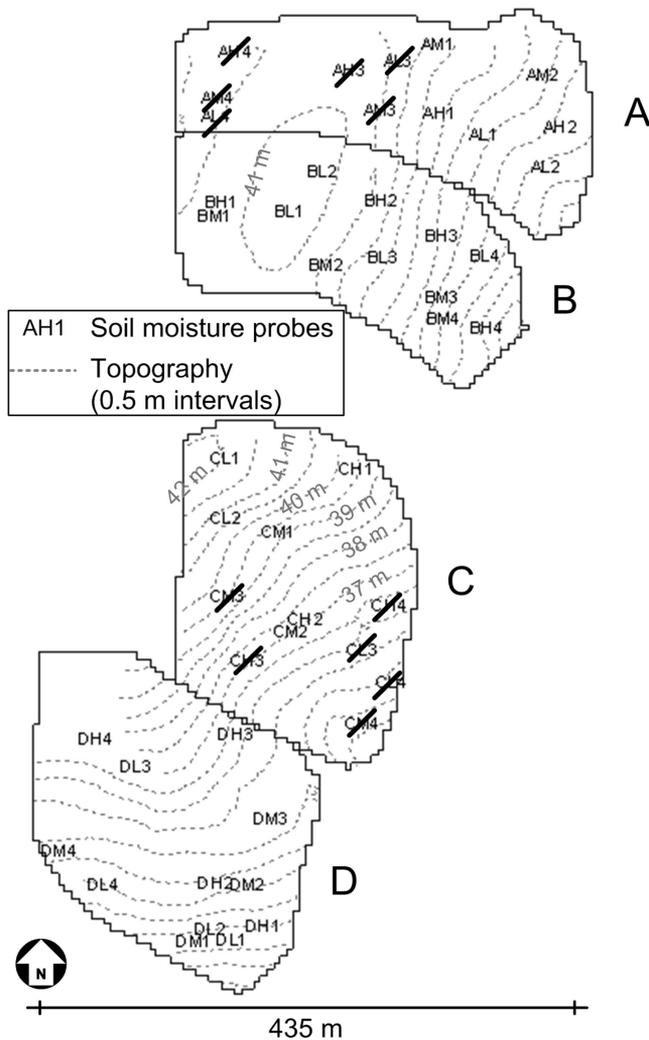
[22] In the original algorithm of *Friedland* [1969], here labeled EnBKF\_1, the state and bias updates are performed, and the bias-blind state is fed back into the model. In other words, the bias-blind a posteriori state estimate  $\hat{\mathbf{x}}_i^+$  given by equation (4) serves as an initial condition for the subsequent model propagation step, i.e., the model run is unchanged relative to EnKF. The bias-corrected output is obtained outside the model run as in equation equation (10).

[23] In EnBKF\_1, the bias-blind a posteriori estimate that is fed back into the model already contains an implicit partial correction for the bias (without any explicit use of the forecast bias estimate). As mentioned above, any bias correction could have an adverse effect on the model's dependent variables or fluxes (such as runoff and evaporation). Instead of computing the innovations in equation (4) from the bias-blind forecast, the implicit bias contribution can approximately be removed, so that the state update approximately corrects for random error only:

$$\hat{\mathbf{x}}_{j,i}^{l+} = \hat{\mathbf{x}}_{j,i}^- + \tilde{\mathbf{K}}_{x,i}[\mathbf{y}_{j,i} - \mathbf{H}_i(\hat{\mathbf{x}}_{j,i}^- + \hat{\mathbf{b}}_i^+)]. \quad (12)$$

[24] In this variant, labeled EnBKF\_2, the analysis that is used to re-initialize the model is not forced to come close to the truth. Rather, we attempt to propagate the model close to its biased state while removing as much of the random error as possible. To obtain the bias-corrected output state, we add the bias correction to the a priori and a posteriori estimates outside of the model run (Table 1).

[25] If the best a priori state estimates are generated when the model is initialized by a state close to the truth, then it seems logical to feed back the bias-corrected analyses into the model. Note again, however, that dependent model variables such as runoff and evaporation may be adversely



**Figure 1.** Digital elevation model with location of soil moisture probes in the field. Defective probes during the study period are crossed out.

affected by imposing the true state in a model that is tuned to work for biased states. In method EnBKF\_3, the bias-corrected analyses (equation (10)) are fed back and it is assumed that, starting from an unbiased or bias-corrected initial state or analysis, the forecasts are only negligibly biased. This method was called *on-line forecast-bias estimation and correction with feedback* by Dee and da Silva [1998] and applied by Dee and Todling [2000], Keppenne *et al.* [2005] and Chepurin *et al.* [2005]. Because in several experiments we observed a fast return of the model state to its biased climatology, in the additional variant EnBKF\_3<sup>+</sup> we correct the forecast outputs for bias without feedback of this bias correction in between analyses. Here, the model run is unchanged relative to EnBKF\_3.

#### 2.4. Remark

[26] Forecast bias could also be approximated as the time mean forecast error, or by a function that reflects cyclic components. Forecast bias could then be computed off-line by comparing long records of forecasts with those of (re-)analyses or observations (assuming they are unbiased) and such bias estimates would be readily available to correct

new forecasts. The off-line estimation of climatological differences between the model and the observations is also the basis for the rescaling approaches that have been developed recently [Reichle and Koster, 2004; Drusch *et al.*, 2005; Toth and Peña, 2006]. Such approaches reduce the discrepancy between model forecasts and observations through rescaling of the observations to the model climatology prior to data assimilation. In other words, rescaling approaches imply that only the anomaly information from the observations is assimilated. By design, rescaling reduces the average magnitude of the innovations and avoids that the model is pushed away from its own climatology through data filtering, in line with the idea behind EnBKF\_2 in our study.

[27] There are, however, important differences between the bias estimation variants discussed here and the rescaling approaches. Obviously, the climatological rescaling techniques are based on a priori knowledge, while in our on-line bias estimation methods the forecast bias is updated dynamically at each assimilation event. Moreover, rescaling does not per se assign the systematic errors to either the model or the observations, while our bias estimation variants assume unbiased observations and only account for (model) forecast bias. Another difference is that the rescaling approaches, in particular the approach based on matching cumulative distribution functions, also address systematic errors in higher order moments such as errors in the variability and skewness. By contrast, our on-line bias estimation variants with a persistent bias model estimate the mean errors.

### 3. Experimental Setup

#### 3.1. Model and Observations

[28] The filtering algorithms outlined above were tested with soil moisture observations at different depths and different locations in the 21 ha corn field in Maryland, USA, where the project for Optimizing Production Inputs for Economic and Environmental Enhancement (OPE<sup>3</sup>, <http://hydrolab.arsusda.gov/ope3/>) is conducted. In each of the four sub-watersheds (A, B, C, D) in the field, there are 12 capacitance probes for soil moisture measurements (EnviroSCAN, SENTEK Pty Ltd., South Australia), but during our study period only 36 of the 48 probes were operating (Figure 1). The probes were named following a 3 digit system Gish *et al.* [2002]. The first letter represents the sub-watershed (A, B, C, D). The second letter (L, M, H) refers to the estimated infiltration rate at the point of installation (Low, Moderate, High clay content). The third digit (1, 2, 3, 4) discerns between the different probes of a specific infiltration regime. H-probes have sensors at 10, 30, and 80 cm depth. L- and M-probes have sensors at 10, 30, 50, 120, 150, and 180 cm depth. L-probes have an additional sensor at 80 cm depth. The sensors measure in a soil volume with a radius of approximately 10 cm. There were 12 working probes of each type (H, M, L). The 10-min observations were aggregated into hourly time steps for comparison with model results.

