Hydrol. Earth Syst. Sci. Discuss., 8, 5849–5890, 2011
www.hydrol-earth-syst-sci-discuss.net/8/5849/2011/
doi:10.5194/hessd-8-5849-2011

Hydrology and Earth System Sciences Discussions

The importance of parameter resampling for soil moisture data assimilation into hydrologic models using the particle filter

D. A. Plaza[1], R. De Keyser[2], G. J. M. De Lannoy[1], L. Giustarini[3], P. Matgen[3], and V. R. N. Pauwels[1]

[1]Laboratory of Hydrology and Water Management, Ghent University, Ghent, Belgium
[2]Department of Electrical energy, Systems and Automation, Ghent University, Ghent, Belgium
[3]Centre de Recherche Public – Gabriel Lippmann, Département Environnement et Agrobiotechnologies, Belvaux, Luxembourg

Received: 1 June 2011 – Accepted: 10 June 2011 – Published: 21 June 2011

Correspondence to: D. A. Plaza (douglasantonio.plazaguingla@ugent.be)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

The Ensemble Kalman filter (EnKF) and the Sequential Importance Resampling (SIR) particle filter are evaluated for their performance in soil moisture assimilation and the consequent effect on discharge. With respect to the resulting soil moisture time series, both filters perform similarly. However, both filters have a negative effect on the discharge due to inconsistency between the parameter values and the states after the assimilation. In order to overcome this inconsistency, parameter resampling is applied along with the SIR filter, to obtain consistent parameter values with the analyzed soil moisture state. Extreme parameter replication, which could lead to a particle collapse, is avoided by the perturbation of the parameters with white noise. Both the modelled soil moisture and discharge are improved if the complementary parameter resampling is applied. The SIR filter with parameter resampling offers an efficient way to deal with biased observations. The robustness of the methodology is evaluated for 3 model parameter sets and 3 assimilation frequencies.

Overall, the results in this paper indicate that the particle filter is a promising tool for hydrologic modelling purposes, but that an additional parameter resampling may be necessary to consistently update all state variables and fluxes within the model.

1 Introduction

It is widely recognized that hydrologic models are useful tools for a number of purposes, ranging from flood forecasting (Andersson, 1992) to numerical weather prediction and climate studies (Zhang et al., 2008). Due to uncertainties in the meteorological forcings and model parameters, and errors or oversimplifications in the model physics, these models are always prone to a certain level of uncertainty. One way to reduce the predictive uncertainty of hydrologic models is to regularly update these models using externally obtained data sets, which is commonly referred to as Data Assimilation (DA). The improvement of hydrologic model results through the assimilation of soil moisture

HESSD

8, 5849–5890, 2011

data has been the subject of numerous studies (Entekhabi et al., 1994; Walker et al., 2002; Pauwels et al., 2002; De Lannoy et al., 2007a). The underlying idea of data assimilation is to calculate a weighted average between the observations and the model results. The simplest way to perform this is to simply replace the model results by the observations, which is defined as direct insertion (Heathman et al., 2003). More advanced assimilation methods include nudging of the model results to the observations (Houser et al., 1998; Pauwels et al., 2001; Paniconi et al., 2002) and optimal interpolation (Seuffert et al., 2004). These techniques are in fact simplifications of the Kalman filter (Kalman, 1960), in which the model error is calculated explicitly throughout the simulation.

Originally developed for linear systems, and later extended for nonlinear systems, a great deal of attention has been paid to this assimilation method for hydrologic data assimilation. The extended Kalman filter, in which the forecast error covariance is calculated through a linearization of the model, and the ensemble Kalman filter, in which this model error covariance is calculated using the spread of an ensemble of model realizations, have been intercompared by Reichle et al. (2002). At this point, it can be argued that the ensemble Kalman filter is the most frequently used assimilation method in hydrology. A variation to this method is the ensemble Kalman smoother (Dunne and Entekhabi, 2005), in which observations that are distributed in time are used to update the model state variables. This method is comparable to variational assimilation (Caparrini et al., 2004), in which observations within a predefined window are used to estimate the initial state variables. One problem with the frequently used ensemble Kalman filter is the underlying assumption of Gaussianity of both the forecast and observation error structure. As it is evident that this assumption is not realistic for hydrologic systems, assimilation methods have been developed that relax this assumption.

One method that is receiving increasing attention in hydrology is the particle filter, which has been developed to function for any kind of model error (Liu and Chen, 1998). This method has been used to assimilate discharge records into relatively

HESSD

8, 5849–5890, 2011

simple rainfall-runoff models (Weerts and El Serafy, 2006; Moradkhani et al., 2005) and to assimilate water stage records into hydraulic models (Matgen et al., 2010; Giustarini et al., 2011). Recently, this method is also used for the assimilation of soil moisture data, for the estimation of model parameters (Montzka et al., 2011), and the estimation of root-zone soil moisture conditions (Nagarajan et al., 2010).

According to Moradkhani et al. (2005), Nagarajan et al. (2010), and Montzka et al. (2011), it is clear that the trend towards the application of particle filters is not limited to only the state estimation problem, but it can also be used for the identification of model parameter values, by exploiting the advantage of the flexible structure of the particle filter algorithms. In this study, state and parameter estimation are performed within the framework of the particle filter, aiming at an improvement of the model performance in terms of both soil moisture and discharge, through the assimilation of soil moisture data. The particle filter is used, because of its flexibility in the structure of the model and observational errors.

The organization of the paper is as follows: first, the study site and the description of the model are presented. The description of the experiment is presented. Then, the data assimilation methodologies are explained. The EnKF is used as the baseline methodology to which the Sequential Importance Resampling (SIR) particle filter is compared, after which the results from the study are explained. Finally, the conclusions from this study are summarized.

2 Site description

The area (Fig. 1) to be studied is located in the Grand Duchy of Luxembourg and includes the drainage area expanded from the head of the Alzette River basin, 4 km south of the French-Luxembourg border, to the stream gauge located in Pfaffenthal (Luxembourg City).

The discharge area covers a surface of 356 km^2 and consists of about 50 % cultivated land, 22 % urban centers and 28 % woodland. The topography of the floodplain is characterized by a natural sandstone bottleneck which is located near Luxembourg city. The valley located upstream of the bottleneck is up to 2.5 km wide, while in the Luxembourg sandstone the valley is only 75 m wide. The geological substratum is dominated by marls on the left bank and by limestone and sandstones deposits on the right bank. Sand and gravel, as well as marls and clay alternate in the alluvial deposits covering the stratum. A gauging station, operated since 1996, is located around the village of Livange providing accumulated precipitation amounts with a sampled frequency of 15 min. The meteorological station at Findel Airport is operated in the vicinity of the catchment.

3 Model description

The Community Land Model (CLM2.0) is the hydrologic model used in this study. CLM2.0 simulates land surface processes by calculating water and heat fluxes for each grid cell separately, without any interaction between cells. Each grid cell can be subdivided into several patches, containing one single land cover type such us urban, vegetated, wetlands, glacier and lake. The vegetated fraction is further subdivided into patches of plant functional types, which maintain their own prognostic variables. In this study, CLM2.0 was adapted in order to be able to use the individual patches as ensemble members according to De Lannoy et al. (2006a).

The meteorological forcings required by the model are the air temperature, wind speed, specific humidity, incoming solar radiation, and precipitation. The meteorological forcings were assumed to be spatially uniform over the complete study area. Vertical layers in CLM2.0 embody one vegetation layer, up to ten soil layers and up to 5 snow layers. In this application, soil layers depths were set to 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 cm.

In CLM2.0, each grid cell contains around 30 model parameters related to the different physical processes represented by the model. From these 30 parameters, only 10 parameters were found to be highly correlated with soil moisture and baseflow. The

HESSD
8, 5849–5890, 2011

reduction of the parameter set allows for the application of automatic calibration algorithms, such as the Shuffled complex evolution approach (Duan et al., 1993) which was used in this study. Table 1 presents the description of the selected parameters and three corresponding sets of optimal parameter values (set 1, set 2, set 3) which yield a similar good model performance. The optimal values were identified by minimizing the Root Mean Square Error (RMSE) between observed and simulated discharge during year 2006. The three parameter sets will be used to validate the data assimilation methodology for different model configurations and different synthetic observations in this work.

The model is applied using a constant hourly time step and the study area is represented by 4 grid cells at a 10 km × 10 km resolution which is consistent with the resolution of large scale models. For the sake of clarity in the presentation of the algorithm performances, results corresponding to the cell located in the lower left quadrant in Fig. 1 are presented.

4 Experiment setup

A synthetic soil moisture data assimilation study is performed to assess the performance of different filter algorithms (see below). Soil moisture assimilation has received a lot of attention during the last decades, but insight in the impact of soil moisture assimilation on dependent variables, for instance dicharge, has been limited (Pauwels et al., 2002; De Lannoy et al., 2007b; Brocca et al., 2010).

Synthetic volumetric soil moisture observations, corresponding to the top 10 cm soil layer, are generated with the CLM2.0 in a deterministic model run using parameter values which are different from the optimal values (parameter sets 1-2-3) identified by the automatic calibration algorithm. The use of parameter values that differ from the optimal is chosen to obtain synthetic observations which are unlikely to be part of the model ensemble. The way observations are generated introduces bias in the observations themselves. The rationale behind this is also to evaluate how accurately the filters can deal with this. Depending on the algorithm, either an ensemble of synthetic observations is generated (for the EnKF) or only a single realization (for the PF).

