Hydrology and Earth System Sciences Discussions

Applying sequential Monte Carlo methods into a distributed hydrologic model: lagged particle filtering approach with regularization

S. J. Noh[1], Y. Tachikawa[2], M. Shiiba[2], and S. Kim[2]

[1]Department of Urban and Environmental Engineering, Kyoto University, Kyoto, Japan
[2]Department of Civil and Earth Resources Engineering, Kyoto University, Kyoto, Japan

Correspondence to: S. J. Noh (seongjin.noh@gmail.com)

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

HESSD

8, 3383–3420, 2011

Applying sequential Monte Carlo methods into a DHM

S. J. Noh et al.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◄ | ►

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Abstract

Applications of data assimilation techniques have been widely used to improve hydrologic prediction. Among various data assimilation techniques, sequential Monte Carlo (SMC) methods, known as "particle filters", provide the capability to handle non-linear and non-Gaussian state-space models. In this paper, we propose an improved particle filtering approach to consider different response time of internal state variables in a hydrologic model. The proposed method adopts a lagged filtering approach to aggregate model response until uncertainty of each hydrologic process is propagated. The regularization with an additional move step based on Markov chain Monte Carlo (MCMC) is also implemented to preserve sample diversity under the lagged filtering approach. A distributed hydrologic model, WEP is implemented for the sequential data assimilation through the updating of state variables. Particle filtering is parallelized and implemented in the multi-core computing environment via open message passing interface (MPI). We compare performance results of particle filters in terms of model efficiency, predictive QQ plots and particle diversity. The improvement of model efficiency and the preservation of particle diversity are found in the lagged regularized particle filter.

1 Introduction

Data assimilation is a way to integrate information from a variety of sources to improve prediction accuracy, considering the uncertainty in both a measurement system and a prediction model. There has been considerable advances in hydrologic data assimilation for streamflow prediction (e.g., Kitanidis and Bras, 1980; Georgakakos, 1986; Vrugt et al., 2006; Clark et al., 2008; Seo et al., 2003, 2009). State-space filtering methods based on variations of the Kalman filter (KF) approach have been proposed and implemented due to their potential ability to explicitly handle uncertainties in hydrologic predictions. However, the KF approaches for a non-linear system such as

the extended Kalman filter (EKF) have limitations in the practical application due to their instability for strong non-linearity and high computational cost of model derivative equations, especially for high-dimensional state-vector problems, such as spatially distributed models. To cope with the drawbacks of the EKF, the ensemble Kalman filter (EnKF) was introduced by Evensen (1994). The EnKF is computationally efficient due to no need of model covariance estimation, but it is still based on the assumption that all probability distributions involved are Gaussian. Further reviews of Kalman filter based applications for hydrologic models are shown in Vrugt et al. (2006), Moradkhani (2005b, 2008) and Evensen (2009).

Another approach to data assimilation is the variational assimilation (VAR), which has achieved in widespread application to weather and oceanographic prediction models. In hydrologic investigations, VAR is implemented for estimating spatial soil-moisture distributions by Reichle et al. (2001) and for assimilating potential evaporation and real-time observations of streamflow and precipitation to improve streamflow forecasts by Seo et al. (2003, 2009). Although variational methods are computationally efficient than KF-based methods, the derivation of the adjoint model needed for minimization of a cost function is a difficult task, especially in the case of non-linear, high dimensional hydrological applications (e.g., Liu and Gupta, 2007).

Among data assimilation techniques, the sequential Monte Carlo (SMC) methods, known as particle filters, are a Bayesian learning process in which the propagation of all uncertainties is carried out by a suitable selection of randomly generated particles without any assumptions about the nature of the distributions (Gordon et al., 1993; Musso et al., 2001; Arulampalam et al., 2002; Johansen, 2009). Unlike the various Kalman filter-based methods that are basically limited to the linear correction step and the assumption of Gaussian distribution errors, SMC methods have the advantage of being applicable to non-linear and non-Gaussian state-space models. The application of these powerful and versatile methods has been increasing in various areas, including pattern recognition, target tracking, financial analysis, and robotics.

Applying sequential Monte Carlo methods into a DHM

S. J. Noh et al.

In recent years these methods have received considerable attention in hydrology and earth sciences (e.g., Moradkhani et al., 2005a; Weert and El Serafy, 2006; Zhou et al., 2006; van Delft et al., 2009; van Leeuwen, 2009; Karssenberg et al., 2010). Since their first introduction to the rainfall-runoff modeling by Moradkhani et al. (2005a), Weerts and El Serafy (2006) compared the ensemble Kalman filtering and the particle filtering for state updating of hydrological conceptual rainfall-runoff models. The SMC methods have also been applied for parameter estimation and uncertainty analysis of hydrological models. Smith et al. (2008) evaluate structural inadequacy in hydrologic models, Qin et al. (2009) estimate both soil moisture and model parameters, and Rings et al. (2010) implement hydrogeophysical parameter estimation. Uncertainty of a distributed hydrological model is analyzed by Salaman and Feyen (2009, 2010) and dual state-parameter updating of a conceptual hydrologic model is applied for flood forecasting by Noh et al. (2011). The diversity of assimilated data and models has been increasing; snow water equivalent prediction model (Leisenring and Moradkhani, 2010) and assimilation with remote sensing-derived water stages (Montanari et al., 2009). However, the framework to deal with the delayed response, which originates from different time scale of hydrologic processes, routing and spatial heterogeneity of catchment characteristics and forcing data especially in a distributed hydrologic model, has not been thoroughly addressed in the hydrologic data assimilation. Furthermore, alternative methods proposed in the literature to mitigate loss of sample diversity (e.g., Musso et al., 2001; Arulampalam et al., 2002), which may cause collapse of filtering system, have not been studied in hydrology.