[29] The vertical soil moisture profiles at all working 36 observation locations were modeled as independent point profiles with the Community Land Model (CLM2.0). The model simulates land surface processes by calculating

**Table 2.** Overview of Data Assimilation Frequencies and Number of Assimilation Events (Evenly Spaced Between 2 October 2001 and 19 March 2002)

| Frequency                     | a     | b      | c      | d      | e       | f       | g       |
|-------------------------------|-------|--------|--------|--------|---------|---------|---------|
| Interval                      | 1 day | 2 days | 4 days | 1 week | 2 weeks | 4 weeks | 8 weeks |
| Number of assimilation events | 169   | 85     | 43     | 25     | 13      | 7       | 4       |

vertical water and heat fluxes and states for each profile separately without interaction between profiles [Dai *et al.*, 2003]. CLM2.0 has one vegetation layer, a user-defined number of soil layers (we used 10), and up to 5 snow layers (depending on the snow depth). The depths of the soil nodes were set to 2.5, 5, 10, 20, 30, 50, 80, 120, 150, and 180 cm, so that soil moisture at all observation depths was included in the model state vector. The interface between the layers was halfway in between two consecutive nodes (except for the upper and lower layer) and hence the layer thickness was variable (schematic shown by De Lannoy *et al.* [2006a]). The land surface model was forced with hourly observed meteorological data, which were assumed to be spatially homogeneous over all profiles.

[30] Data from 1 May 2001 through 1 October 2001 were used for system identification (parameter calibration and initial state estimation), while during the period from 2 October 2001 through 30 April 2002 there was state and bias estimation and validation (split sample) for all 36 profiles. A detailed analysis of the soil moisture data set [Gish *et al.*, 2002; De Lannoy *et al.*, 2006b] revealed a complex subsurface hydrology, mainly caused by an irregular shaped clay layer at 1 to 3 m depth. The calibration and initialization of the model was performed for each profile individually [De Lannoy *et al.*, 2006a], but it was not possible to achieve zero bias for all soil layers in all profiles, mainly because lateral flow is not modeled by CLM2.0. Because of the structural deficiency and the a priori model calibration, further parameter estimation as part of the assimilation would not reduce the bias problem. Therefore, the soil moisture observations were assimilated to estimate the total forecast bias resulting from several error sources, including model structural errors and forcing errors.

[31] The OPE<sup>3</sup> soil moisture observations were assimilated into CLM2.0 during a fixed time period: the first assimilation event was always on 2 October 2001, the last one on 19 March 2002, and varying numbers of assimilation events/frequencies (Table 2) were considered. For the assimilation, the state vector for each profile consists of 22 prognostic variables, namely soil moisture and soil temperature at 10 levels, canopy water storage, and vegetation temperature. Assimilation of soil moisture thus also influences the other state variables, like soil temperature, which is important for balancing and to study the total effect on evapotranspiration. Since snow and soil freezing were negligible, model prognostic variables related to snow and frozen soil water were not included in the state vector. In a typical experiment, we assimilated in situ observations in one soil layer for all 36 profiles at selected time steps (Table 2) and verified against the (withheld) in situ data corresponding to this or the other profile layers over all time steps within the validation period (2 October 2001 through 30 April 2002). The effect of data assimilation in an individual soil layer on the performance of the complete profile was studied,

and for subsection 4.2 the effect of assimilation of all available data per profile on the evapotranspiration and runoff was investigated.

### 3.2. Practical Considerations

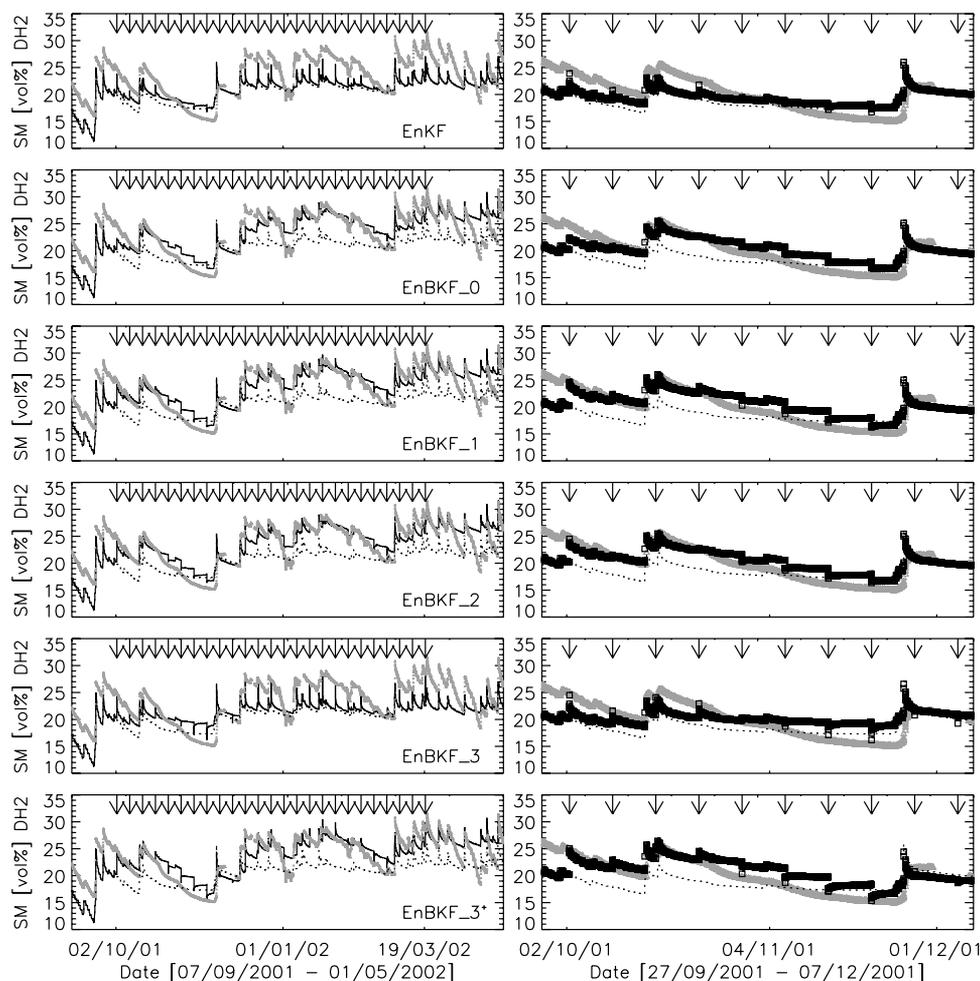
[32] The first practical obstacle encountered in the assimilation experiments was the estimation of the model and observation error statistics. The ensemble a priori state error covariance  $\tilde{\mathbf{P}}_{x,i}^-$  was computed as the sample covariance from an ensemble integration with  $N = 64$  members, because more members did not significantly alter the description of the predicted univariate state variable pdfs. Each member experienced realistic perturbations of all initial state values, parameters and forcings, as described by De Lannoy *et al.* [2006a]. Because of the model structure (independent profiles) and our choice of uncorrelated forcing perturbations, a priori state errors at different locations were effectively uncorrelated. The spurious elements of  $\tilde{\mathbf{P}}_{x,i}^-$  outside of the diagonal blocks (corresponding to the variables in one profile) were set to 0. This strategy amounts to a severe covariance localization that increases the rank of  $\tilde{\mathbf{P}}_{x,i}^-$  [Hamill *et al.*, 2001]. As mentioned above, the a priori bias error covariance  $\mathbf{P}_{b,i}^-$  was chosen proportional to  $\tilde{\mathbf{P}}_{x,i}^-$  for all experiments. The observations were assumed unbiased with error covariance  $\mathbf{R} = (0.022 \text{ m}^3 \text{ m}^{-3})^2 \cdot \mathbf{I}$ , assuming zero cross-correlation between the observation errors [De Lannoy *et al.*, 2007].