The forecast uncertainty is introduced through the generation of soil moisture random samples, which is referred to ensemble generation in the EnKF methodology (Eq. 8) and particles propagation in the PF methodology (see SIR algorithm description). For the sake of comparison of the methodologies, a unified criterium has been adopted which is the ensemble generation.

The meteorological forcings and the model parameters were disturbed with an additive zero mean white Gaussian noise in order to obtain the soil moisture ensemble. The standard deviation of this random number for the parameters was set to a predefined fraction of the parameter value. In order to check for the correctness of the ensemble, two different ensemble verification measures were used (De Lannoy et al., 2006a). The ensemble spread (ensp_t), the ensemble mean square error (mse_t), and the ensemble skill (ensk_t) have to be computed first and at each time step t:

$$
\begin{aligned}
\mathrm{ensp}_t &= \frac{1}{N}\sum_{i=1}^{N}(\hat{z}_{t,i}-\bar{\hat{z}}_t)^2 \\
\mathrm{mse}_t &= \frac{1}{N}\sum_{i=1}^{N}(\hat{z}_{t,i}-z_t)^2 \qquad (1)\\
\mathrm{ensk}_t &= (\bar{\hat{z}}_t - z_t)^2
\end{aligned}
$$

where $\hat{z}_t$ is the variable to be estimated and z_t is the corresponding observation of the estimated variable at time step t. In order to have a large enough ensemble spread, on average the ensemble mean differs from the observation by a value that is equal to the time average of the ensemble spread. Therefore, the following expression should be true:

$$\frac{<\mathrm{ensk}>}{<\mathrm{ensp}>} \approx 1 \qquad (2)$$

HESSD
8, 5849–5890, 2011

where $< . >$ indicates an average over the simulation period. Furthermore, if the truth is statistically indistinguishable from a member of the ensemble, the following expression should be true:

$$\frac{< \sqrt{\text{ensk}} >}{< \sqrt{\text{mse}} >} \approx \sqrt{\frac{N+1}{2N}} \tag{3}$$

An ensemble size of 64 members was used and the optimal disturbance fractions correspond to 0.10 for the model parameters, and 0.01 for the meteorological forcings. Figure 5 shows the soil moisture ensemble and the corresponding baseflow ensemble, the ratio $< \text{ensk} > / < \text{ensp} >$ is equal to 1.09 which approximates 1 and the ratio $< \sqrt{\text{ensk}} > / < \sqrt{\text{mse}} >$ is equal to 0.72 which approximates the value of $\sqrt{1/2}$ with the simulation period corresponding to year 2007.

The hydrologic assimilation study presented in this paper is part of a sequential assimilation study where CLM2.0 is coupled to a 1-D hydraulic model and a joint assimilation experiment is carried out. The link between the hydrologic and the hydraulic model is the discharge. Therefore, a discharge ensemble is generated in order to complete the model sequence. Figure 6 shows the discharge ensemble and the discharge observations obtained from a gauging station located in Pfaffenthal (Fig. 1) connecting the hydrologic study site (green patch) to the hydraulic study area (river reach between Pfaffenthal and Mersch). Preliminary results of the joint hydrodynamic study are reported in Matgen et al. (2010). The appropriateness of discharge ensemble spread for covering the discharge observations and for hydraulic data assimilation purposes is checked according to Eqs. (2) and (3) with values of 0.9619 and 0.7002, respectively, indicating a correct ensemble.

A robustness test of the assimilation algorithms will be performed by considering the impact of the data assimilation frequency and of different optimal parameter values for the model integration and the synthetic observation generation. Discussion on the filter performances for these scenarios will be extended in the results section.

HESSD
8, 5849–5890, 2011

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◂◂ | ▸▸
◂ | ▸
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

5 Assimilation algorithms

In nonlinear estimation, the dynamic system in discrete time is described by the state evolution equation given by:

$$\boldsymbol{x}_t = \boldsymbol{f}_t(\boldsymbol{x}_{t-1}, \boldsymbol{u}_{t-1}, \boldsymbol{v}_{t-1}) \tag{4}$$

where t is the discrete time index, $\boldsymbol{x}$ is the state vector, $\boldsymbol{f}_t(.)$ is the nonlinear function, $\boldsymbol{u}$ is the input vector and $\boldsymbol{v}$ is the process noise. In this study, the state vector consists of 22 variables for each vertical profile, i.e., canopy water storage, vegetation temperature and soil temperature and moisture at 10 levels, CLM2.0 represents the nonlinear function $\boldsymbol{f_t(.)}$ and $\boldsymbol{u_t}$ is the vector of meteorological forcings.

The state estimation is accomplished when the information from the measurement is assimilated into the model. The relationship between the measurements and the system states (the observation model) is given by:

$$\boldsymbol{y}_t = \mathbf{H}_t \boldsymbol{x}_t + \boldsymbol{n}_t \tag{5}$$

Equation (5) represents the observation model, where $\boldsymbol{y}$ is a vector which contains the measurements, $\mathbf{H}_t$ is a diagonal matrix containing values of 0 and 1 and $\boldsymbol{n}_t$ is the noise affecting the observations. In this study the observation model is linear, because the assimilated soil moisture observations will correspond directly to the soil moisture state variables.

In recursive Bayesian filtering the solution to the estimation problem consists of two steps: the prediction and correction steps. These steps are formulated as follows:

$$p(\boldsymbol{x}_t|\boldsymbol{y}_{1:t-1}) = \int p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}) p(\boldsymbol{x}_{t-1}|\boldsymbol{y}_{1:t-1}) d\boldsymbol{x}_{t-1} \tag{6}$$

$$p(\boldsymbol{x}_t|\boldsymbol{y}_{1:t}) = \frac{p(\boldsymbol{y}_t|\boldsymbol{x}_t) p(\boldsymbol{x}_t|\boldsymbol{y}_{1:t-1})}{\int p(\boldsymbol{y}_t|\boldsymbol{x}_t) p(\boldsymbol{x}_t|\boldsymbol{y}_{1:t-1}) d\boldsymbol{x}_t} \tag{7}$$

In the prediction step (Eq. 6), the prior probability density function (pdf) $p(\boldsymbol{x}_t|\boldsymbol{y}_{1:t-1})$ is obtained based on the fact that the transition pdf $p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})$ and the prior pdf at time

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |
HESSD
8, 5849–5890, 2011
Assimilation of soil moisture using the particle filter
D. A. Plaza et al.

step $t-1$ are known, whereas in the correction step (Eq. 7), the prior pdf is corrected using the information from the likelihood pdf $p(\boldsymbol{y}_t|\boldsymbol{x}_t)$ and the posterior pdf $p(\boldsymbol{x}_t|\boldsymbol{y}_{1:t})$ is derived. The analytical solution of Eqs. (6) and (7) is difficult to determine since the evaluation of the integrals might be intractable. Monte Carlo (MC) simulation is a powerful tool to approximate the Bayesian solution.

In this application, two MC-based methods are evaluated as data assimilation methods, the well-known Ensemble Kalman Filter (EnKF) and the Particle Filter (PF), which, as stated above, is increasingly receiving attention. Both methods aim to approximate the posterior pdf by a set of random samples, hereafter referred to as ensemble members or particles. In the EnKF, the posterior pdf and the likelihood pdf are considered to be Gaussian, and therefore characterized by the mean and covariance. The latter is approximated by the sample covariance. On the other hand, in the PF, the point-mass representation of the posterior pdf does not require a parameterization of the pdf. The latter implies the relaxation from the assumption of Gaussianity, allowing to extend the PF to nonlinear and nonGaussian applications.

5.1 Ensemble Kalman Filter

The state propagation represented by Eq. (4) can be extended for a probabilistic model governing the ensemble state evolution (transition pdf) by perturbing of all the possible contributions to the model error, i.e., model parameters, forcings and initial conditions according to:

$$\hat{\boldsymbol{x}}_{t,i}^{-} = \boldsymbol{f}_t(\hat{\boldsymbol{x}}_{t-1,i}^{+}, \boldsymbol{u}_{t-1}, \boldsymbol{v}_{t-1,i}) \tag{8}$$

with $\{\hat{\boldsymbol{x}}_{t,i}^{-}, i = 1,...,N\}$ the a priori (forecast) ensemble state vector, i the ensemble member index and N the size of the ensemble equal to 64 in this study. The best estimate of $\hat{\boldsymbol{x}}_{t,i}^{-}$ is given by the ensemble mean:

$$\bar{\hat{\boldsymbol{x}}}_t^{-} = \frac{1}{N}\sum_{i=1}^{N} \hat{\boldsymbol{x}}_{t,i}^{-} \tag{9}$$

HESSD
8, 5849–5890, 2011

and the ensemble state error matrix is defined by:

$$\mathbf{E}_t^- = [\hat{\boldsymbol{x}}_{t,1}^- - \bar{\hat{\boldsymbol{x}}}_t^-, \ldots, \hat{\boldsymbol{x}}_{t,N}^- - \bar{\hat{\boldsymbol{x}}}_t^-] \tag{10}$$

By means of the MC approach the a priori error covariance can be approximated by the sample error covariance as follows:

$$\hat{\mathbf{P}}_t^- = \frac{1}{N-1}\mathbf{E}_t^-(\mathbf{E}_t^-)^{\mathrm{T}} \tag{11}$$

As reported in Burgers et al. (1998), the observations $\boldsymbol{y}_t$ should be perturbed in order to assure sufficient spread according to:

$$\boldsymbol{y}_{t,i} = \boldsymbol{y}_t + \boldsymbol{n}_{t,i} \quad with \quad \{i = 1, \ldots, N\} \tag{12}$$

with $\boldsymbol{n}_{t,i}$ a white Gaussian noise characterized by a zero mean and a covariance $\mathbf{R}$. The matrix $\mathbf{R}$ should represent the uncertainty of the observations and is set to 0.0005 $\mathbf{I}$, with $\mathbf{I}$ the identity matrix and units in $(\mathrm{cm}^3\,\mathrm{cm}^{-3})^2$.