In this paper, we apply particle filters for a distributed hydrologic model to enhance forecasting capability of streamflow. A lagged particle filtering approach is proposed to consider different response time of internal states in a distributed hydrologic model. The regularized particle filter with the Markov chain Monte Carlo (MCMC) move step is also adopted to improve sample diversity under the lagged filtering approach. The proposed filtering approach is evaluated with its prediction accuracy, predictive QQ plots and particle diversity in comparison with other types of filtering approaches.

A process-based distributed hydrologic model, WEP (Jia and Tamai, 1998; Jia et al., 2001, 2009), is implemented for sequential data assimilation through state updating of internal hydrologic variables. Particle filtering is parallelized and implemented in the multi-core computing environment via open message passing interface (MPI).

⁵ The paper is organized in the following way. Section 2 outlines the Bayesian filtering theory and particle filters. In Sect. 3, a lagged filtering approach is introduced with additional regularization step to reflect different response of internal processes in sequential data assimilation. Section 4 presents the case study results demonstrating the applicability of proposed particle filtering approach. Several SMC filters are evaluated for real-time streamflow forecasting in the Katsura River catchment using the WEP model. Sequential data assimilation is performed by several different schemes of the particle filters: lagged regularized particle filter (lagged RPF), lagged sequential importance resampling (SIR) and SIR particle filter. The performance results of various SMC methods are compared via model efficiency, predictive QQ plots and ratio of the effective particle number. Section 5 summarizes the results and conclusions.

2 Method of particle filters

In this section we briefly describe the theory of Bayesian filtering and sequential Monte Carlo (SMC) filtering for its suboptimal solution in non-linear and non-Gaussian cases. Several variants of SMC filters are explained including sequential importance resampling (SIR) and regularized particle filter (RPF), which are based on the sequential importance sampling (SIS). The detailed description of sequential Monte Carlo methods can be found in Arulampalam et al. (2002) and Moradkhani et al. (2005a).

2.1 Bayesian filtering theory and basic particle filtering methods

To define the problem of the Bayesian filtering, consider a general dynamic state-space model which is described as follows:

$$x_k = f\left(x_{k-1}, \theta, u_k\right) + \omega_k \quad \omega_k \sim N(0, W_k) \tag{1}$$

$$y_k = h(x_k, \theta') + v_k \qquad v_k \sim N(0, V_k) \tag{2}$$

where $x_k \in \Re^{n_x}$ is the n_x dimensional vector denoting the system state at time k. The operator $f : \Re^{n_x} \to \Re^{n_x}$ expresses the system transition in response to the forcing data u_k (e.g. rainfall, weather data) and parameters θ. $h : \Re^{n_x} \to \Re^{n_y}$ expresses the measurement function having parameters θ'. ω_k and v_k represent the model error and the measurement error, respectively and W_k and V_k represent the covariance of the error.

In the Bayesian recursive estimation, if the system and measurement models are non-linear and non-Gaussian, it is not possible to construct the posterior probability density function (PDF) of the current state x_k given all the measurement $y_{1:k} = \{y, i = 1, \ldots, k\}$ analytically. When the analytic solution is intractable, an optimal solution can be approximated by SMC filters.

Sequential Monte Carlo (SMC) filters are a set of simulation-based methods that provide a flexible approach to computing the posterior distribution without any assumptions about the nature of the distributions. The key idea of SMC filters is based on point mass ("particle") representations of probability densities with associated weights as:

$$p(x_k|y_{1:k}) \approx \sum_{i=1}^{n} w_k^i \delta\left(x_k - x_k^i\right) \tag{3}$$

where x_k^i and w_k^i denote the ith posterior state ("particle") and its weight, respectively, and $\delta(\cdot)$ denotes the Dirac delta function. Since it is usually impossible to sample from the true posterior PDF, an alternative is to sample from a proposal distribution, also

called importance density, denoted by $q(x_k|y_k)$. After the several steps of computation, the recursive weight updating can be derived as follows:

$$w_k^i \propto w_{k-1}^i \frac{p\left(y_k|x_k^i\right)p\left(x_k^i|x_{k-1}^i\right)}{q\left(x_k^i|x_{k-1}^i,y_k\right)} \tag{4}$$

The choice of importance density is one of the most critical issues in the design of SMC methods. The most popular choice is the transitional prior as:

$$q\left(x_k^i|x_{k-1}^i,y_k\right) = p\left(x_k^i|x_{k-1}^i\right) \tag{5}$$

By substituting Eq. (5) into Eq. (4), the weight updating becomes

$$w_k^i \propto w_{k-1}^i p\left(y_k|x_k^i\right) \tag{6}$$

The sequential importance sampling (SIS) algorithm shown above is a Monte Carlo filter that forms the basis for most SMC filters. A common problem with the SIS algorithm is the degeneracy phenomenon, where after a few iterations, all but one particle will have negligible weight. A suitable measure of the degeneracy is the effective sample size n_{eff} estimated as follows (Kong et al., 1994):

$$n_{\text{eff}} = \frac{1}{\sum_{i=1}^{n}\left(w_k^i\right)^2} \tag{7}$$

If the weights is uniform (i.e. $w_k^i = 1/n$ for $i = 1,\dots,n$) then $n_{\text{eff}} = n$. If all but one particle have 0 weight, then $n_{\text{eff}} = 1$. Ratio of the effective particle number n_{ratio} is estimated as follows:

$$n_{\text{ratio}} = \frac{n_{\text{eff}}}{n} \tag{8}$$

The maximum of n_{ratio} is 1 when the weights are uniform. Small n_{ratio} indicates a severe degeneracy and vice versa. n_{ratio} is used as an indicator of degeneracy in this study because it can be used easily regardless of the particle number.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

|◄ | ►|

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

The degeneracy phenomenon can be reduced by performing the resampling step whenever a significant degeneracy is observed. Thus, the sequential importance resampling (SIR) particle filter is drived from the SIS algorithm by performing the resampling step at every time index. The idea of resampling is simply that particles with very low weights are abandoned, while multiple copies of paraticles are kept with the uniformly weighted measure $\{x_k^j, n^{-1}\}$, which still approximates the posterior PDF, $p(x_k|y_{1:k})$ (van Leeuwen, 2009).