[33] For all algorithms it is necessary to make sure that the updated state variables are physically realistic. Whenever estimates are used to re-initialize the model, they must also be compatible with the model's parameters. Consequently, each state estimate (forecast/analysis) was checked and possibly altered before it was fed back into the model. For example, a corrected soil moisture that exceeded the model's porosity was re-set to the porosity before feedback. For the output state (possibly corrected outside the model run), the values were simply limited to physically realistic numbers, without regard to model-specific limitations. In any case, only a negligible amount of these kinds of interventions was required.

## 4. Results

### 4.1. Soil Moisture

[34] We begin our analysis by highlighting a few key features of the soil moisture estimates produced by the various assimilation algorithms. Figures 2 and 3 show examples of soil moisture time series at two distinct soil depths and locations where different types of forecast bias were present. For the assimilation integrations, only observations in the single soil layer that is shown in the figure have been assimilated, and the output analyses/forecasts of Table 1 are plotted. For reference, we also show the ensemble mean of an ensemble integration without assi-

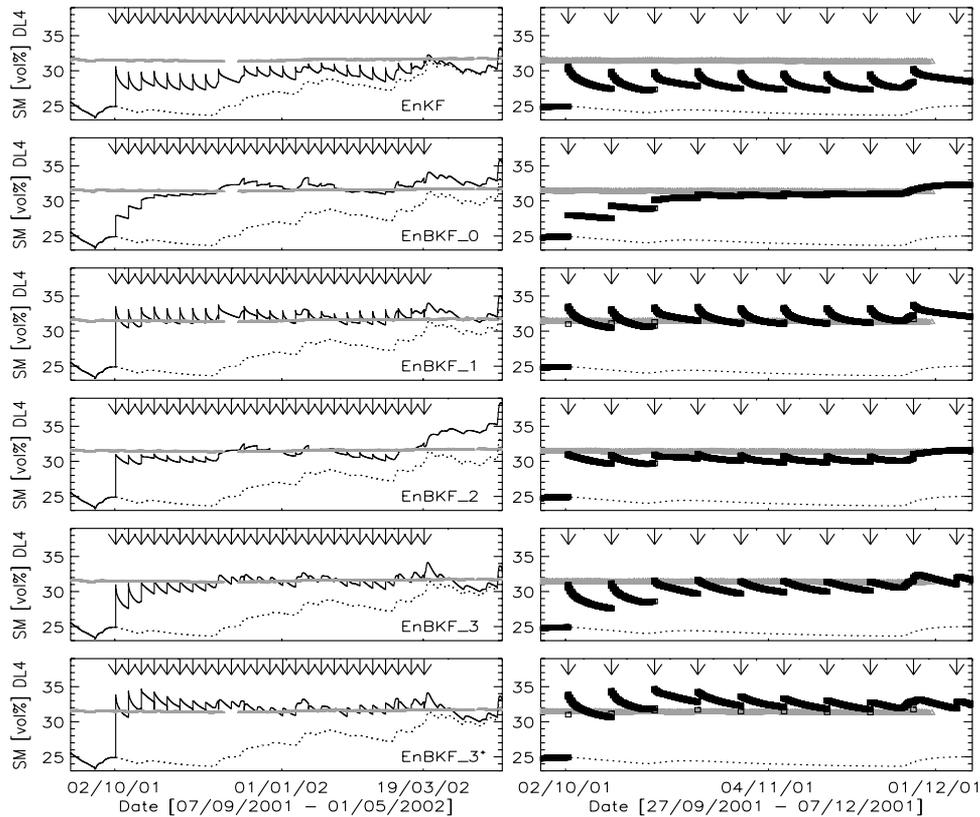


**Figure 2.** Soil moisture at 10 cm depth for DH2 for weekly assimilation (indicated by arrows) at 10 cm only and for the six different estimation algorithms. Right column plots are zooms of left column plots. Shown are (gray) observations, (dotted) ensemble mean without filtering, and (black solid line in the left panels and black squares in the right panels) filtering runs.

milation, which is referred to as the control model integration. In both examples the control model results are typically drier than the observations. In the first example (Figure 2), the 10 cm in situ data show a larger dynamic range than the corresponding control model estimates, and the bias varies on timescales of a few weeks. In the second example (Figure 3), the 80 cm control model soil moisture is persistently drier than the observations, although the bias gradually decreases toward the end of the experiment period.

[35] As expected, the assimilation integrations provide soil moisture estimates that are typically closer to the observations than the control model estimates. In most cases, however, the assimilation estimates drift back toward the control model estimates immediately after each update because of errors in the model formulation or parameterization. This behavior is most clearly visible in Figure 3, where the drift was most pronounced for EnKF and EnBKF\_1, and consequently this also affected the output for EnBKF\_1 and EnBKF\_3<sup>+</sup>. In these methods, the model was re-initialized with a state estimate that was forced to move from the model climatology toward the observation climatology.

[36] Figure 3 also shows that algorithms EnBKF\_1 and EnBKF\_3<sup>+</sup> produced overshoots, i.e., the output was found outside the range spanned by the observation and model results. More specifically, the analyses of algorithms EnBKF\_1 and EnBKF\_3<sup>+</sup> can be clearly distinguished as isolated squares in the zoomed plots of Figure 3 and were well estimated, while the estimation of the (output) forecasts after the assimilation events produced the overshoots. The problem of overshoots indicates that a persistent bias model might be appropriate from one analysis to the next, but not for correcting the forecasts immediately following the last analysis. During a limited period after the state update, the forecast bias is typically limited by the effect of the stage 1 filter and growing quickly, until the model forecasts reach the ‘stable’ model climatology and the persistent bias model is valid again. Instead of the persistence model, a zero-initialized growing bias model that reaches the last updated  $\hat{\mathbf{b}}_i^+$  at the next assimilation step could perhaps be used. Overshoots were not observed for the output of EnBKF\_2. However, the underlying model run (without addition of the forecast bias estimate) showed that the model state diverged from the desired (model) trajectory. The algorithm with bias estimation only (EnBKF\_0) did not suffer from this prob-



**Figure 3.** Same as Figure 2, but at 80 cm depth for DL4.

lem and was able to find a good estimate of the bias, so that the output state came very close to the observations.

[37] While Figures 2 and 3 illustrate specific issues, we now turn to profile-integrated and area-averaged performance measures in Figure 4, namely the root mean square error ( $RMSE$ ) and the time series correlation coefficient ( $R$ ) between observed and modeled/assimilated soil moisture. For each of the filtering methods, seven separate assimilation integrations were performed by assimilating in situ data from one of seven layers once per week (Table 2) while withholding the rest of the in situ observations. Next, the profile-average  $RMSE$  and  $R$  were computed for each entire probe (with equal weights for all observed layers) for all assimilation integrations during the validation period (2 October 2001 through 30 April 2002) and subsequently normalized with the corresponding values from the control integration (no filtering). Figure 4 then shows the spatial average normalized  $RMSE$  and  $R$  across all L-probes, plus/minus one (spatial) standard deviation.