Although the methods are described with the general vector notation, it should be noted that in this study, the assimilated soil moisture observations are scalars, i.e. a single observation is assimilated to update the state vector.

For the correction step of the filter, the Kalman gain has to be computed. Here, the approximated Kalman gain $\hat{\mathbf{K}}_t$ is computed since the sample covariance is used in the calculation. $\hat{\mathbf{K}}_t$ is given by:

$$\hat{\mathbf{K}}_t = \hat{\mathbf{P}}_t^- \mathbf{H}_t^{\mathrm{T}} [\mathbf{H}_t \hat{\mathbf{P}}_t^- \mathbf{H}_t^{\mathrm{T}} + \mathbf{R}]^{-1} \tag{13}$$

Finally, the a posteriori (after correction) state ensemble is given by:

$$\hat{\boldsymbol{x}}_{t,i}^+ = \hat{\boldsymbol{x}}_{t,i}^- + \hat{\mathbf{K}}_t[\boldsymbol{y}_{t,i} - \mathbf{H}_t \hat{\boldsymbol{x}}_{t,i}^-], \tag{14}$$

HESSD
8, 5849–5890, 2011

5.2 Particle Filter

Particle filters are a set of algorithms which approximate the posterior pdf by a group of random samples. In more detail, the integrals are mapped to discrete sums:

$$p(\boldsymbol{x}_t|\boldsymbol{y}_{1:t}) \approx \hat{p}(\boldsymbol{x}_t|\boldsymbol{y}_{1:t}) = \frac{1}{N}\sum_{i=1}^{N}\delta(\boldsymbol{x}_t - \boldsymbol{x}_{t,i}) \tag{15}$$

where the particles $\{\boldsymbol{x}_{t,i}; i = 1 \ldots N\}$ should be sampled from the posterior pdf and δ is the Dirac measure. The Dirac measure is given by:

$$\delta_x(X) = \begin{cases} 0 & \text{if } x \notin X, \\ 1 & \text{if } x \in X. \end{cases} \tag{16}$$

where x is a possible element of set X.

At this point, drawing particles is unfeasible since the posterior pdf is unknown. Nevertheless, it is viable to draw particles from a known proposal pdf (also called importance pdf). This is the basis of the importance sampling principle. Sequential Importance Sampling (SIS) is the recursive version of the importance sampling MC method and the particle filters are based on the SIS approach.

5.2.1 Sequential Importance Sampling (SIS)

The SIS approach approximates the posterior pdf by a set of weighted particles as follows:

$$\hat{p}(\boldsymbol{x}_t|\boldsymbol{y}_{1:t}) = \sum_{i=1}^{N}\tilde{\boldsymbol{w}}_{t,i}\delta(\boldsymbol{x}_t - \boldsymbol{x}_{t,i}) \tag{17}$$

where $\tilde{\boldsymbol{w}}_{t,i}$ are the normalized importance weights associated to the particles which are drawn from the proposal pdf. Considering that the system state evolves according

HESSD
8, 5849–5890, 2011

Assimilation of soil moisture using the particle filter

D. A. Plaza et al.

to a Markov process, and applying recursion to the filtering problem, the recursive expression for the importance weights is given by:

$$\boldsymbol{w}_{t,i} = \boldsymbol{w}_{t-1,i} \cdot \frac{p(\boldsymbol{y}_t|\boldsymbol{x}_{t,i})p(\boldsymbol{x}_{t,i}|\boldsymbol{x}_{t-1,i})}{q(\boldsymbol{x}_{t,i}|\boldsymbol{x}_{0:t-1,i},\boldsymbol{y}_{1:t})} \tag{18}$$

The selection of the proposal pdf $q(.,\boldsymbol{y}_{1:t})$ is extremely important in the design stage of the SIS filter. The filter performance mainly depends on how well the proposal pdf approximates the posterior pdf. In Doucet et al. (2001), an optimal choice for the importance density function is proposed:

$$q(\boldsymbol{x}_t|\boldsymbol{x}_{0:t-1},\boldsymbol{y}_{1:t}) = p(\boldsymbol{x}_t|\boldsymbol{x}_{0:t-1},\boldsymbol{y}_{1:t}) \tag{19}$$

This pdf is optimal in the sense that it minimizes the variance of the importance weights, however, the application of equation 19 is complex from the implementation point of view. A common choice of the proposal is the transition prior function (Gordon et al., 1993; Kitagawa, 1996):

$$q(\boldsymbol{x}_{t,i}|\boldsymbol{x}_{0:t-1,i},\boldsymbol{y}_{1:t}) = p(\boldsymbol{x}_{t,i}|\boldsymbol{x}_{t-1,i}) \tag{20}$$

The choice of the transition prior as the proposal simplifies Equation 18 resulting in an expression where the importance weights depend on their past values and also on the likelihood pdf. In this application, the likelihood pdf is considered to be Gaussian. Thus, the particles are weighted according to:

$$p(\boldsymbol{y}_t|\hat{\boldsymbol{x}}^-_{t,i}) = \frac{\exp\left(-\frac{1}{2}(\boldsymbol{y}_t - \mathbf{H}_t\hat{\boldsymbol{x}}^-_{t,i})^{\mathbf{T}}\mathbf{R}^{-1}(\boldsymbol{y}_t - \mathbf{H}_t\hat{\boldsymbol{x}}^-_{t,i})\right)}{(2\pi)^{m/2}|\mathbf{R}|^{1/2}} \tag{21}$$

where $\mathbf{R}$ is again the measurement error covariance matrix, which is set to $0.0005\mathbf{I}$, $|\mathbf{R}|$ is the determinant of matrix $\mathbf{R}$ and m is the dimension of vector $\boldsymbol{y}_t$. The normalized weights are given by:

$$\tilde{\boldsymbol{w}}_{t,i} = \frac{\boldsymbol{w}_{t,i}}{\sum_{i=1}^{N}\boldsymbol{w}_{t,i}} \tag{22}$$

HESSD

8, 5849–5890, 2011

Finally, the best estimate of the state consists of the weighted mean of the particle set $\{\hat{\boldsymbol{x}}^{-}_{t,i}, \tilde{\boldsymbol{w}}_{t,i}\}$. The SIS filter poses the problem of particle depletion, this problem is caused by the increase of the variance over time as stated in Kong et al. (1994) and Doucet et al. (2001).

The plots in the upper part of Fig. 2 show the importance weight transition from a uniform distribution at $t = 0$ to a normal distribution according to the Gaussian likelihood pdf at $t = 1$. While in the plots located in the lower part, it is clearly noticeable that after a few model time steps, only one of the normalized importance weights reaches the value of 1, and the remaining set of weights are reduced to negligible values. Consequently, a large number of samples are removed from the sample space, because their weights become numerically insignificant, generating a wrong approximation of the posterior pdf.

A heuristic approach to mitigate the degeneracy problem by increasing the particle set is impractical in most cases. The approach adopted in this work is the Sequential Importance Resampling approach.

5.2.2 Sequential Importance Resampling (SIR)

Resampling is basically the selection and replication of those particles with high importance weights. This additional step to the SIS filter involves mapping the Dirac random measure $\{\boldsymbol{x}_{t,i}, \tilde{\boldsymbol{w}}_{t,i}\}$ into an equally weighted random measure $\{\boldsymbol{x}_{t,i}, N^{-1}\}$. Gordon et al. (1993) proposed a methodology which consists of drawing samples uniformly from the discrete set $\{\boldsymbol{x}_{t,i}, \tilde{\boldsymbol{w}}_{t,i}\}$ and it is referenced as the Sequential Importance Resampling method (SIR). The algorithm consists in the construction of the cumulative distribution of the particles set and the projection of a uniformly drawn sampling index i onto the domain of the distribution. As a result of the projection, the resampling index j is obtained and the particle set $\{\hat{\boldsymbol{x}}^{-}_{t,i}\}$ is resampled according to this index, the resulting particle set $\{\hat{\boldsymbol{x}}^{+}_{t,j}\}$ contains replications of those particles with high importance weight. A detailed description of the SIR resampling algorithm is given in Arulampalam et al. (2002) and extended in Moradkhani et al. (2005). Residual resampling is an improved version of the SIR method and was proposed by Higuchi (1997) and Liu and Chen (1998), but is not used here.