Resampling is one of the key issues in the SMC filters and there are various resampling approaches that have been introduced in the literature such as multinomial resampling, residual resampling, stratified resampling and systematic resampling, etc. A comparative anlaysis and review of resampling approaches can be found in Douc et al. (2005) and van Leeuwen (2009). Systematic resampling, also known as stochastic universal sampling, is often preffered due to its computational simplicity and good empirical performance. It has been also shown that systematic resampling has the lowest sampling noise (Kitagawa, 1996). Hence, we use systematic resampling for all particle filtering cases in this study. It is worth noting that there are several choices in resampling methods and the proper method may be different according to the characteristics of hydrologic models. See Weerts and El Serafy (2006), Rings et al. (2010), Salaman and Feyman (2009) for residual resampling, see also Salaman and Feyman (2010), Moradkhani et al. (2005a) for systematic resampling. Although the SIR method has the advantage that the importance weights are easily evaluated, as resampling is applied at every iteration, this filter may lead to a sudden loss of diversity in particles and is sensitive to outliers.

2.2 Regularized particle filter

The positive effects of the resampling step are to automatically concentrate particles in regions of interest of the state-space and to reduce particle degeneracy. However, the particles resampled from high weights are statistically selected many times. This leads to another problem, known as sample impoverishment, which means a loss of

diversity among the particles as the resultant sample will contain many repeated points (Ristic et al., 2004). There have been some systematic techniques proposed to solve the problem of the sample impoverishment. An alternative solution is to introduce the regularization step when the sample impoverishment becomes severe. Regularized particle filter (RPF) is based on regularization of the empirical distribution associated with the particle system using the kernel method (Musso et al., 2001). The main idea of RPF consists in changing the discrete approximation of posteri or distribution to a continuous approximation such that the resampling step is changed into simulating an absolutely continuous distribution, hence producing a new particle system with n different particle locations. The concept of discrete and continuous approximation of particle density is illustrated in Fig. 1. If the weights are concentrated on the limited number of particles, the resampling in the discrete approximation (e.g. SIR particle filter) may lead to a poor representation of the posterior density, while a continuous approximation in regularized measure improves the diversity in the resampling step.

In the RPF, samples are drawn from the approximation

$$p\left(x_k|y_{1:k}\right) \approx \sum_{i=1}^{n} w_k^i K_h\left(x_k - x_k^i\right) \tag{9}$$

where

$$K_h(x) = \frac{1}{h^{n_x}} K\left(\frac{x}{h}\right) \tag{10}$$

is the rescaled kernel density $K(\cdot)$; $h > 0$ is the bandwidth; and n_x is the dimension of the state vector x. The kernel density is a symmetric probability density function on $\Re^{n_x}$, such that

$$K > 0, \quad \int K(x)dx = 1, \quad \int xK(x)dx = 0, \quad \int \|x\|^2 K(x)dx < \infty. \tag{11}$$

The kernel $K(\cdot)$ and bandwidth h are chosen so as to minimise the mean integrated square error (MISE) between the true posterior density and the corresponding regularized weighted empirical measure in Eq. (9), which is defined as

$$\text{MISE}(\hat{p}) = E\left[\int \left[\hat{p}(x_k|y_{1:k}) - p(x_k|y_{1:k})\right]^2 dx_k\right] \tag{12}$$

where $\hat{p}(\cdot|\cdot)$ denotes the approximation to $p(x_k|y_{1:k})$ given by the right-hand side of Eq. (9). In the special case of equally weighted samples, $w^i = 1/n$ for $i = 1,\dots,n$, the optimal choice of the kernel is the Epanechnikov kernel,

$$K_{\text{opt}} = \begin{cases} \frac{n_x+2}{2c_{n_x}}(1 - \|x\|^2) & \text{if} \quad \|x\| < 1 \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

where c_{n_x} is the volume of the unit sphere of $\Re^{n_x}$. It is worth noting that the use of kernel approximation becomes increasingly less appropriate as n_x (dimensionality of the state) increases. To reduce the computational cost, the samples can be generated from the Gaussian kernel instead of the Epanechikov kernel. The optimal bandwidth with unit covariance matrix is

$$h_{\text{opt}} = A \cdot N^{-\frac{1}{n_x+4}} \qquad \text{with} \qquad A = \left[4/(n_x+2)\right]^{\frac{1}{n_x+4}} \tag{14}$$

The RPF differs from the SIR only in additional regularization steps when sample impoverishment happens (effective particle number is less than threshold). The key step is

$$x_k^{i*} = x_k^i + h_{\text{opt}} D_k \varepsilon^i \tag{15}$$

where x_k^{i*} is new particle generated from kernel density; and S_k is the empirical covariance matrix such that $D_k D_k^T = S_k$. Theoretical disadvantage of the RPF is that its samples are no longer guaranteed to asymptotically approximate those from the posterior. This can be mitigated including the Markov chain Monte Carlo (MCMC) move step

(Gilks and Berzuini, 2001) based on the Metropolis-Hastings algorithm (Robert and Casella, 1999). The key idea is that a resampled particle is moved to a new state according to Eq. (15), only if $u \leq \alpha$, where $u \sim U[0, 1]$ and α is the acceptance probability. Otherwise, the move is rejected.

$$\alpha = \min\left\{ 1, \frac{p\left(y_k | x_k^{i^*}\right)}{p\left(y_k | x_k^{i}\right)} \right\} \tag{16}$$

In Eq. (16), α becomes 1.0 when the likelihood of new particle is greater than that of the previous particle. That means that the MCMC move step contributes for screening bad particles in the regularization step ensuring particles asymptotically approximate samples from the posterior. Although this approach is frequently found to improve performance, despite a less rigorous deviation, RPF has not been introduced in hydrologic data assimilation. In this study, the regularization and the MCMC move step are implemented only when loss of sample diversity is detected in the updating stage.