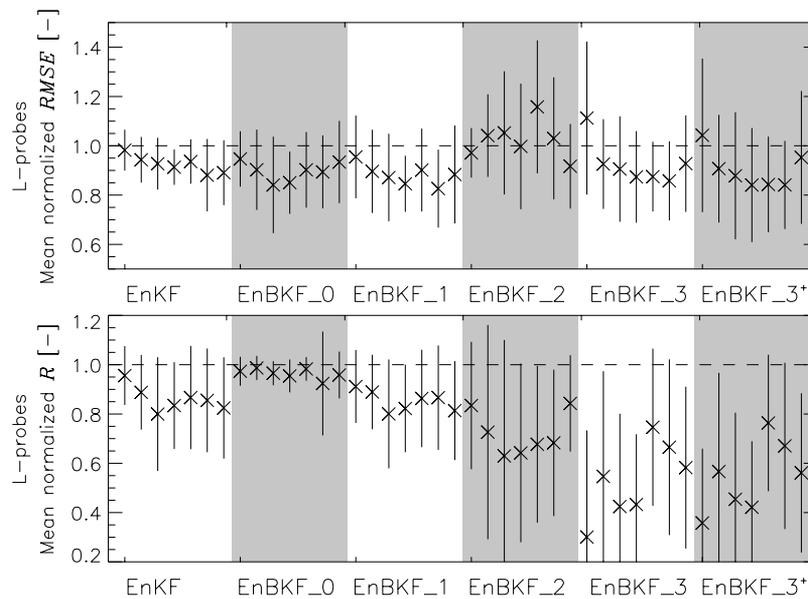
[38] Figure 4 shows that, in general, it was possible to improve (relative to the control) the  $RMSE$  of the complete profile slightly through assimilation in any single layer, both with and without bias estimation (except for EnBKF\_2). However, the improvement was modest and bias correction did not improve the results much more than conventional EnKF filtering without explicit bias estimation.

[39] The time series correlation  $R$  was typically degraded through filtering (Figure 4), because each assimilation event caused a jump in the time series, which decreased the similarity with the observed time series. EnKF, EnBKF\_0, and EnBKF\_1 retained acceptable correlation. These algo-

rithms have less impact on the model evolution, and the control model already provided a good estimate of the temporal evolution based on the high-quality input precipitation and other forcing data. By contrast, EnBKF\_3/3<sup>+</sup> greatly reduced  $R$ , because the complete bias correction was fed back in all profile layers, thereby causing more significant jumps in the time series.

[40] In general, assimilation in a single layer had only a limited impact on the complete profile. This may be because the state and bias error covariances ( $\mathbf{P}_{x,i}^-$  and  $\mathbf{P}_{b,i}^-$ ) were poorly known and error corrections were therefore inadequately propagated through the soil moisture profile. Moreover, the observability matrix of the bias system did not reach full rank when only one layer was observed, because the persistence model for the bias implies that there is no vertical “flow” of bias.

[41] Figure 5 shows the spatial average of the normalized  $RMSE$  and  $R$  for two individual soil layers into which observations were assimilated at different frequencies. The spatial average is over sensors with observations at the assimilation depth (i.e., all 36 probes at 10 cm, and 24 H- and L-probes at 80 cm, because M-probes have no observations at 80 cm depth). Obviously, the performance for the assimilation layer itself was greatly improved through inclusion of bias estimation, even though the profile-average performance improvement was modest (Figure 4). For EnBKF\_0, EnBKF\_1, EnBKF\_2 and EnBKF\_3<sup>+</sup>, the simulations at all sensor depths yielded better results than the control integration (without filtering) and the EnKF. By contrast, EnBKF\_3 was no better than EnKF. For EnBKF\_2, it was found that the underlying model run



**Figure 4.** Profile-integrated, area-average, normalized (top)  $RMSE$  and (bottom)  $R$  for different algorithms with weekly assimilation at 7 different depths (10, 30, 50, 80, 120, 150, and 180 cm) per algorithm. Normalization of  $RMSE$  and  $R$  is with respect to the control run. Averages were calculated over all layers and all 12 L-probes. One spatial standard deviation is also shown.

(without bias addition) sometimes introduced additional bias (not shown). Of all methods, simple bias estimation without state estimation (EnBKF\_0) was most efficient: it reduced the  $RMSE$  and kept the time series correlation  $R$  highest. For algorithms with bias estimation, the performance measures did not generally improve when the assimilation was done more frequently than once per week and less intensive assimilation is needed than for EnKF.

#### 4.2. Other Variables

[42] Unfortunately, in situ observations were not available for evapotranspiration, runoff, and drainage. Nevertheless, it is instructive to assess the impact of the various assimilation and bias feedback options on these dependent variables. Table 3 shows soil moisture, evapotranspiration (and its component fluxes), runoff, and drainage (averaged over the profile locations) for 2-weekly assimilation of complete observed profiles. The table also lists the profile-integrated  $RMSE$  between observed and modeled/assimilated soil moisture, the total column absolute increment added to the soil water, and spatial standard deviations for all variables that provide a measure of the variability across all locations. Note that flux totals given in Table 3 are for the 211-day validation period from 2 October 2001 through 30 April 2002 that includes fall, winter, and early spring with relatively little evapotranspiration. For reference, the total observed precipitation during the validation period is 279 mm (of which 25 mm fell as snow), and the total saturated water content of the control model, averaged across all profiles, is  $869 \pm 90$  mm.

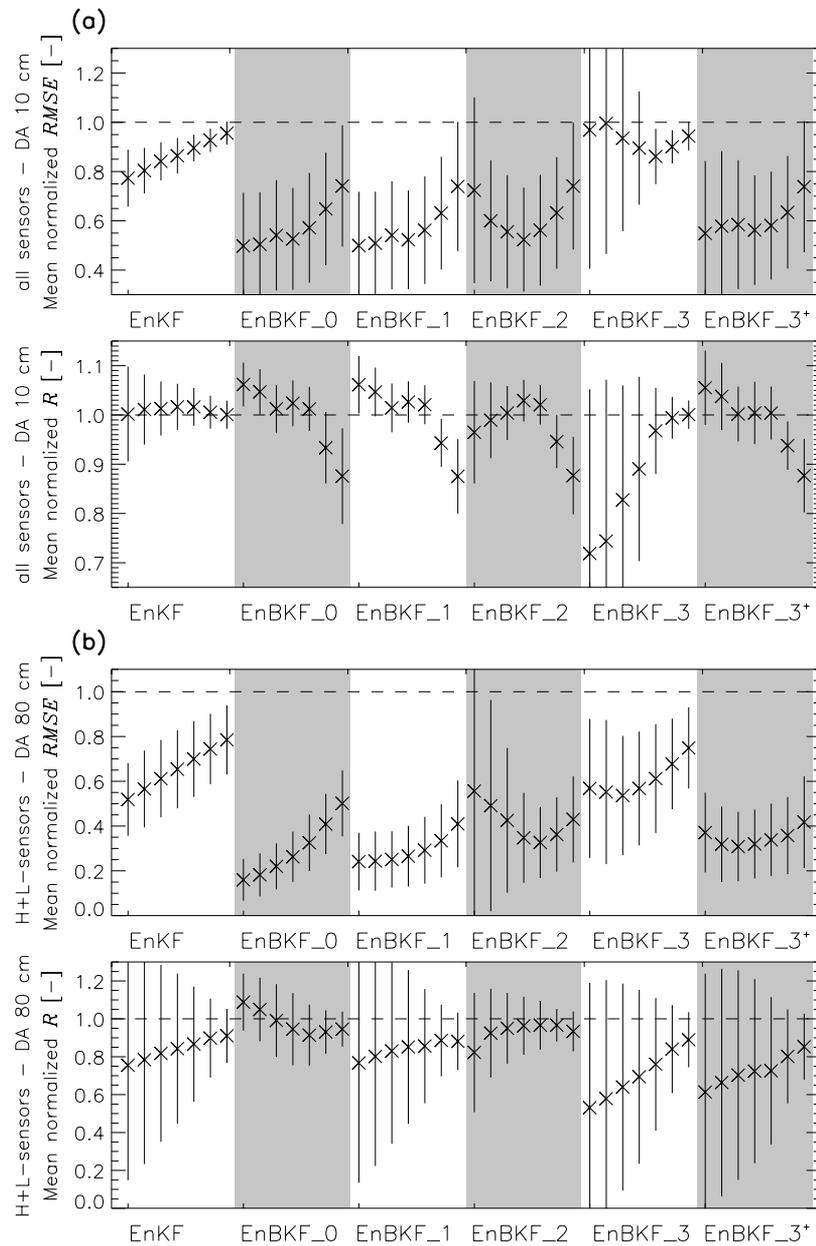
[43] Table 3 indicates that the average soil moisture is typically increased through data assimilation, from 16.5 vol% for the control run to between 18 and 20 vol%. Algorithms that include bias correction in the outputs produce the highest soil moisture estimates. For all assimilation algorithms, the total evapotranspiration, its component fluxes,