The SIR algorithm used in this study is summarized as follows:

- FOR $t = 1, 2, \ldots$
  - Propagate the particles in time.
    Draw $\hat{\boldsymbol{x}}^{-}_{t,i} \sim p(\hat{\boldsymbol{x}}_{t,i}|\hat{\boldsymbol{x}}^{+}_{t-1,i})$
  - IF (t corresponds to a DA time step)
    - Importance sampling step
      - For $i = 1 : N$
        Compute $\boldsymbol{w}_{t,i} = p(\boldsymbol{y}_t|\hat{\boldsymbol{x}}^{-}_{t,i})$
        Normalize $\tilde{\boldsymbol{w}}_{t,i} = \frac{\boldsymbol{w}_{t,i}}{\sum_{i=1}^{N}\boldsymbol{w}_{t,i}}$
    - Resampling step
      - For $i = 1 : N$
        Obtain the resampling index j vector.
        Resample $\{\hat{\boldsymbol{x}}^{-}_{t,i}\} \Rightarrow \{\hat{\boldsymbol{x}}^{+}_{t,j}\}$
        Assign $\tilde{\boldsymbol{w}}_{t,i} = \frac{1}{N}$
  - END IF
- END FOR

The replication of the particles during the resampling step poses a problem when the set of resampled particles collapses in the worst case to a single particle due to a wrong selection of the importance pdf or due to a narrow likelihood pdf. Figure 3 shows the resampling index j, which indicates the location of the particles to be resampled, at 4 DA events. Subfigures (a) and (b) indicate a proper performance of the resampling algorithm where the particle replication is not extreme. On the other hand, the resampling index j, as a result of the application of a hypothetical and too narrow

HESSD
8, 5849–5890, 2011

likelihood pdf, is presented in subfigures (c) and (d), subfigure (c) indicates that the sample collapses to the particle values located at positions 27th, 53th, and 60th. The extreme replication problem is noticeable mostly in subfigure (d) where almost all the particles collapse to the value of the particle located at the 54th position.

The particle degenaracy problem can be handled using Markov chain Monte Carlo (MCMC) steps (Andrieu et al., 1999) after the resampling step. However, the use of MCMC steps increases the computational time considerably due to need of new proposed particles sampled from the prior density function. The scope of this paper is limited to the application of the SIR filter assuming proper importance and likelihood density functions.

5.3 Differences between the EnKF and the SIR-PF

Although the EnKF and the SIR filter are MC-based methods there is a considerable difference in the way MC simulation is applied. The upper part of Fig. 4 represents the prediction (forecast) step and the correction (analysis) step of the EnKF while in the lower part, the SIR filter application is shown. Here, it is important to remark the fact that in the EnKF all the ensemble members at time step t are updated using the same approximated Kalman gain and the innovation factor which depends on the perturbed measurements and the observation model (Eq. 14). On the other hand, in the SIR filter all the particles at time step t are weighted and resampled.

The replication and suppression of particles decrease the particles variance, which limits the state space representation and possibly affects the behaviour of variables related to the assimilated state in a negative way.

6 SIR filter with parameter resampling (SR+PR)

In the EnKF and the PF, the uncertainty in the model is represented through samples referred to as ensemble members or particles, respectively. These samples are drawn

HESSD
8, 5849–5890, 2011

for the EnKF according to Eq. (8) and for the PF from the importance density function (equal to the prior density function for the standard particle filter).

The uncertainty in the model is caused by uncertainty in the meteorological forcings, initial conditions and parameters. Thus, the generation of ensembles, presented in the experiment setup section, is fundamental since the ensemble should represent this model uncertainty. The perturbation of the parameters plays an important role in the generation of the ensemble due to the contribution of the parameters to the modelling errors.

The state estimation method aims at finding the optimal state value based on the information from the measurements. The estimated state value can positively or negatively affect the behaviour of other variables in the model. In this study, soil moisture is the state variable that will affect the baseflow. The key idea of the SIR filter with parameter resampling is that a combination of estimated state values with consistent parameter values can result in a positive impact on the affected variable.

The operation of the parameter resampling step is the following: after the resampling of the states, the same vector/matrix containing the particle indices to be resampled is used to resample the parameter set. The last action leads to a selection (replication or suppression) of parameters that are tied to a particular state realization.

An extreme replication of the parameter values poses the same problem as in the case of the state replication. Moreover, the ensemble will fail in the representation of the model uncertainty since the spread of the ensemble is decreased after the parameter resampling. In order to overcome this problem, the resampled parameter values are perturbed with the addition of white Guassian noise and the variance (*var*) of the noise is set to a fraction of the optimal parameter value.

The SIR with parameter resampling filter is summarized as follows:

- FOR $t = 1, 2, ...$
  - Propagate the particles in time.
  - IF (t corresponds to a DA time step)

- Importance sampling step
- Resampling step
  - For $i = 1 : N$
    Obtain the resampling index j vector.
    Resample $\{\hat{\boldsymbol{x}}^{-}_{t,i}\} \Rightarrow \{\hat{\boldsymbol{x}}^{+}_{t,j}\}$
    Resample the parameter set θ: $\{\theta_i\} \Rightarrow \{\theta_j\}$
    Perturb the resampled parameter set $\theta_j + N(0, var)$
    Assign $\tilde{\boldsymbol{w}}_{t,i} = \frac{1}{N}$

– END IF

– END FOR

7 Results

The data assimilation experiments are validated by comparing soil moisture and discharge assimilation results against synthetic observed soil moisture and baseflow values. The reference model integration without data assimilation is performed with parameter set 2, while the assimilation integrations are performed with a slightly different configuration (parameter sets 1-2-3).

Data assimilation is performed every week, with the first DA event at 8 February 2007 and the last at 24 May 2007. Every DA event is indicated by a black arrow in the figures and the simulation period corresponds to the first half of year 2007 (1 January–1 July). Root Mean Square Error (RMSE), between the synthetic observed and modelled soil moisture and baseflow, is computed over the time period starting 1 day before the first DA event and 1 day after the last DA event.

HESSD
8, 5849–5890, 2011

7.1 EnKF

The EnKF has been widely used and accepted as a sequential data assimilation method in Geosciences. Therefore, this study includes the results from an unsophisticated implementation of the EnKF in order to highlights some limitations and compare them against other filter implementations.

The upper part of Fig. 7 shows the evolution over time of the 64 soil moisture ensemble members in light gray color, the ensemble mean in a black dashed line corresponding to the EnKF performance and the synthetic observations in red dotted line. From the figure, one can note the correction of the soil moisture values at every DA assimilation event. However, the correction is not persistent, and after some simulation time steps the effect of the assimilation disappears. This typical short persistence of the update is largely due to bias as described by De Lannoy et al. (2007a). A number of methodologies to reduce the effect of bias were suggested by De Lannoy et al. (2007a) and Reichle and Koster (2004). The bias originates from integrating the model with a different parameter set than the synthetic observations and it is thus not a priori clear how to partition the bias between forecast and observation bias. While the effect here could have been overcome by a separate state and bias-filter (De Lannoy et al., 2007a), it was opted to use the simplified setup to illustrate how the effect of biased soil moisture assimilation on the depending discharge with different filters. The RMSE between the assimilation results and the observed soil moisture with a value of 2.50 vol% indicates a low improvement when contrasting against the RMSE between the a priori ensemble mean and the observations with 3.07 vol%.

A rather poor assimilation performance with respect to the discharge (not shown) has been found. This can be explained by a poor assimilation analysis of the baseflow in the lower part of Fig. 7. There, the influence of the soil moisture assimilation on the baseflow is very obvious at each arrow. At first glance one can see the presence of peaks in the behaviour of the baseflow as a consequence of the assimilation. The RMSE values of $4.79 \times 10^{-6}\,\mathrm{mm\,s^{-1}}$ for the baseflow a priori ensemble mean and

HESSD
8, 5849–5890, 2011

6.91×10^{-6} mm s^{-1} for the assimilation results corroborates the negative impact on the baseflow after the assimilation. A detailed analysis around the negative impact on the baseflow is presented further.

7.2 SIR filter

Figure 8 shows the performance of the SIR filter for soil moisture assimilation and the corresponding impact of the assimilation on the baseflow. According to the RMSE values: 3.07 vol% without assimilation and 2.90 vol% after assimilation, the improvement obtained from the SIR filter application is non significant. Although the RMSE for the EnKF performance is slightly better than the RMSE for the SIR filter, indicating a slightly better performance of the EnKF, the difference between these values is not so high as to consider a substantial improvement when using the EnKF.

On the other hand, when comparing the assimilation impact on the baseflow between the SIR filter (lower part of Fig. 8) and the EnKF a different performance can be observed. Both filters perform negatively according to the RMSE values (6.91×10^{-6} mm s^{-1} for the EnKF and 9.10×10^{-6} mm s^{-1} for the SIR filter) when compared to the model run without assimilation (4.79×10^{-6} mm s^{-1}), and the negative impact is enlarged with the SIR filter. This is expected because the state updating procedures are different according to what is explained in Sect. 5.3. The replication of those state particles with higher weight in combination with the parameter values affect the baseflow behaviour negatively. In order to assign to each resampled state particle a consistent parameter value, the application of the parameter resampling is evaluated as an alternative to improve the filter performance and to have a positive impact on the baseflow.

7.3 SIR filter with parameter resampling

The SIR filter with parameter resampling aims at a combination of estimated state values with consistent parameter values. This procedure should result in a positive impact on the land surface variables that dynamically depend (through the model, including the parameter configuration) on the assimilated soil moisture state variable.

Figure 9 shows the performance of the soil moisture assimilation and the impact of the assimilation on the baseflow for the SIR+PR filter without the perturbation of the resampled parameters. Looking at Figs. 8 and 9, the decrease in the dispersion of the soil moisture and baseflow particles is noticeable when the parameter resampling is performed. This reduction is indicated by the time-averaged ensemble spread $< ensp >$ (Eq. 1), calculated over the entire validation period with inclusion of the DA time steps, with values of 3.13×10^{-4} and 1.41×10^{-4} $(\mathrm{mm^3\,mm^{-3}})^2$ for soil moisture and values of 1.56×10^{-09} and 2.22×10^{-11} $(\mathrm{mm\,s^{-1}})^2$ for the baseflow.