3 Particle filter with lag time approach

In general, there are many types of state variables in a distributed hydrologic model and each variable interacts with each other based on different time scales. For example, in catchment modeling, internal state variables may refer to two-dimensional distribution of soil moisture content, evapotranspiration and overland flow; and an observable state may refer to streamflow flux at the monitoring sites. There is a time lag until the changes of soil moisture distribution affect infiltration and sub-surface/surface runoff processes and generated runoff is routed as streamflow into the measurement site. Hydrologic components in a model have usually different time scales, which needs to be considered in the data assimilation process.

As stated by Salaman and Feyen (2010), this response time is usually larger than the high-frequency discharge measurements. One simple approach is to use the delayed

updating, which gives larger time intervals before updating. However, the delayed updating leads to omitting large quantities of measurement information and fixed delay assumption may result in inappropriate estimation, because a response time always changes according to the current spatial distributions of the state and forcing variables.

5 Furthermore, when system dynamic is relatively fast (e.g. hourly-based hydrologic or hydraulic modeling cases), delayed updating may lead to missing of proper timing of assimilation. That can make it hard to implement sequential data assimilation techniques into hydrologic modeling. Thus, we propose a new lagged particle filtering approach not only to consider different catchment responses but also to use whole measurement

10 information for data assimilation.

Figure 2 shows an example of a newly proposed lagged particle filtering approach. Here, k is the current time step and j is the lag time required for response of internal state variables to be transmitted into the observable variables. Note that it is better to set the lag time j large enough to cover plausible ranges as the system response is

15 time-variant. The assimilation window of the lagged filtering is defined from $k - j$ to k time step. The procedure of the lagged filtering is as follows. (1) To have prediction at the time step k, simulation starts from the time step $k - j$. (2) The weights of particles are estimated according to the measurement at each time step. (3) When particles arrive at the current time, the lagged weights are calculated through aggregating past

20 weights. (4) Resampling is executed according to the lagged weights. Note that state variables at the time step $k - j + 1$ are resampled simultaneously with those at the current time step. (5) For the next time step $k + 1$, simulation starts from the time step $k - j + 1$ and has the same procedure from 1) to 4). In this manner, the sequential data assimilation procedure is implemented at every time step without loss of measurement

25 information.

Lagged weight, w_{lag}^{i}, can be calculated through the aggregation of the past weights as:

$$w_{\text{lag}}^{i} \propto \prod_{t=1}^{j} \left(w_{k-j+t}^{i} \right)^{\alpha_t} \qquad (17)$$

A lagged weight is normalized according to their summation. Weighting factor α_t can be decided via the prior information of system response or calibration processes. In this study, we use a square root form of weighting factor as:

$$\alpha_t = \sqrt{t} \qquad t = 1, \ldots, j \tag{18}$$

5 Using Eq. (18) means that we provide more confidence for the latest information. Compared to conventional particle filtering, an additional procedure needed in the lagged filtering is only that state variables at the time step $k - j + 1$ should be stored and re-sampled according to lagged weights.

We combine the lagged filtering approach with the regularized particle filter (RPF) to
10 enhance the sample diversity. Figure 3 illustrates the regularization step in the lagged filtering window. If the effective particle number n_{eff} is less than threshold ($n_{eff} < n_{thr}$), the regularization step is executed performing the simulation in the same time loop with different particle members, which are generated from kernel. Note that diversity of particles is enhanced in this step because distribution of kernel density is usually
15 broader than process noise. As mentioned in the previous section, the Markov chain Monte Carlo (MCMC) move step is also implemented along with the lagged filtering to make perturbed particle system via regularization asymptotically approximate samples from the posterior. When each particle arrives in the current time step k, acceptance probability α is calculated according to the lagged likelihood as shown in Eq. (16). If
20 a particle is rejected ($u > \alpha$), state variables before regularization will be used without kernel perturbation.

Figure 4 summaries one cycle of RPF with the Markov chain Monte Carlo (MCMC) move step under the lagged filtering approach. The assimilation window and basic procedure before regularization are the same as explained above. In the regularization
25 step, new particles from kernel are propagated along the same time window. Arrived particles are accepted according to the acceptance probability of the MCMC move step and vice versa. It is worth mentioning that the regularization step can be executed not just in the sample impoverishment case but also in the particle collapse case, which

means all particle have negligible weights falling outside of the measurement PDF. In this case, the regularization step is used effectively for re-initialization of the particle system.

4 Implementation

4.1 Study area

The SMC methods are applied to the Katsura River catchment (Fig. 5) to demonstrate the improvement of the accuracy of streamflow forecasting. This catchment is located in Kyoto, Japan, and covers an area of 1100 km^2 (887 km^2 at the Katsura station) (see Fig. 5). Topography is characterized by mountainous upstream in the North and a flatter plain in the south. The elevation in the catchment ranges from 4 to 1158 m, with an average of about 325 m. The land use consists of forest (70%), agricultural area (14%) and residential area (8%), respectively. There are 13 rainfall observation stations, 1 meteorological observation station and 4 river flow observation stations. Annual precipitation and temperature are about 1422 mm and 16.2 °C in Kyoto city (2001 ~ 2010). Precipitation is concentrated in the summer season from May to September. The Hiyoshi dam is located upstream, and the controlled outflow record from the dam reservoir is given as inflow to the hydrologic model and the model simulates rainfall-runoff processes for the downstream of the dam.