and surface runoff were only marginally affected when compared to the control run (Table 3). This is a logical consequence of the fact that the complete profile assimilation had only a minor effect on the upper layers of soil moisture and because of the winter time. Deeper layers were influenced more strongly [De Lannoy et al., 2007]. Consequently, the choice of assimilation algorithm had a large impact on subsurface drainage.

[44] The control integration and EnBKF\_0 suggest that 109 mm of water (39% of observed precipitation) left the study site as subsurface drainage. This estimate is roughly in line with an independent estimate that at least 30% of precipitation exits the site as drainage (Timothy Gish, personal communication). Significantly increased drainage was found for the rest of the assimilation algorithms. EnBKF\_2 suggests that 149 mm of water (53% of observed precipitation) drains from the site. EnKF, EnBKF\_1, and EnBKF\_3/3<sup>+</sup> produce drainage far in excess of observed precipitation. Such large drainage cannot be explained by errors in the observed precipitation and are entirely unrealistic.

[45] The unrealistic drainage rates follow directly from the fact that the calibrated model is biased dry and already produces relatively high subsurface drainage. When the data assimilation adds water to bring the biased soil moisture up to the observed values, the model responds with increased drainage. Note that the corresponding dry-out can also be seen in Figure 3 for assimilation in a single layer. The increased drainage acts to restore the bias, which in turn necessitates additional increments of water during the assimilation. In general, depending on the level of bias, the bias correction scheme that is used, and due to nonlinear effects, drainage produced by the assimilation easily departs from realistic values.

[46] The average absolute assimilation increments, also listed in Table 3, provide additional insights. By construc-



**Figure 5.** Normalized single layer  $RMSE$  and  $R$  for different algorithms with 7 different assimilation frequencies (a, b, c, d, e, f, and g; Table 2) per algorithm. Assimilation depth was (a) 10 cm and (b) 80 cm. Normalization of  $RMSE$  and  $R$  is with respect to the control run. Averages were calculated over all probes with sensors at the assimilation depth. One spatial standard deviation is also shown.

tion, data assimilation increments become part of the water balance, but it is difficult, if not impossible, to attribute the increments to specific fluxes. A desirable feature of a data assimilation system is to produce good soil moisture estimates (that is, low  $RMSE$ ) with the smallest possible increments. For methods without output bias correction (EnKF and EnBKF\_3) increments occurred at only 13 analysis times during the study period, with average absolute increments between 7 mm and 38 mm per event. By contrast, EnBKF\_0, EnBKF\_1, EnBKF\_2, and EnBKF\_3<sup>+</sup> use a total of  $211 \times 24 = 5064$  hourly increments for output bias correction during the 211-day experiment period, each averaging between 36 mm and 65 mm depending on the algorithm. EnBKF\_1, for example, thus requires a total

absolute increment of  $5064 \times 39 \text{ mm} \approx 197 \text{ m}$  of water during the experiment period!

[47] Table 3 clearly shows the tradeoff between good soil moisture estimates and small increments. Of all filtering methods, the regular EnKF caused the smallest increments, but shows little improvement in soil moisture over the control run and produced the largest soil moisture  $RMSE$  of all assimilation integrations. State updating with bias removal from the innovations (EnBKF\_2) resulted in the smallest  $RMSE$ , but at the expense of high increments, because the underlying model run with assimilation generated additional bias. Bias correction without state updating (EnBKF\_0) required the largest increments and most distorted the water balance, but delivered low  $RMSE$ . Gene-

**Table 3.** Area-Average States and Total Fluxes (2 October 2001 Through 30 April 2002) for 2-Weekly Assimilation of Complete Observed Profiles, Plus/Minus One (Spatial) Standard Deviation<sup>a</sup>

|   | Units      | EnCtrl<br>EnBKF_0          | EnKF<br>EnBKF_1            | EnBKF_2          | EnBKF_3<br>EnBKF_3 <sup>+</sup> |
|---|------------|----------------------------|----------------------------|------------------|---------------------------------|
| Profile average soil moisture           | [vol%]     | 16.5 ± 2.8<br>19.2 ± 5.1   | 17.9 ± 3.6<br>19.7 ± 5.3   | —<br>19.2 ± 4.8  | 18.2 ± 3.8<br>20.0 ± 5.4        |
| Evapotranspiration                      | [mm]       | 67 ± 11                    | 67 ± 11                    | 67 ± 11          | 67 ± 11                         |
| Ground evaporation                      | [mm]       | 42 ± 10                    | 42 ± 10                    | 42 ± 10          | 42 ± 10                         |
| Canopy evaporation                      | [mm]       | 0.2 ± 0.7                  | 0.1 ± 0.7                  | 0.1 ± 0.7        | 0.1 ± 0.7                       |
| Canopy transpiration                    | [mm]       | 25 ± 10                    | 25 ± 10                    | 25 ± 10          | 25 ± 10                         |
| Surface runoff                          | [mm]       | 38 ± 11                    | 39 ± 11                    | 37 ± 11          | 39 ± 11                         |
| Subsurface drainage                     | [mm]       | 109 ± 32                   | 427 ± 393                  | 149 ± 60         | 571 ± 526                       |
| Soil moisture <i>RMSE</i>               | [vol%]     | 6.74 ± 3.47<br>2.00 ± 0.70 | 4.98 ± 2.74<br>1.94 ± 0.71 | —<br>1.88 ± 0.70 | 4.82 ± 2.73<br>2.08 ± 0.75      |
| Average absolute increment <sup>b</sup> | [mm/event] | n/a<br>65 ± 58             | 26 ± 29<br>39 ± 40         | —<br>51 ± 50     | 38 ± 39<br>36 ± 41              |

<sup>a</sup>Per column, the two listed algorithms yield the same output for all variables except for soil moisture and for the increment. EnCtrl stands for the control, i.e., the ensemble mean run without filtering.

<sup>b</sup>For EnKF and EnBKF\_3 there are 13 assimilation events. For EnBKF\_0, EnBKF\_1, EnBKF\_2, and EnBKF\_3<sup>+</sup> there are 24\*211 = 5064 hourly increments (events) over the 211-day period.

rally, methods with bias correction outside the model result in less biased output (low *RMSE*), but at the expense of large increments and a corresponding distortion in the water budget.