Resampling the parameters along with the state SIR filter causes a reduction of the analysis error (the ensemble spread represents the uncertainty at the analysis step). Moreover, the order of the error reduction magnitude is different for the soil moisture and baseflow, with a strong reduction for the baseflow. The latter can be explained by the nonlinear relationship between baseflow and soil moisture.

An extreme reduction of the ensemble spread due to an extreme state and parameter particles replication needs to be avoided. Here, we propose the perturbation of the resampled parameters by using additive white Gaussian noise as the solution to the particles collapse problem. The predefined standard deviation of the noise is set to a fraction of the optimal parameter values, for the results presented in Fig. 10 the fraction is set to 0.01 of parameter set 2.

Figure 10 shows the SIR+PR filter performance with the perturbation of the resampled parameters. The upper part of this figure presents the performance for the soil moisture assimilation. The dynamics of the state ensemble is positively affected by the parameter resampling improving the overall performance of the filter and keeping the

HESSD
8, 5849–5890, 2011

benefit of the state updating for a long time after the DA events. The benefit is quantified by the RMSE values corresponding to 3.07 vol % without assimilation and 0.60 vol % when the SIR+PR is applied. Moreover, the perturbation of the resampled parameters increases the ensemble spread from 1.41×10^{-4} $(\mathrm{mm}^3\,\mathrm{mm}^{-3})^2$ to 2.02×10^{-4} $(\mathrm{mm}^3\,\mathrm{mm}^{-3})^2$.

Additionally, the plot of the baseflow (see lower part of Fig. 10) shows graphically a considerable improvement on the behaviour when comparing to the assimilation effects of the EnKF and SIR filter application. This improvement can be corroborated with the reduction in the RMSE values from $4.78 \times 10^{-6}\,\mathrm{mm\,s^{-1}}$ when no assimilation is performed to $1.60 \times 10^{-6}\,\mathrm{mm\,s^{-1}}$ when soil moisture DA is performed. The baseflow ensemble spread can be increased by the parameter perturbation. The ensemble spread values indicate an increase from 2.22×10^{-11} $(\mathrm{mm\,s^{-1}})^2$ to 2.71×10^{-11} $(\mathrm{mm\,s^{-1}})^2$.

An overall conclusion based on the good RMSE values obtained for soil moisture and baseflow is that the addition of the parameter resampling to the SIR filter is effective in removing the bias through an indirect calibration of the modelled particles.

7.4 Sensitivity study

The performance of the EnKF, SIR-filter and SIR+PR filter with parameter perturbation are further analyzed for 3 different initial parameter sets, each identified by the automatic calibration algorithm with a similar optimization index value. The filter performance is analyzed through the comparison of the RMSE values.

Table 4 presents the RMSE values between the estimated and observed volumetric soil moisture at the surface for every filter and for every parameter set. Although the SIR+PR RMSE values are different, due to different system dynamics the SIR+PR filter outperforms the rest of the filters indicating robustness of the algorithm. Additionally, according to Table 3 the positive impact on the baseflow persists among the three cases.

Considering the assimilation of remote sensed soil moisture data, the availability of data is of main importance in the application of the assimilation algorithm. Therefore, the SIR+PR performance is tested for 3 DA frequencies. Additionally to the DA frequency corresponding to 16 DA events, the methodology is evaluated for 8 DA events with 1 event every 2 weeks and 4 DA steps with 1 event every four weeks.

Tables 4 and 5 show the RMSE values for the 3 DA frequencies for soil moisture and baseflow respectively. The values indicate a notorious improvement when using the SIR+PR and the positive impact on the baseflow is maintained for the 3 DA frequencies. However, the improvement of the filters decreases as the assimilation frequency is reduced.

8 Summary and conclusions

The EnKF and the SIR filter have been evaluated for their performance in assimilation soil moisture and the impact thereof on baseflow fluxes. Both filters perform relatively good for the correction of the modelled soil moisture, although it should be noted that they were affected by the presence of bias. The impact of the soil moisture assimilation on the baseflow results indicates a strong negative effect. The SIR+PR approach is presented as a solution to this shortcoming in the EnKF and SIR filter performance.

The SIR+PR filter methodology strives on the correction of the consistency between parameters and soil moisture states replicating the consistent parameters and rejecting the erratic parameter values. Results indicate a notorious improvement of the performance not only in the estimation of the soil moisture but also in the influence on the baseflow.

Yet, a severe replication affects the parameter diversity and leads to an improper representation of the posterior pdf when assimilating data. The perturbation of the resampled parameter set by a white Gaussian noise with zero mean and predefined standard deviation mitigates the side-effects of the replication.

The robustness of the SIR+PR filter has been tested through the evaluation of the SIR+PR filter for different parameter sets and different assimilation frequencies. The positive results of this study are promising with respect to the assimilation of real data.

Acknowledgements. The work in this paper has been funded by the Belgian Science Policy for the HYDRASENS project in the frame of the STEREO II programme. Gabriëlle De Lannoy is a postdoctoral researcher funded by the Foundation of Scientific Research of the Flemish Community (FWO-Vlaanderen).

References

Andersson, L.: Improvement of runoff models, what way to go?, Nord. Hydrol., 23(5), 315–332, 1992. 5850

Andrieu, C., de Freitas, J., Freitas, J. D., and Doucet, A.: Sequential Bayesian Estimation And Model Selection Applied To Neural Networks, 1999. 5864

Arulampalam, M. S., Maskell, S., and Gordon, N.: A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking, IEEE T. Signal Proces., 50, 174–188, 2002. 5862

Brocca, L., Melone, F., Moramarco, T., Wagner, W., Naeimi, V., Bartalis, Z., and Hasenauer, S.: Improving runoff prediction through the assimilation of the ASCAT soil moisture product, Hydrol. Earth Syst. Sci., 14, 1881–1893, doi:10.5194/hess-14-1881-2010, 2010. 5854

Burgers, G., van Leeuwen, P. J., and Evensen, G.: On the Analysis Scheme in the Ensemble Kalman Filter, Mon. Weather Rev., 126, 1719–1724, 1998. 5859

Caparrini, F., Castelli, F., and Entekhabi, D.: 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, 2004. 5851

De Lannoy, G. J. M.: Assimilation of soil moisture observations into a spatially distributed hydrologic model, Ph.D. thesis, Ghent University, 2006b. 5876

De Lannoy, G. J. M., Houser, P. R., Pauwels, V. R. N., and Verhoest, N. E. C.: Assessment of model uncertainty for soil moisture through ensemble verification, J. Geophys. Res., 111, D10101, doi:10.1029/2005JD006367, 2006a. 5853, 5855

De Lannoy, G. J. M., Houser, P. R., Pauwels, V. R. N., and Verhoest, N. E. C.: 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, 2007a. 5851, 5867

De Lannoy, G. J. M., Reichle, R. H., Houser, P. R., Pauwels, V. R. N., and Verhoest, N. E. C.:

HESSD
8, 5849–5890, 2011
Assimilation of soil moisture using the particle filter
D. A. Plaza et al.

Correcting for forecast bias in soil moisture assimilation with the ensemble Kalman filter, Water Resour. Res., 43, W09410, doi:10.1029/2006WR005449, 2007b. 5854

Doucet, A., Gordon, N., and Krishnamurthy, V.: Particle filters for state estimation of jump Markov linear systems, IEEE T. Signal Proces., 49, 613–624, doi:10.1109/78.905890, 2001. 5861, 5862

Duan, Q. Y., Gupta, V. K., and Sorooshian, S.: Shuffled complex evolution approach for effective and efficient global minimization, J. Optim. Theory Appl., 76, 501–521, 1993. 5854

Dunne, S. and Entekhabi, D.: An ensemble-based reanalysis approach to land data assimilation, Water Resour. Res., 41, W02013, doi:10.1029/2004WR003449, 2005. 5851

Entekhabi, D., Nakamura, H., and Njoku, E. G.: Solving the inverse-problem for soil moisture and temperature profiles by sequential assimilation of multifrequency remotely sensed observations, IEEE T. Geosci. Remote., 32, 438–448, 1994. 5851

Giustarini, L., Matgen, P., Hostache, R., Montanari, M., Plaza, D., Pauwels, V. R. N., De Lannoy, G. J. M., De Keyser, R., Pfister, L., Hoffmann, L., and Savenije, H. H. G.: Assimilating SAR-derived water level data into a hydraulic model: a case study, Hydrol. Earth Syst. Sci., 8, 2103–2144, doi:10.5194/hessd-8-2103-2011, 2011. 5852

Gordon, N., Salmond, D., and Smith, A.: Novel approach to nonlinear/non-Gaussian Bayesian state estimation, Radar and Signal Processing, IEE Proceedings F, 140, 107–113, 1993. 5861, 5862

Heathman, G. C., Starks, P. J., Ahuja, L. R., and Jackson, T. J.: Assimilation of surface soil moisture to estimate profile soil water content, J. Hydrol., 279, 1–17, 2003. 5851

Higuchi, T.: Monte Carlo filter using the genetic algorithm operators, J. Stat. Comput. Sim., 59, 1–23, 1997. 5863

Houser, P. R., Shuttleworth, W. J., Gupta, H. V., Famiglietti, J. S., Syed, K. H., and Goodrich, D. C.: Integration of Soil Moisture Remote Sensing and Hydrologic Modeling using Data Assimilation, Water Resour. Res., 34, 3405–3420, 1998. 5851

Kalman, R. E.: A New Approach to Linear Filtering and Prediction Problems, Transactions of the ASME–Journal of Basic Engineering, 82, 35–45, 1960. 5851

Kitagawa, G.: Monte Carlo Filter and Smoother for Non-Gaussian Nonlinear State Space Models, J. Comput. Graph. Stat., 5, 1–25, 1996. 5861

Kong, A., Liu, J. S., and Wong, W. H.: Sequential Imputations and Bayesian Missing Data Problems, J. Am. Stat. Assoc., 89, 278–288, 1994. 5862

Liu, J. S. and Chen, R.: Sequential Monte Carlo Methods for Dynamic Systems, J. Am. Stat.