4.2 Hydrological model and particle filtering

The hydrologic model used is the water and energy transfer processes (WEP) model, which is developed for simulating spatially variable water and energy processes in catchments with complex land covers (Jia and Tamai, 1998; Jia et al., 2001). State variables include soil moisture content, surface runoff, groundwater tables, discharge and water stage in rivers, heat flux components, etc. (Fig. 6). The spatial calculation

unit of the WEP model is a square or rectangular grid. Runoff routing on slopes and in rivers is carried out by applying one-dimensional kinematical wave approach from upstream to downstream. The WEP model has been applied in several watersheds in Japan, Korea and China with different climate and geographic conditions (Jia et al., 2001, 2009; Kim et al., 2005; Qin et al., 2008).

Model setup uses 250 m grid resolution and an hourly time step. Simulation is divided into two part; calibration (1 June–31 July 2007) and validation (1 July–31 August 2003). One month of warm-up period is added before the data assimilation starts. We use hourly observed rainfall from 13 observation stations organized by the Ministry of Land, Infrastructure, Transport and Tourism in Japan (http://www1.river.go.jp/) and hourly observed meteorological data from Kyoto station including air temperature, relative humidity, wind speed and duration of sunlight organized by Japan Meteorological Agency (http://www.jma.go.jp/jma/index.html). Hourly observed discharges from Katsura station is used for the data assimilation.

SRTM 90 m digital elevation map (DEM) was adopted (http://srtm.csi.cgiar.org/) and converted into 250 m resolution. Soil distribution was obtained from the website of Food and Agriculture Organization of the United Nations (http://www.fao.org/nr/land/soils/en/). Physical property of soil was derived from soil texture information using the ROSETTA model (Schaap et al., 2001). However, the saturated hydraulic conductivity of several soils was adjusted via calibration period as soil property estimated from large scale soil map has large uncertainty. For other parameters related with aquifer and vegetation, we applied parameter ranges from previous studies mentioned above. Artifical water use was approximately estimated as $3 \, \text{m}^3 \, \text{s}^{-1}$ and subtracted directly from simulated discharge at the Katsura station.

Ensemble simulation of 384 particles was conducted on a multi-processing computer (128 cores in the supercomputing system of Kyoto University) via parallel-computing techniques of open message passing interface (MPI) (http://www.open-mpi.org/). Parallel programming is written using single-program multiple-data (SPMD) approach, which means the same modeling procedure with different state variables. A

master process aggregates particle statistics and controls resampling/regularization steps. Message passing commands of MPI is used effectively to transfer numerous information of spatially distributed state variables from one particle to another in the resampling step. Elapsed time is about 5 h in SIR, 10 h in lagged SIR and 13 h in lagged RPF for 4 month period simulation with hourly time step, respectively.

4.3   Process and measurement error models

Particle filters perform suboptimal estimation of the system states considering the uncertainty in both measurement and modeling system. Therefore, the choice of the error models is crucial to obtain a better estimation (Weerts and El Serafy, 2006). Another important point is to decide hidden state variables for filtering. As there are numerous state variables in a distributed hydrologic model, it is not practical to consider uncertainty of all state variables with limited number of particles. In that case, it is required to choose limited number of state variables, which process error of the modeling system is aggregated in, and is easy to control an observable state. In this study, we select soil moisture content and overland flow in each grid as hidden state variables and streamflow at the Katsura station as an observable variable. Global multipliers are introduced to perturb state variables stochastically and effectively. In case of soil moisture content, total soil moisture depth at the previous time step S_{k-1} is aggregated for three soil layers within the catchment as:

$$S_{k-1} = \sum_{l=1}^{3} \sum_{j=1}^{m} \theta_j^l d_j^l \tag{19}$$

where θ_j^l and d_j^l are is volumetric soil moisture content ($m^3 m^{-3}$) and depth (m) in the layer l and m represent the number of soil layers and the total number of grids within the catchment, respectively. Then, process noise of soil moisture content w_{soil_k} is added to aggregated variable S_{k-1} as:

$$\hat{S}_k = S_{k-1} + w_{soil_k} \tag{20}$$

w_{soil_k} is assumed as Gaussian distribution $N(0, \sigma^2_{soil_k})$ having a heteroscedastic standard deviation as:

$$\sigma_{soil_k} = \alpha_{soil} S_{k-1} + \beta_{soil} \tag{21}$$

In this study, we apply for 0.05 of α_{soil} and 50 mm of β_{soil}, which means proceess errors of 5% of standard deviation having small constant error. When process error is generated for each particle, it is applied in a multiplicative way using Eqs. (22) and (23).

$$\gamma_s = \frac{\hat{S}_k}{S_{k-1}} \tag{22}$$

$$\hat{\theta}^l_j = \gamma_s \theta^l_j \tag{23}$$

It is worth noting that non-linearity of the distributed hydrologic model can alleviate loss of spatial diversity in the pertubation process, which is one of disadvantages of global multipliers. For example, even if the same noise is applied, spatial pattern of state variables can be different due to the condition of previous time step and non-linear system response for that. The perturbation of overland flow is also applied in a multiplicative way as:

$$\hat{q}_{ov_j} = (1 + w_{ov_k}) q_{ov_j} \tag{24}$$

where process noise of overland flow w_{ov_k} is assumed as Gaussian distribution $N(0, \sigma^2_{ov_k})$. The standard deviation of overland flow noise σ_{ov_k} is parameterized as follows:

$$\sigma_{ov_k} = c_{ov} 10^{\alpha_{ov} \exp(-y_{sim_{k-1}}/\beta_{ov})} \tag{25}$$

where, α_{ov} and β_{ov} are adaptable parameters with setting -10 and $5\,m^3\,s^{-1}$, respectively. $y_{sim_{k-1}}$ is the simulated discharge of data assimilation point at the previous time step. c_{ov} is the constant coefficent. The value of c_{ov} is estimated through the calibration periods and set as 0.03 for the whole simulation. This formulation has been originally

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◄ | ►

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

proposed by Seo et al. (2009) to enhance the forecast in periods of low flow. Equation (25) specifies progressively smaller uncertainty if the simulated flow falls below the threshold, β_{ov} (m^3 s^{-1}). We adopt this error formulation because error of overland flow routing is expected to reduce in low flow periods.