### 4.3. Forecast Bias Evolution

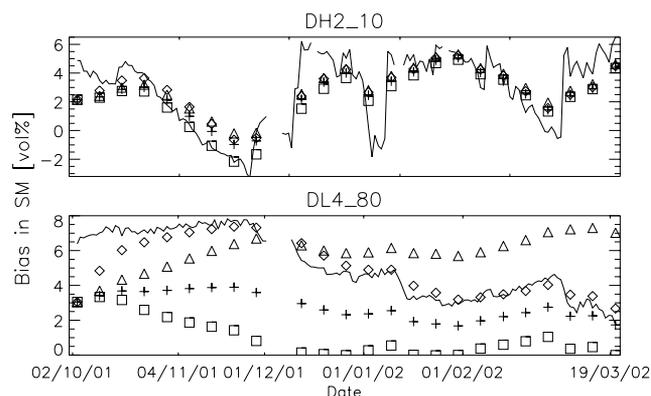
[48] Figure 6 shows the temporal evolution of the forecast bias estimates for two representative locations at two depths (10 cm and 80 cm) with high and low temporal variability in the bias (compare to Figures 2 and 3). For reference, Figure 6 also shows the difference between the observations and the control model integration. It is important to note that this difference is distinct from the forecast bias that we are trying to estimate and that the bias estimates should not necessarily match the observations minus control difference. While the different algorithms clearly yielded distinct values for the bias estimates, the temporal evolution was quite similar across the algorithms. In the case of bias estimation only (EnBKF\_0), the bias estimates were typically larger than for EnBKF\_1, because in EnBKF\_0 no random error was assumed and all error was attributed to bias. The estimated bias for the case of bias correction in the innovations only (EnBKF\_2) was larger than for other algorithms and growing in time for some locations, mainly at 80 cm. In the case of bias correction with complete feedback (EnBKF\_3/3<sup>+</sup>), the estimated forecast bias at 80 cm became gradually less than for the other algorithms as a consequence of the correction feedback. This desirable behavior is not seen at 10 cm, because the model returns to its biased climatology after the update with bias correction, as shown by the spikes for EnBKF\_3 at location DH2 in Figure 2. As will be shown in the next section, the imposed perturbations resulted in an a priori state error covariance which was relatively large. When  $\hat{\mathbf{P}}_{x,i}^-$  and hence  $\mathbf{P}_{b,i}^-$  were reduced in additional sensitivity experiments (not shown), the bias was identified more slowly, and less of the discrepancy between observations and predictions was assigned to forecast bias.

### 4.4. Consistency of Filter Operations

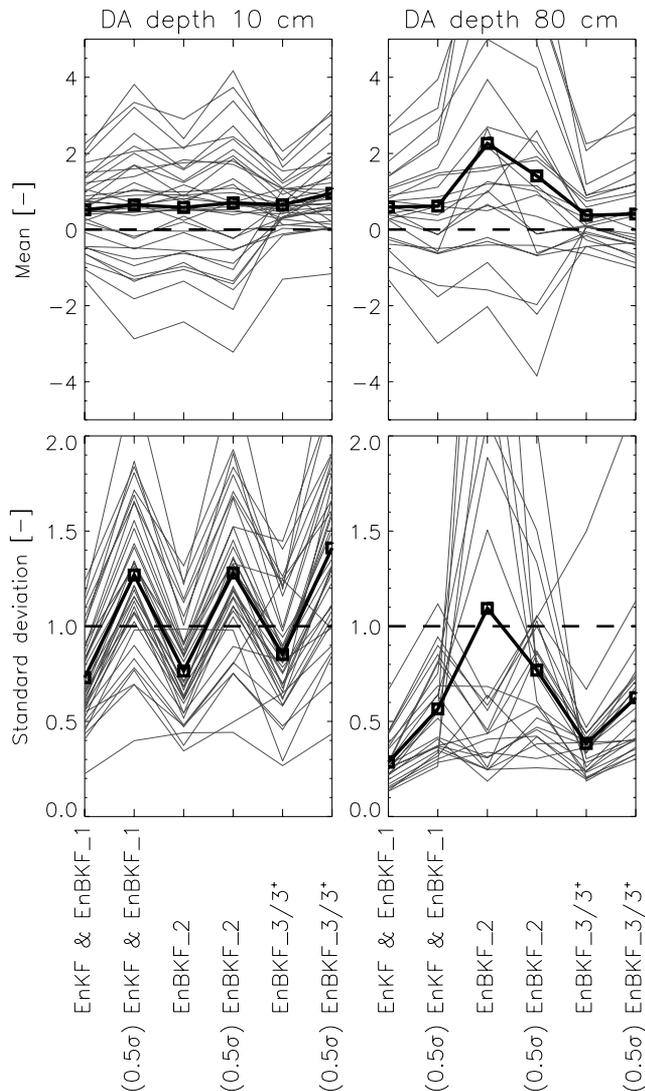
[49] Innovations permit the indirect validation of the stage 1 and stage 2 filters. If the model is linear and the filter operates in accordance with its underlying assumptions, appropriately normalized innovations should obey a

standard-normal distribution (Gaussian with mean zero and variance one), indicating that the filter parameters (model and observation error parameters) are appropriate. Because of the non-linearity of the land surface model and because of the approximate (scalar) normalization of the innovations (see below), we do not expect that the normalized innovations follow a standard-normal distribution perfectly even for very good filter parameters, i.e.,  $\hat{\mathbf{P}}_{x,i}^-$ ,  $\mathbf{P}_{b,i}^-$  and  $\mathbf{R}_i$ . Nevertheless, closeness of the innovations distribution to the standard-normal distribution can be considered a sign of consistent filter operation.

[50] For our analysis, we normalize each stage 1 (for bias-blind state estimation) ensemble mean innovation  $[\mathbf{y}_i - \mathbf{H}_i \hat{\mathbf{x}}_i^-]_k$  (corresponding to a scalar observation at a given data assimilation step) by the square-root of its filter-estimated standard deviation  $[\mathbf{H}_i \hat{\mathbf{P}}_{x,i}^- \mathbf{H}_i^T + \mathbf{R}_i]_{kk}^{1/2}$ . Figure 7 shows for all available observation points in space (one per line) at 10 cm and 80 cm depth the time mean and standard deviations of the normalized ensemble mean innovations for weekly assimilation of observations in these individual layers. Generally, we found that the mean of the normalized stage 1 innovations across all locations was close



**Figure 6.** Estimated forecast bias for weekly assimilation in a single layer at DH2 (10 cm) and DL4 (80 cm) for (◇) EnBKF\_0, (+) EnBKF\_1, (△) EnBKF\_2, and (□) EnBKF\_3/3<sup>+</sup>. The solid line shows the difference between the observations and the control model integration.



**Figure 7.** (Upper row) mean and (lower row) standard deviation of normalized stage 1 innovations time series at (left) all 36 and (right) 24 H- and L-probe locations (one line per location) for weekly assimilation in a single layer at (left) 10 cm and (right) 80 cm for the different algorithms with the full and halved ( $0.5\sigma$ ) ensemble perturbations. The thick line shows the average across all locations.

to zero. For individual locations, however, the mean was very different from zero, in particular for the filters that feed back bias-blind estimates into the model (EnKF, EnBKF\_1, EnBKF\_2). The area-average spread of the normalized stage 1 innovations at 10 cm depth was less than 1 for all types of filtering when averaged across all locations, indicating that the ensemble spread was larger than necessary. At 80 cm, EnBKF\_2 resulted in a normalized innovation spread larger than 1, indicating that the state's uncertainty was underestimated. This means that the filter is overly confident of its estimates, less responsive to observations, and at risk of diverging from the truth. This behavior of EnBKF\_2 was more prominent with more frequent assimilation and mainly for deeper assimilation layers. In these cases, the elements in  $\hat{\mathbf{P}}_{x,i}^-$  became very small because a number of ensemble members hit limiting values, thereby

causing a very limited state and bias update. The limiting values were reached because the repeated (slightly overestimated) bias removal pushed the states toward unrealistic values. This problem could be reduced by re-initializing the bias to zero before each new assimilation update.