Assoc., 93, 1032–1044, 1998. 5851, 5863
Matgen, P., Montanari, M., Hostache, R., Pfister, L., Hoffmann, L., Plaza, D., Pauwels, V. R. N., De Lannoy, G. J. M., De Keyser, R., and Savenije, H. H. G.: Towards the sequential assimilation of SAR-derived water stages into hydraulic models using the Particle Filter: proof of concept, Hydrol. Earth Syst. Sci., 14, 1773–1785, doi:10.5194/hess-14-1773-2010, 2010. 5852, 5856
Montzka, C., Moradkhani, H., Weihermller, L., Franssen, H.-J. H., Canty, M., and Vereecken, H.: Hydraulic parameter estimation by remotely-sensed top soil moisture observations with the particle filter, J. Hydrol., 399, 410–421, 2011. 5852
Moradkhani, H., Hsu, K.-L., Gupta, H., and Sorooshian, S.: Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter, Water Resour. Res., 41, W05012, doi:10.1029/2004WR003604, 2005. 5852, 5862
Nagarajan, K., Judge, J., Graham, W. D., and Monsivais-Huertero, A.: Particle Filter-based assimilation algorithms for improved estimation of root-zone soil moisture under dynamic vegetation conditions, Adv. Water Resour., 34, 433–477, 2010. 5852
Paniconi, C., Maroccu, M., Putti, M., and Verbunt, M.: Newtonian nudging for a Richards equation-based distributed hydrological mode, Adv. Water Resour., 26, 161–178, 2002. 5851
Pauwels, V. R. N., Hoeben, R., Verhoest, N. E. C., and De Troch, F. P.: 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, 2001. 5851
Pauwels, V. R. N., Hoeben, R., Verhoest, N. E. C., De Troch, F. P., and Troch, P. A.: Improvement of TOPLATS-based discharge predictions through assimilation of ERS-based remotely sensed soil moisture values, Hydrol. Process., 16, 995–1013, 2002. 5851, 5854
Reichle, R. H. and Koster, R. D.: Bias reduction in short records of satellite soil moisture, Geophys. Res. Lett., 31, L19501, doi:10.1029/2004GL020938, 2004. 5867
Reichle, R. H., Walker, J. P., Koster, R. D., and Houser, P. R.: Extended versus Ensemble Kalman filtering for land data assimilation, J. Hydromet., 3, 728–740, 2002. 5851
Seuffert, G., Wilker, H., Viterbo, P., Drusch, M., and Mahfouf, J. F.: The usage of screen-level parameters and microwave brightness temperature for soil moisture analysis, J. Hydromet., 5, 516–531, 2004. 5851
Walker, J. P., Willgoose, G. R., and Kalma, J. D.: Three-dimensional soil moisture profile retrieval by assimilation of near-surface measurements: Simplified Kalman filter covariance forecasting and field application, Water Resour. Res., 38, 1301,

doi:10.1029/2002WR001545, 2002. 5851

Weerts, A. H. and El Serafy, G. Y. H.: Particle filtering and ensemble Kalman filtering for state updating with hydrological conceptual rainfall-runoff models, Water Resour. Res., 42, W09403, doi:10.129/2005WR004093, 2006. 5852

Zhang, J., Wang, W., and Leung, L.: Contribution of land-atmosphere coupling to summer climate variability over the contiguous United States, J. Geophys. Res., 113, D22109, doi:10.129/2008JD010136, 2008. 5850

HESSD

8, 5849–5890, 2011

Table 1. Optimal parameter sets: NwRb and NwRs were converted into tuneable parameters (De Lannoy, 2006b), k is the soil layer index.

Description	set 1	set 2	set 3
Fraction of model area with high water table (wtfact[fraction])	0.280	0.704	0.742
Water table depth scale parameter (fz [m^{-1}])	49.173	3.423	3.475
Saturated soil hydraulic conductivity (kd [$mm\,s^{-1}$])	0.827	0.095	0.099
Base flow parameter for saturated fraction of watershed (ld [$mm\,s^{-1}$])	0.0071	0.0034	0.0027
First bottom layer contributing to the calculation of base flow (NwRb [–])	5	5	6
Last top layer contributing to the calculation of the surface runoff (NwRs [–])	3	4	4
Clapp and Hornberger constant (bsw_k [–])	5.487	4.659	4.623
Volumetric soil water at saturation ($watsat_k$ [–])	0.638	0.597	0.600
Hydraulic conductivity at saturation ($hksat_k$ [$mm\,s^{-1}$])	0.047	0.011	0.010
Minimum soil suction ($sucsat_k$ [mm])	284.76	557.17	606

HESSD
8, 5849–5890, 2011

Table 2. RMSE [vol%] between the observed and simulated soil moisture for 3 parameter sets.

Filter	set 1	set 2	set 3
Ensemble	2.35	3.07	2.41
EnKF	1.88	2.50	1.91
SIR filter	2.12	2.90	2.33
SIR+PR filter	1.19	0.60	0.96

HESSD

8, 5849–5890, 2011

Assimilation of soil moisture using the particle filter

D. A. Plaza et al.

Table 3. RMSE [mm s^{-1}] between the observed and simulated baseflow for 3 parameter sets.

Filter	set 1	set 2	set 3
Ensemble	7.26×10^{-6}	4.79×10^{-6}	4.61×10^{-6}
EnKF	7.97×10^{-6}	6.91×10^{-6}	5.58×10^{-6}
SIR filter	9.35×10^{-6}	9.10×10^{-6}	7.91×10^{-6}
SIR+PR filter	6.27×10^{-6}	1.60×10^{-6}	3.62×10^{-6}

HESSD
8, 5849–5890, 2011

HESSD

8, 5849–5890, 2011

Assimilation of soil moisture using the particle filter

D. A. Plaza et al.

Table 4. RMSE [vol%] between the observed and simulated soil moisture for 3 DA frequencies.

Filter	16 DA steps	8 DA steps	4 DA steps
Ensemble	3.07	3.07	3.07
EnKF	2.50	2.67	2.88
SIR filter	2.90	2.99	3.01
SIR+PR filter	0.60	0.70	1.08

HESSD
8, 5849–5890, 2011

Assimilation of soil moisture using the particle filter

D. A. Plaza et al.

Table 5. RMSE [mm s^{-1}] between the observed and simulated baseflow for 3 DA frequencies.

Filter	16 DA steps	8 DA steps	4 DA steps
Ensemble	4.79×10^{-6}	4.79×10^{-6}	4.79×10^{-6}
EnKF	6.91×10^{-6}	5.82×10^{-6}	4.61×10^{-6}
SIR filter	9.10×10^{-6}	7.60×10^{-6}	7.60×10^{-6}
SIR+PR filter	1.60×10^{-6}	1.52×10^{-6}	1.22×10^{-6}

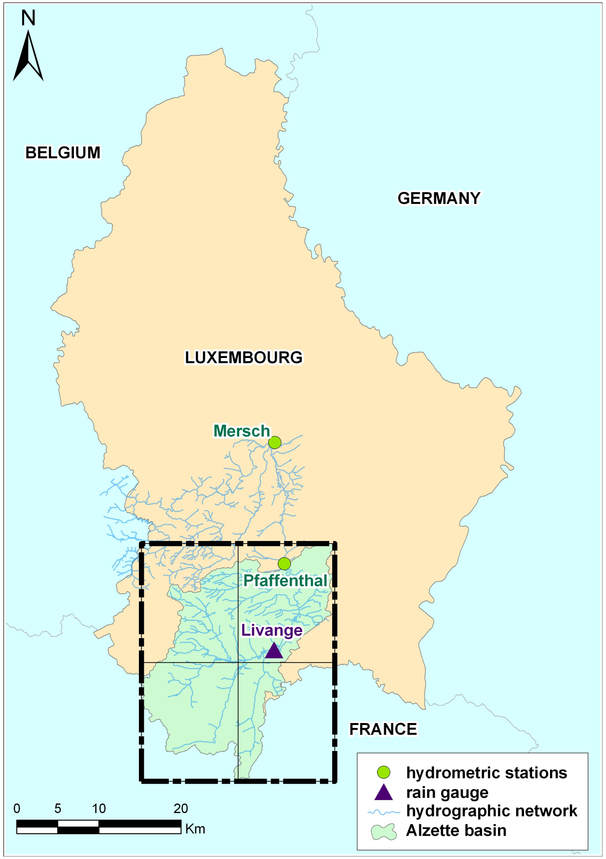

Fig. 1. The study area: the discharge area in the Alzette river basin is indicated by the green patch.