5   The measurement error of the discharge is assumed as Gaussian distribution $N(0, \sigma^2_{obs_k})$ similar to previous studies (Georgakakos, 1986; Weerts and El Serafy, 2006; Salaman and Feyen, 2010). The standard deviation of the measurement error is chosen as:

$$\sigma_{obs_k} = \alpha_{obs} y_k + \beta_{obs} \qquad (26)$$

10 In Eq. (26) α_{obs} is set as 0.1 which means 10% of the measurement error, and the constant coefficient β_{obs} is applied as 5 m^3 s^{-1} to consider uncertainty in periods of low flow such as artificial water use, dam reservior control and etc. The uncertainty of forcing data is not considered in this study to make it easy to evaluate the difference of each particle filter. 15% of perturbation from the uniform distribution is applied for the
15 initial soil moisture condition.

5   Results and discussion

We implement three different versions of particle filters, which are SIR, lagged SIR and lagged RPF, for the streamflow forecasting using the WEP model. Lagged filtering is implemented for SIR and regularized particle filter (RPF). Based on our knowledge on
20 the WEP model behaviour and various tests in the calibration period, the lag time is set as 6 h. Thus, the lagged SIR/RPF starts its prediction from 6 h ago, executing resampling at every updating time step. The lagged RPF has an additional regularization step when sample impoverishment happens or all the particle fall outside of the measurement PDF. For the evaluation of the forecasting accuracy, Nash-Sutcliffe efficiency
25 is calculated as:

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

|◄ | ►|

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

$$E = 1 - \frac{\sum_{k=1}^{T} \left(y_k - y_{\text{sim}_k} \right)^2}{\sum_{k=1}^{T} (y_k - \bar{y})^2} \tag{27}$$

where y is observation, $\bar{y}$ is the mean of observation, y_{sim_k} is the forecasted streamflow at the measurement site and T is the total number of time steps.

As an additional evaluation method on the performances of probabilistic prediction, we use the predictive QQ plot. This method has originated from the economic field (e.g., Diebold et al., 1998) and has been adapted for the verification tools used for probabilistic forecasts of hydrological and meteorological variables (Gneiting et al., 2007; Laio and Tamea, 2007; Thyer et al., 2009; Salaman and Feyen, 2009, 2010). The predictive QQ plot is established by comparing the empirical cumulative density function (CDF) of the sample of p values, which is the probability of "observation" on the CDF of the "prediction", with the CDF of a uniform distribution. It is based on the hypothesis that the p value is a realization from a uniform distribution on $[0, 1]$, under the assumption that the observation y_k is a realization from the predictive distribution. Detailed description on the predictive QQ plot about construction and interpretation can be found in Laio and Tamea (2007) and Thyer et al. (2009).

The diversity of particle system is an important index in a particle filtering process, because the posterior distribution is estimated from the distribution of each particle associated with its weight. This sample impoverishment problem is checked with a ratio of the effective particle number n_{ratio} using Eq. (8).

Figure 7 illustrates the results of the lagged regularized PF using the WEP model during the validation period (1 June to 31 August 2003). Observed discharge is compared with two-step-ahead probablistic forecast of the streamflow at the Katsura station (Fig. 7a). The forcasted streamflow shows good confirmity between observation and simulation, while the deterministic modeling show significant underestimation in the overall rainfall events. Traces of volumetric soil moisture ($m^3 m^{-3}$) is estimated via averaging soil moisture of each soil layer for all grids within the catchment (Fig. 7b). Volumetric soil moisture show sharp rising in the events, followed by fluctuating recession.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

|◀ | ▶|

◀ | ▶

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

The Nash-sutcliffe model efficiency for calibration and validation period is shown in Table 1. In terms of model efficiency, forecast via PFs outperforms deterministic modeling and the accuracy of lagged RPF is higher than other filters in the calibration period.

Figure 8 illustrates forecasted discharge for two rainfall events (8 to 19 August 2003) within the validation period. While the mean values of three particle filters show similar pattern, 90% probabilistic ranges of the streamflow via the SIR particle filter is larger than those via the lagged filtering.

Traces of updated hidden states and their distribution via several PFs are shown in Fig. 9 and the plotted time period (8 to 19 August 2003) is identical to Fig. 8. Behavior of internal state (soil moisture) can be seen through volumetric soil moisture averaged within the catchment (Fig. 9a). Averaged volumetric soil moisture shows similar pattern in three particle filters, whereas the posterior distribution of the state variables via SIR is more diffusive than those via the lagged RPF. Narrow confidential intervals of internal and observable states via the lagged RPF show the enhancement of the probabilistic forecast. On the contrary, wide-spread confidential bands via SIR mean that particles generated from state perturbation are not properly filtered out during the updating step.