[51] Since overestimation of the spread implies underestimation of the information content of the assimilation estimates, we also tried halving the magnitude of the ensemble perturbations for all parameters, initial states and forcings. Results for the normalized innovations from these experiments are also shown in Figure 7. For assimilation at 10 cm depth, the now reduced  $\hat{\mathbf{P}}_x^-$  underestimated the state estimate's uncertainty at most locations (at this depth) and unduly increased the risk of filter divergence in individual locations. At 80 cm depth, however, only the state estimates at a few locations showed too little spread. Paradoxically, the spread of the normalized innovations for EnBKF\_2 at 80 cm was smaller for the smaller perturbations. This is because the reduced spread in the a priori state estimates implied smaller bias updates and fewer instances when limiting values were reached, that is the smaller perturbations effectively resulted in larger a priori sample error covariances and thus in smaller spread of the normalized innovations. This finding serves as a warning that it is often difficult to use data assimilation methods with non-linear models.

[52] The innovations analysis was also performed for the stage 2 filter (for bias estimation). In this case, the ensemble mean a priori bias-corrected innovations  $[\mathbf{y}_i - \mathbf{H}_i(\hat{\mathbf{x}}_i^- + \hat{\mathbf{b}}_i^-)]_k$  for all scalar innovations were normalized by their filter-estimated standard deviations  $[\frac{1}{1-\gamma}\mathbf{H}\hat{\mathbf{P}}_{x,i}^-\mathbf{H}^T + \mathbf{R}_i]_{kk}^{1/2}$ . We found that  $\gamma = 0.5$  was larger than necessary, even when the magnitude of the perturbations was cut in half (not shown). This implies that the uncertainty of the bias estimates is considerably smaller than that of the state estimates, and that the resulting bias-corrected state estimate is less uncertain than the filter (with  $\gamma = 0.5$ ) estimates it is. A smaller  $\gamma$  should thus improve overall filter performance and yield increased information content for the bias-corrected state estimates. It was found that for higher assimilation frequencies,  $\hat{\mathbf{P}}_{b,i}^-$  could be reduced further.

## 5. Summary and Conclusions

[53] We tested different practical algorithms for state and forecast bias estimation for real soil moisture assimilation in a small agricultural field and investigated feedback of bias-blind or bias-corrected state estimates into a biased model. For assimilation in a single layer, the soil moisture estimates for the assimilation layer could be greatly improved through inclusion of explicit bias estimation (Figure 5), even though the profile-average performance improvement was modest (Figure 4). For assimilation of complete observed profiles, significant improvements in soil moisture estimates were achieved through explicit bias estimation (Table 3).

[54] Some but not all of the forecast bias is implicitly removed in the conventional EnKF. Better results for soil moisture were achieved with the original separate state and bias estimation algorithm (EnBKF\_1) but this method produced undesirable overshoots in the output. The overshoots are caused by the use of a persistence model for the bias and may be ameliorated with a more sophisticated bias model. When the bias was removed from the innovations in

the state update equation (EnBKF\_2), the filter gradually overestimated the bias and diverged from the validating observations. This problem could be alleviated by re-initializing the bias to zero before updating. Complete feedback of the bias correction into the model (EnBKF\_3/3<sup>+</sup>) produced mixed results because in this case the model was driven away from its preferred soil moisture climatology. Post-processing the model results by the estimated bias without updating the state (EnBKF\_0) yielded the best results for soil moisture, because the major part of forecast error was systematic, rather than random.

[55] For EnBKF\_0, information from the observations never reaches model-produced dependent variables by construction. For all other assimilation algorithms, it is important to consider the impact of soil moisture assimilation on diagnostic processes such as drainage, runoff, and evapotranspiration. Because of the specifics of our experiment, evapotranspiration and surface runoff were almost unchanged by all assimilation algorithms (with respect to the control run). By design, EnBKF\_2 affected subsurface drainage only marginally, but all other methods (EnKF, EnBKF\_1, and EnBKF\_3/3<sup>+</sup>) showed a strong and negative impact on subsurface drainage.

[56] Problems with dependent fluxes are often linked to excessive assimilation increments and thus require a trade-off between the quality of the soil moisture estimates and that of the flux estimates. In a biased model, large and frequent increments are required to produce low errors in soil moisture estimates. The large increments, in turn, result in unrealistic fluxes as the model continually drifts back to its biased soil moisture climatology. This problem is particularly acute if bias-corrected soil moisture estimates are fed back into the model (EnBKF\_3/3<sup>+</sup>). Moreover, when bias corrections are applied at every output time step, huge total increments result.

[57] In summary, the best choice of algorithm for assimilation and bias correction depends on the quality of the model, the nature of the model bias and the intended use of the estimated model state variables or fluxes. If, like in our study, correct fluxes are produced only when the land model is operating within its biased soil moisture climatology, EnBKF\_0 appears to be the method of choice. Alternatively, EnBKF\_2 could be considered, but would require further tuning to prevent filter divergence. If the land model were to produce the correct fluxes when its soil moisture is close to the true soil moisture, EnBKF\_1 or EnBKF\_3<sup>+</sup> would be an obvious choice.

[58] Further filter tuning could improve the performance of the various assimilation methods and may also affect the relative performance among the algorithms. Much more research is also needed into the choice of the bias model, the estimation of the a priori bias error covariance matrix, and the need for unbiased observations. The undeniable presence of bias in land surface models requires ongoing development of advanced bias estimation and correction methods if land data assimilation systems are to succeed.

[59] **Acknowledgments.** We thank the USDA Beltsville Agricultural Research Center/Agricultural Research Service for their data and the NASA Hydrological Sciences Branch for hosting the first author during part of the research. This research is supported by PhD-fund of the Bijzonder Onderzoeksfonds of Ghent University. Wade Crow and the two other reviewers are thanked for their suggestions to improve this paper.