HESSD
8, 5849–5890, 2011

Assimilation of soil moisture using the particle filter

D. A. Plaza et al.

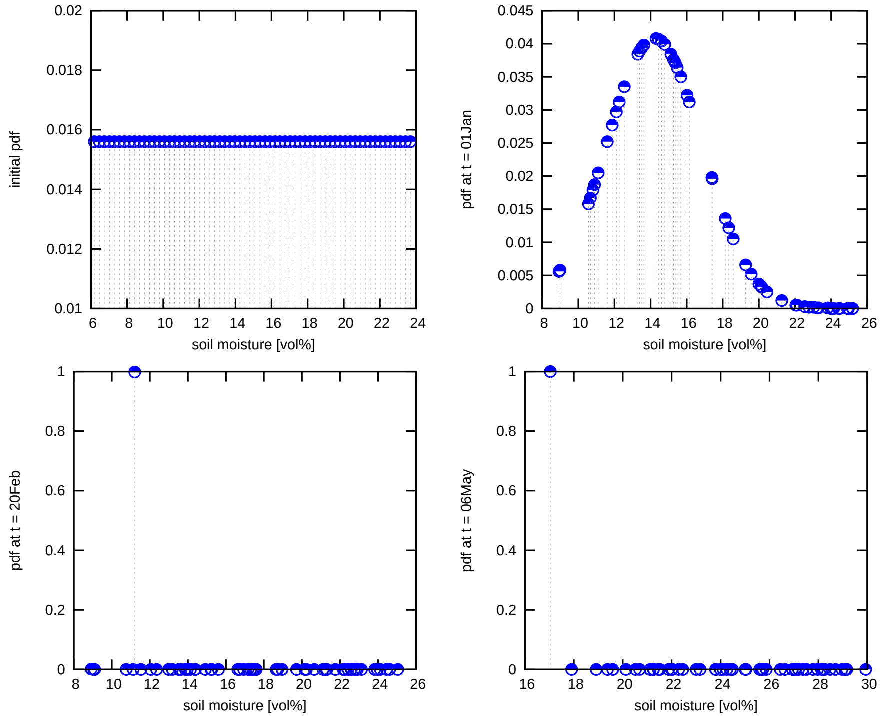

Fig. 2. Monitoring the importance weights: axis x represents the particles location (volumetric soil moisture [vol%]) and axis y the importance weights values at four different daily model time steps 0, 1 (1 January), 51 (20 February), 126 (6 May).

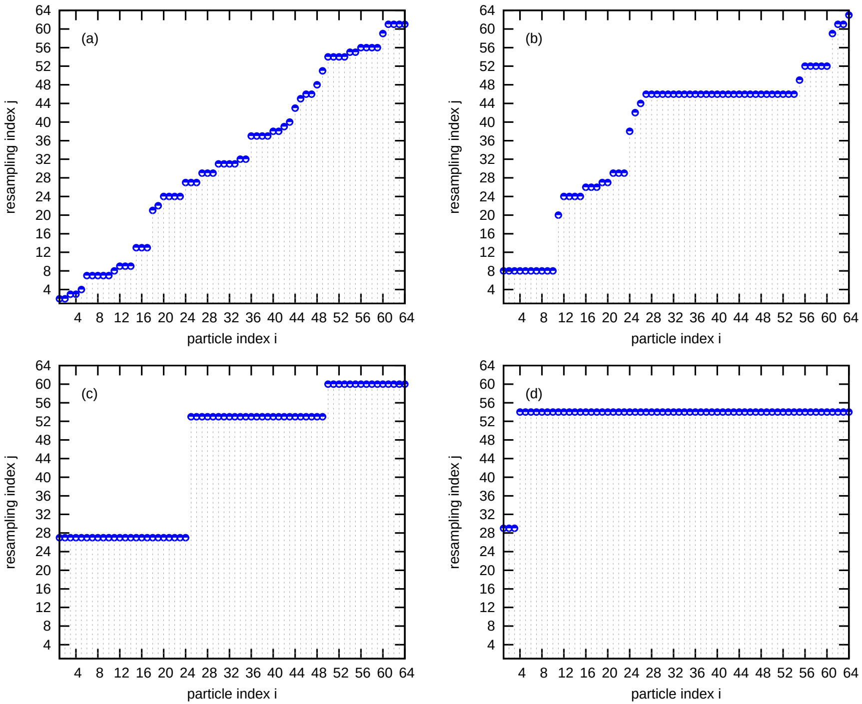

Fig. 3. Extreme particle replication example with $N = 64$: panels **(a)** and **(b)** present a proper resampling performance whereas panels **(c)** and **(d)** show the extreme replication problem. The sample set collapses to particles located at 27th, 53th and 60th positions in panel **(c)** and almost all the set collapses to the value of the particle at the 54th position in panel **(d)**.

HESSD
8, 5849–5890, 2011

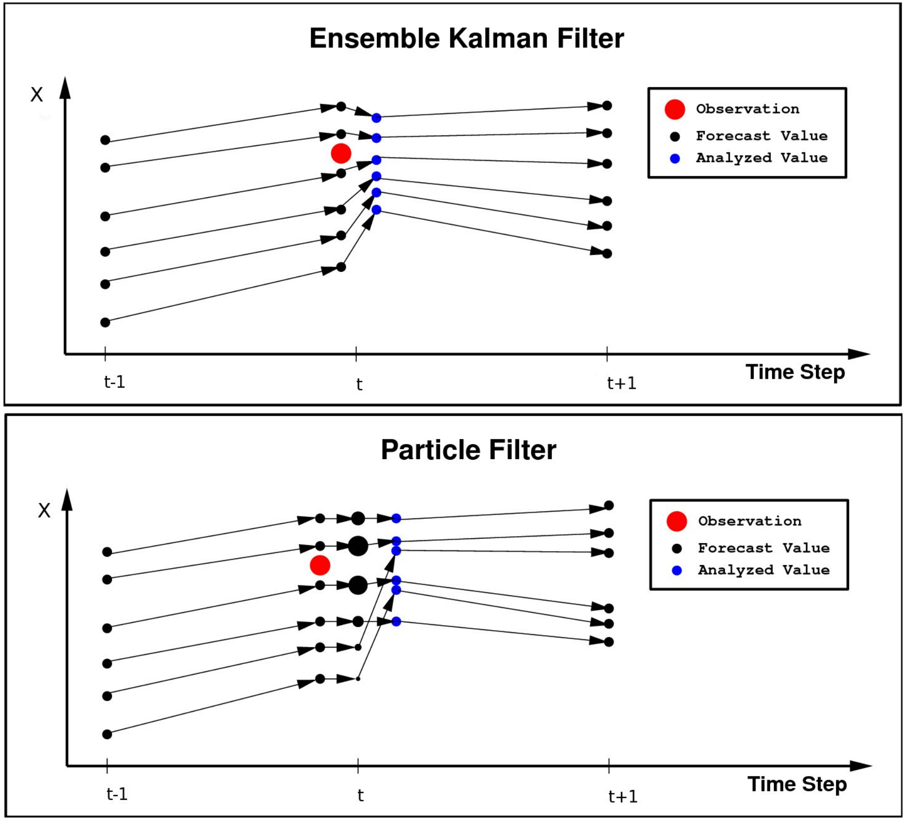

Fig. 4. Forecast (prediction) and analysis (correction) steps at time step t for the EnKF in the upper part of the plot and the SIR filter in the lower part of the plot. In the EnKF, all the ensemble members (blue dots) are updated with the same Kalman gain whereas in the PF, the update consists in the replication/suppression of the particles.

HESSD
8, 5849–5890, 2011

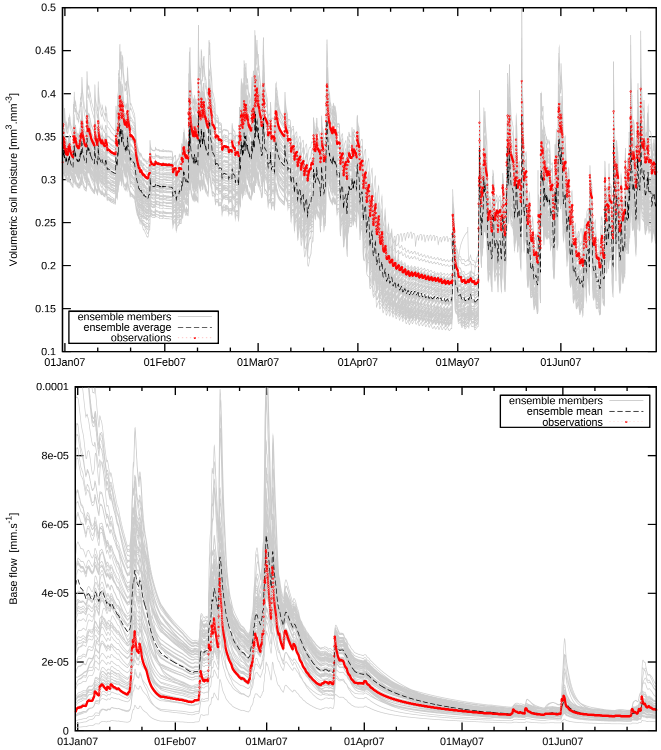

Fig. 5. Soil moisture and baseflow ensembles: the upper plot corresponds to the generation of the volumetric soil moisture ensemble with the ensemble members in gray line, the ensemble mean in black line and the synthetic soil moisture observations in red dotted line. The lower plot corresponds to the baseflow ensemble.