As we see through Figs. 8 and 9, in case of probabilistic modeling, we need a tool to check the adequacy of the predictive distribution. The predictive QQ plot is one of alternatives, providing a better summary of the performance of probabilistic predictions than simple hydrograph (e.g., Thyer et al., 2009; Salaman and Feyen, 2009, 2010). Figure 10 show the predictive QQ plots during the calibration and validation period at the Katsura station. Over-estimated predictive uncertainty is clearly seen in the results of SIR, while the results of two lagged filters show slightly under-predicted forecasts. In case of SIR, pertubed internal state of a ditributed hydrologic model cannot be properly updated within one time step interval. Therefore, the uncertainty of internal state diverses until the effect of the perturbation is propogated to the observation station usually located in the downstream. On the other hand, updating can be properly executed in the lagged filtering approach, considering response time of internal hydrologic

processes. However, as estimation accuracy may be lowered in lagged filtering, more advanced calculation of lagged weight are required.

The particle diversity within the simulation is evaluated via ratio of the effective particle number n_{ratio} shown in Fig. 11. The lowest value of the effective particle number represents whether significant loss of sample diversity happens during the simulation. We can see sample impoverishment in case of both lagged and simple SIR filters. As stated by other researches, SIR approach is very prone to sample impoverishment, reducing potentials of filtering. However, the lagged RPF does not show evidence of particle collapse in this implementation, executing the regularization step when a loss of diversity happens. In this study, the threshold for n_{ratio} is set to 0.9. Also note that the RPF can be applicable without a lagged filtering approach but proper particles are unrecognizable in the MCMC move step.

6   Conclusions

A lagged particle filtering approach was proposed to consider different response of internal states in the distributed hydrologic modeling system and the regularized particle filter with the MCMC move step was implemented to preserve sample diversity under the lagged filtering approach. As a process-based distributed hydrologic model, WEP was implemented to enhance forecasting capability of the streamflow.

Two step ahead prediction by particle filters reproduced the streamflow properly compared to that of deterministic modeling in terms of model efficiency index. The enhancement of the probabilistic forecast via the lagged RPF was consistently seen through relatively narrow confidential intervals of internal and observable states. In terms of the predictive QQ plot, lagged filtering is evaluated to have more proper probabilistic bands, whereas SIR reproduced more diffuse probable density. The preservation of particle diversity was shown in the regularized particle filter with MCMC move step.

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

SMC methods have significant potential for high non-linearity problems, especially for process-based distributed models in the hydrologic investigation. However, the computational cost and vague adequacy of SMC for distributed modeling have been bottleneck for their practical implementation. As shown in this study, particle filtering process can be effectively parallelized and implemented in the multi-core computing environment via MPI library. The lagged RPF is expected to be used as one of frameworks for sequential data assimilation of process-based distributed modeling. However, as Weerts and El Serafy (2006) stated, attention also should be focused on development of error models to account for distributed internal processes properly. More advanced treatment of lagged weights and effective sequential estimation method of model parameters are remained as open problems, indeed.

References

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

Clark, P. M., Rupp, D. E., Woods, R. A., Zheng, X., Ibbitt, R. P., Slater, A. G., Schmidt, J., and Uddstrom, M. J.: Hydrological data assimilation with the ensemble Kalman filter: use of streamflow observations to update states in a distributed hydrological model, Adv. Water Resour., 31, 1309–1324, 2008.

Diebold, F. X., Gunther, T. A., and Tay, A. S.: Evaluating density forecasts with applications to financial risk management, Int. Econ. Rev., 39, 863–883, 1998.

Douc, R., Cappe, O., and Moulines, E.: Comparison of resampling scheme for particle filtering, in: Proceedings of the 4th International Symposium on Image and Signal Processing and Analysis, Zagreb, Croatia, 15–17 September 2005, art. no. 1521264, 64–69, 2005.

Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, J. Geophys. Res., 99, 10143–10162, 1994.

Evensen, G.: Data Assimilation: the Ensemble Kalman Filter, Springer, Berlin, 2009.

Georgakakos, K. P.: A generalized stochastic hydrometeorological model for flood and flash-flood forecasting, Water Resour. Res., 22, 2096–2106, 1986.

Gilks, W. R. and Berzuini, C.: Following a moving target – Monte Carlo inference for dynamic Baysian models, J. Roy. Stat. Soc. B, 63, 127–146, 2001.

Gneiting, T., Balabdaoui, F., and Raftery, A. E.: Probabilistic forecasts, calibration and sharpness, J. Roy. Stat. Soc. B, 69, 243–268, 2007.

Gordon, N. J., Salmond, D. J., and Smith, A. F. M.: Novel approach to nonlinear/non-Gaussian Bayesian state estimation, Proc. Inst. Electr. Eng., 140, 107–113, 1993.

Jia, Y. and Tamai, N.: Integrated analysis of water and heat balance in Tokyo metropolis with a distributed model, J. Japan Soc. Hydrol. Water Resour., 11, 150–163, 1998.

Jia, Y., Ni, G., Kawahara, Y., and Suetsugi, T.: Development of WEP model and its application to an urban watershed, Hydrol. Process., 15, 2175–2194, 2001.

Jia, Y., Ding, X., Qin, C., and Wang, H.: Distributed modeling of landsurface water and energy budgets in the inland Heihe river basin of China, Hydrol. Earth Syst. Sci., 13, 1849–1866, doi:10.5194/hess-13-1849-2009, 2009.

Johansen, A. M.: SMCTC: sequential Monte Carlo in C++, J. Stat. Software, 30, 1–41, 2009.

Karssenberg, D., Schmitz, O., Salaman, P., de Jong, K., and Bierkens, M. F. P.: A software framework for construction of process-based stochastic spatio-temporal models and data assimilation, Environ. Model. Softw., 25, 489–502, 2010.

Kim, H., Noh, S., Jang, C., Kim, D., and Hong, I.: Monitoring and analysis of hydrological cycle of the Cheonggyecheon watershed in Seoul, Korea, in: Proceedings of International Conference on Simulation and Modeling, Bangkok, Thailand, 17–19 January 2005, C4-03, 2005.