## References

- Baek, S., B. R. Hunt, E. Kalnay, E. Ott, and I. Szunyogh (2006), Local ensemble Kalman filtering in the presence of model error, *Tellus A*, 58, 293–306.
- Berg, A., J. Famiglietti, J. Walker, and P. R. Houser (2003), Impact of bias correction to reanalysis products on simulations of North American soil moisture and hydrological fluxes, *J. Geophys. Res.*, 108(D16), 4490, doi:10.1029/2002JD003334.
- Bosilovich, M. G., J. D. Radakovich, A. D. Silva, R. Todling, and F. Verter (2006), Skin temperature analysis and bias correction in a coupled land-atmosphere data assimilation system, *J. Meteorol. Soc. Jpn.*, conditionally accepted.
- Burgers, G., P. J. van Leeuwen, and G. Evensen (1998), On the analysis scheme in the ensemble Kalman filter, *Mon. Weather Rev.*, 126, 1719–1728.
- Chepurin, G. A., J. A. Carton, and D. Dee (2005), Forecast model bias correction in ocean data assimilation, *Mon. Weather Rev.*, 133, 1328–1342.
- 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.
- Crow, W. T., R. D. Koster, R. H. Reichle, and H. O. Sharif (2005), Relevance of time-varying and time-invariant retrieval error sources on the utility of spaceborne soil moisture products, *Geophys. Res. Lett.*, 32, L24405, doi:10.1029/2005GL024889.
- Dai, Y., X. Zeng, R. Dickinson, I. Baker, G. B. Bonan, M. G. Bosilovich, A. S. Denning, P. A. Dimeyer, P. R. Houser, G. Niu, K. W. Oleson, C. A. Schlosser, and Z.-L. Yang (2003), The Common Land Model, *Bull. Am. Meteorol. Soc.*, 84, 1013–1023.
- De Lannoy, G. J. M., P. R. Houser, V. R. N. Pauwels, and N. E. C. Verhoest (2006a), Assessment of model uncertainty for soil moisture through ensemble verification, *J. Geophys. Res.*, 111, D10101, doi:10.1029/2005JD006367.
- De Lannoy, G. J. M., N. E. C. Verhoest, P. R. Houser, T. Gish, and M. Van Meirvenne (2006b), Spatial and temporal characteristics of soil moisture in an intensively monitored agricultural field (OPE<sup>3</sup>), *J. Hydrol.*, 331, 719–730, doi:10.1016/j.jhydrol.2006.06.01.
- De Lannoy, G. J. M., P. R. Houser, V. R. N. Pauwels, and N. E. C. Verhoest (2007), State and bias estimation for soil moisture profiles by an ensemble Kalman filter: Effect of assimilation depth and frequency, *Water Resour. Res.*, 43, W06401, doi:10.1029/2006WR005100.
- Dee, D. P., and A. M. da Silva (1998), Data assimilation in the presence of forecast bias, *Q. J. R. Meteorol. Soc.*, 124, 269–295.
- Dee, D. P., and R. Todling (2000), Data assimilation in the presence of forecast bias: The GEOS moisture analysis, *Mon. Weather Rev.*, 128, 3268–3282.
- Drécourt, J.-P., and H. Madsen (2002), Uncertainty estimation in groundwater modelling using Kalman filtering, in *ModelCARE 2002, Proceedings of the 4th International Conference on Calibration and Reliability in Groundwater Modelling*, edited by K. Kovar and Z. Hrkal, vol. 46 (2/3), pp. 306–309, Acta Universitatis Carolinae, Prague, Czech Republic.
- Drécourt, J.-P., H. Madsen, and D. Rosbjerg (2006), Bias aware Kalman filters: Comparison and improvements, *Adv. Water Resour.*, 29, 707–718.
- Drusch, M., E. F. Wood, and H. Gao (2005), Observation operators for the direct assimilation of TRMM microwave imager retrieved soil moisture, *Geophys. Res. Lett.*, 32, L15403, doi:10.1029/2005GL023623.
- Evensen, G. (2003), The ensemble Kalman filter: Theoretical formulation and practical implementation, *Ocean Dynamics*, 53, 343–367.
- Friedland, B. (1969), Treatment of bias in recursive filtering, *IEEE Trans. Autom. Control*, AC-14, 359–367.
- Gish, T., W. Dulaney, K.-J. S. Kung, C. Daughtry, J. Doolittle, and P. Miller (2002), Evaluating use of ground-penetrating radar for identifying subsurface flow pathways, *Soil Sci. Soc. Am. J.*, 66, 1620–1629.
- Griffith, A. K., and N. K. Nichols (2000), Adjoint techniques in data assimilation for estimating model error, *J. Flow, Turbulence and Combustion*, 65, 469–488.
- Hamill, T. M., J. Whitaker, and C. Snyder (2001), Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter, *Mon. Weather Rev.*, 129, 2776–2790.
- 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 hydrologic modeling using data assimilation, *Water Resour. Res.*, *34*, 3405–3420.
- Jazwinski, A. H. (1970), *Stochastic processes and filtering theory*, vol. 64 of *Mathematics in science and engineering*, Academic Press.
- Keppenne, L. C., M. M. Rienecker, N. P. Kurkowski, and D. A. Adamec (2005), Ensemble Kalman filter assimilation of temperature and altimeter data with bias correction and application to seasonal prediction, *Non-linear Process. Geophys.*, *12*, 491–503.
- Lorenc, A., and O. Hammon (1988), Objective quality control of observations using Bayesian methods: Theory, and a practical implementation, *Q. J. R. Meteorol. Soc.*, *114*, 515–543.
- Madsen, H., and C. Skotner (2005), Adaptive state updating in real-time river flow forecasting - A combined filtering and error forecasting procedure, *J. Hydrol.*, *308*, 302–312.
- Margulis, S. A., D. McLaughlin, D. Entekhabi, and S. Dunne (2002), Land data assimilation of soil moisture using measurements from the Southern Great Plains 1997 Field Experiment, *Water Resour. Res.*, *38*(12), 1299, doi:10.1029/2001WR001114.
- Ni-Meister, W., J. Walker, and P. R. Houser (2005), Soil moisture initialization for climate prediction: Characterization of model and observation errors, *J. Geophys. Res.*, *110*, D13111, doi:10.1029/2004JD005745.
- Pauwels, V. R. N., and G. J. M. De Lannoy (2006), Improvement of modeled soil wetness conditions and turbulent fluxes through the assimilation of observed discharge, *J. Hydrometeorol.*, *7*, 458–477.
- Pauwels, V. R. N., N. E. C. Verhoest, G. J. M. De Lannoy, P. Defourny, V. Guissard, and C. Lucau (2007), Optimization of a coupled hydrology/crop growth model through the assimilation of observed soil moisture and LAI values using an Ensemble Kalman filter, *Water Resour. Res.*, *43*, W04421, doi:10.1029/2006WR004942.
- Reichle, R. H., and R. Koster (2004), Bias reduction in short records of satellite soil moisture, *Geophys. Res. Lett.*, *31*, L19501, doi:10.1029/2004GL020938.
- Reichle, R. H., and R. Koster (2005), Global assimilation of satellite surface soil moisture retrievals into the NASA Catchment land surface model, *Geophys. Res. Lett.*, *32*, L02404, doi:10.1029/2004GL021700.
- Reichle, R. H., D. Entekhabi, and D. B. McLaughlin (2001), Downscaling of radio brightness measurements for soil moisture estimation: A four dimensional variational data assimilation approach, *Water Resour. Res.*, *37*, 2353–2364.
- Reichle, R. H., D. B. McLaughlin, and D. Entekhabi (2002), Hydrologic data assimilation with the Ensemble Kalman filter, *Mon. Weather Rev.*, *120*, 103–114.
- Reichle, R. H., R. Koster, J. Dong, and A. Berg (2004), Global soil moisture from satellite observations, land surface models, and ground data: Implications for data assimilation, *J. Hydrometeorol.*, *5*, 430–442.
- Toth, Z., and M. Peña (2006), Data assimilation and numerical forecasting with imperfect models: the mapping paradigm, *Physica D*, submitted.
- Vidard, P., A. Piacentini, and F.-X. Le Dimet (2004), Variational data analysis with control of the forecast bias, *Tellus*, *56*, 177–188.
- Walker, J. P., and P. R. Houser (2004), Requirements of a global near-surface soil moisture satellite mission: Accuracy, repeat time, and spatial resolution, *Adv. Water Resour.*, *27*, 785–801.
- Walker, J. P., G. R. Willgoose, and J. D. Kalma (2001), One-dimensional soil moisture profile retrieval by assimilation of near-surface measurements: A simplified soil moisture model and field application, *J. Hydrometeorol.*, *2*, 356–373.
- Walker, J. P., P. R. Houser, and R. Reichle (2003), New technologies require advances in hydrologic data assimilation, *EOS*, *84*, 545–551.
- Zupanski, D., and M. Zupanski (2006), Model error estimation employing ensemble data assimilation approach, *Mon. Weather Rev.*, *134*, 1337–1354.

---

G. J. M. De Lannoy, V. R. N. Pauwels, and N. E. C. Verhoest, Laboratory of Hydrology and Water Management, Ghent University, Coupure links 653, B-9000 Ghent, Belgium. (gabrielle.delannoy@ugent.be)

P. R. Houser, George Mason University and Center for Research on Environment and Water, 4041 Powder Mill Road, Suite 302, Calverton, MD 20705-3106, USA.

R. H. Reichle, Global Modeling and Assimilation Office, NASA Goddard Space Flight Center, (Code 610.1) Greenbelt, MD 20771, USA.