HESSD
8, 5849–5890, 2011

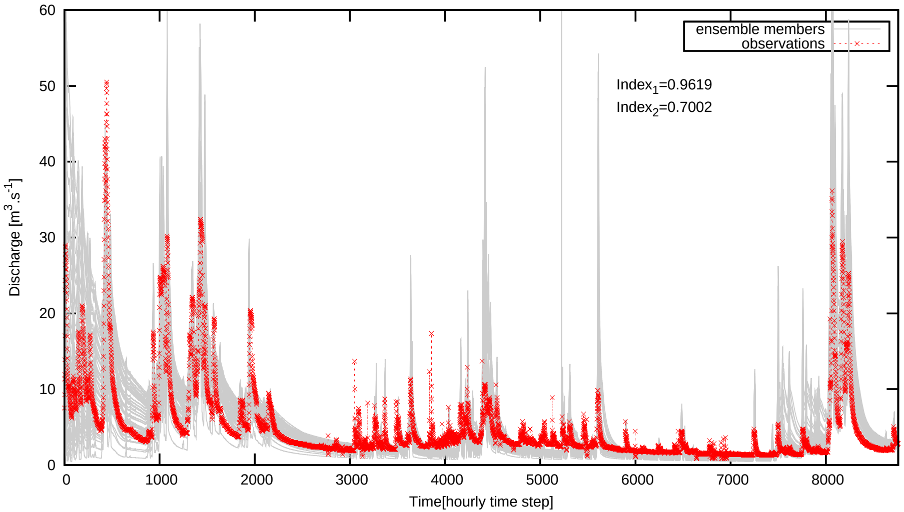

Fig. 6. Generation of the discharge ensemble: the ensemble quality is controlled by index$_1$ (Eq. 2) and index$_2$ (Eq. 3).

HESSD
8, 5849–5890, 2011

Assimilation of soil moisture using the particle filter

D. A. Plaza et al.

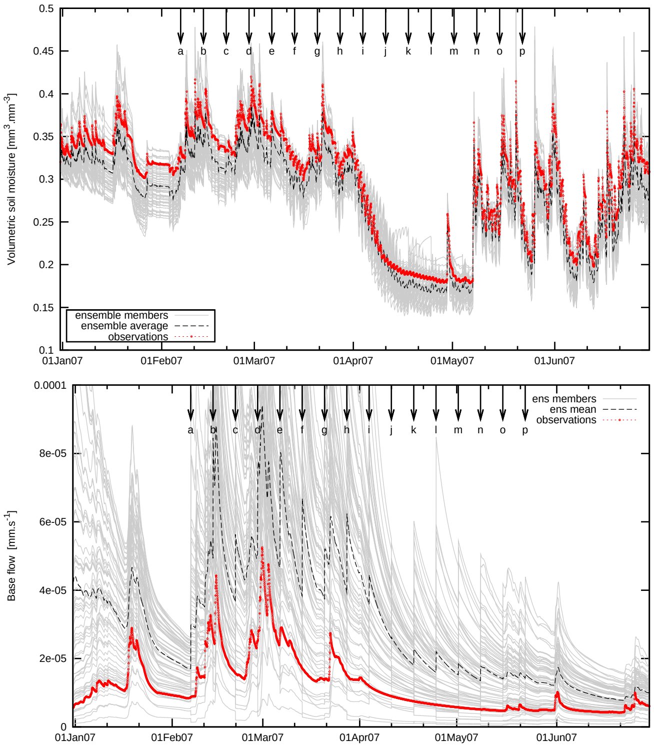

Fig. 7. Kalman filter performance and assimilation impact on the baseflow. The soil moisture and baseflow time series correspond to the DA study performed with set 2 as the initial parameter set.

HESSD
8, 5849–5890, 2011

Assimilation of soil moisture using the particle filter

D. A. Plaza et al.

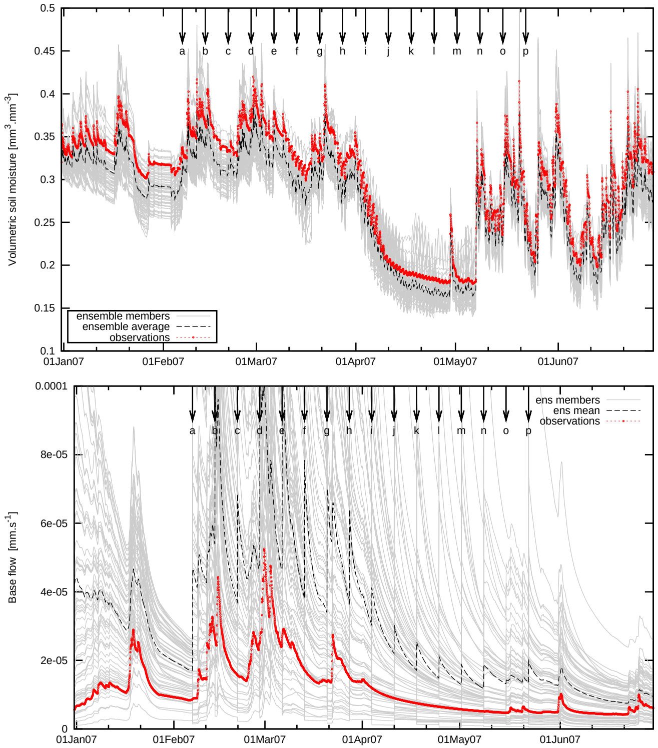

Fig. 8. SIR filter performance and assimilation impact on the baseflow. The soil moisture and baseflow time series correspond to the DA study performed with set 2 as the initial parameter set.

HESSD
8, 5849–5890, 2011

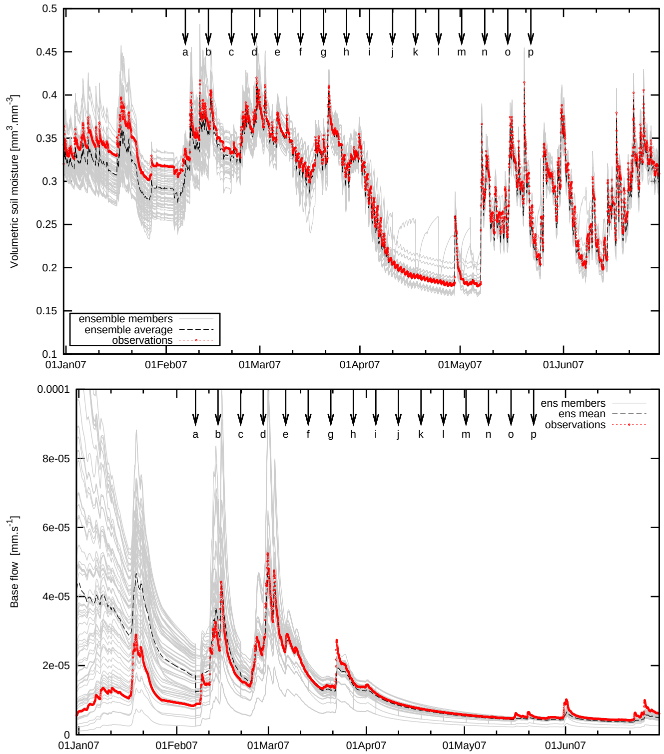

Fig. 9. SIR+PR filter performance and assimilation impact on the baseflow without parameter perturbation. The soil moisture and baseflow time series correspond to the DA study performed with set 2 as the initial parameter set.

HESSD
8, 5849–5890, 2011

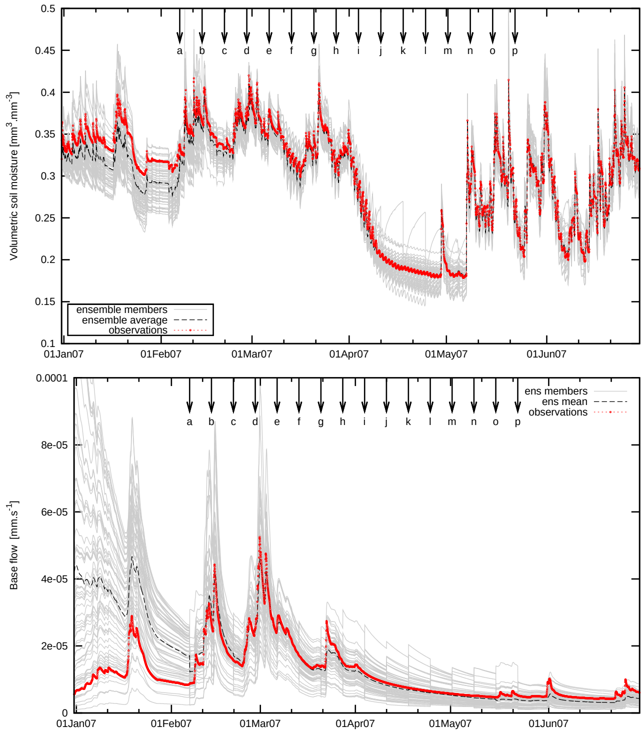

Fig. 10. SIR+PR filter performance and assimilation impact on the baseflow with parameter perturbation. The soil moisture and baseflow time series correspond to the DA study performed with set 2 as the initial parameter set.

HESSD
8, 5849–5890, 2011

Assimilation of soil moisture using the particle filter

D. A. Plaza et al.