Kitagawa, G.: Monte-Carlo filter and smoother for non-Gaussin non-linear state-space models, J. Comput. Graph. Stat., 5, 1–25, 1996.

Kitanidis, P. K. and Bras, R. L.: Real-time forecasting with a conceptual hydrologic model 2. applications and results, Water Resour. Res., 16, 1034–1044, 1980.

Kong, A., Liu, J. S., and Wong, W. H.: Sequential imputations and Bayesian missing data problems, J. Am. Stat. Assoc., 89, 278–288, 1994.

Laio, F. and Tamea, S.: Verification tools for probabilistic forecasts of continuous hydrological variables, Hydrol. Earth Syst. Sci., 11, 1267–1277, doi:10.5194/hess-11-1267-2007, 2007.

Leisenring, M. and Moradkhani, H.: Snow water equivalent prediction using Bayesian data as-

simialtion methods, Stoch. Environ. Res. Risk Assess., 25(2), 253–270, doi:10.1007/s00477-010-0445-5, 2010.

Liu, Y. and Gupta, H. V.: Uncertainty in hydrologic modeling: toward an integrated data assimilation framework, Water Reour. Res., 43, W07401, doi:10.1029/2006WR005756, 2007.

5   Montanari, M., Hostache, R., Matgen, P., Schumann, G., Pfister, L., and Hoffmann, L.: Calibration and sequential updating of a coupled hydrologic-hydraulic model using remote sensing-derived water stages, Hydrol. Earth Syst. Sci., 13, 367–380, doi:10.5194/hess-13-367-2009, 2009.

Moradkhani, H.: Hydrologic remote sensing and land surface data assimilation, Sensors, 8,
10   2986–3004, doi:10.3390/s8052986, 2008.

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, 2005a.

Moradkhani, H., Sorooshian, S., Gupta, H. V., and Houser, P. R.: Dual state-parameter estimation of hydrological models using ensemble Kalman filter, Adv. Water Resour., 28, 135–147,
15   2005b.

Musso, C., Oudjane, N., and LeGland, F.: Improving regularized particle filters, in: Sequential Monte Carlo in Practice, edited by: Doucet, A., de Freitas, N., and Gordon, N., Springer-Verlag, New York, 247–271, 2001.

20   Noh, S. J., Tachikawa, Y., Shiiba, M., and Kim, S.: Dual state-parameter updating scheme on a conceptual hydrologic model using sequential Monte Carlo filters, Ann. J. Hyd. Engin. JSCE, 55, 1–6, 2011.

Qin, C., Jia, Y., Su, Z., Zhou, Z., Qiu, Y., and Shen, S.: Integrating remote sensing information into a distributed hydrological model for improving water budget predictions in largescale
25   basins through data assimilation, Sensors, 8, 4441–4465, doi:10.3390/s8074441, 2008.

Qin, J., Liang, S., Yang, K., Kaihotsu, I., Liu, R., and Koike, T.: Simultaneous estimation of both soil moisture and model parameters using particle filtering method through the assimilation of microwave signal, J. Geophys. Res., 114, D15103, doi:10.1029/2008JD011358, 2009.

Reichle, R. H., Koster, R. D., Entekhabi, D.: Variational data assimilation of microwave radio-
30   brightness observations for land surface hydrology applications, IEEE T. Geosci. Remote, 39, 1708–1718, 2001.

Rings, J., Huisman, J. A., and Vereecken, H.: Coupled hydrogeophysical parameter estimation using a sequential Bayesian approach, Hydrol. Earth Syst. Sci., 14, 545–556,

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◄◄ | ►►

◄ | ►

Back | Close

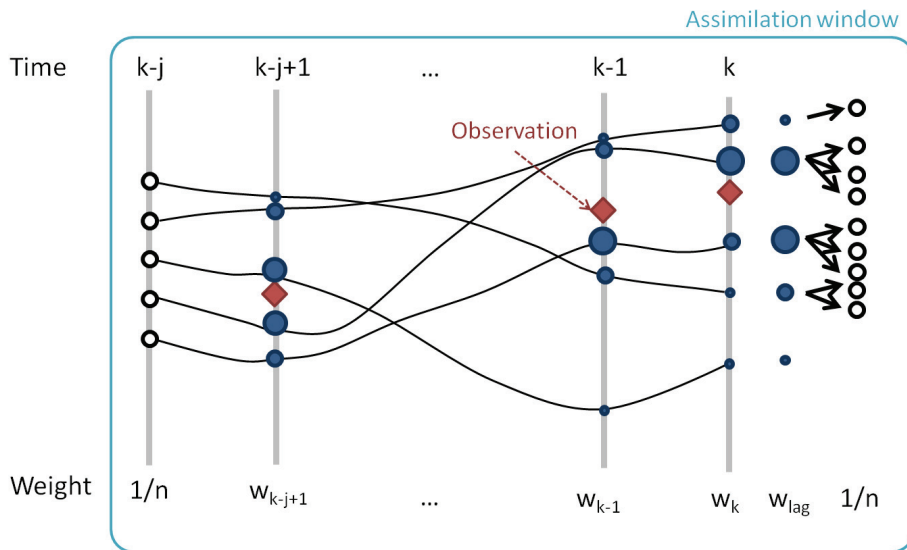Full Screen / Esc

Printer-friendly Version

Interactive Discussion

doi:10.5194/hess-14-545-2010, 2010.

Ristic, B., Arulampalam, S., and Gordon, N.: Beyond the Kalman Filter: Particle Filters for Tracking Applications, Artech House, London, UK, 2004.

Robert, C. P. and Casella, G.: Monte Carlo Statistical Methods, Springer, New York, 1999.

Salaman, P. and Feyen, L.: Assessing parameter, precipitation, and predictive uncertainty in a distributed hydrological model using sequential data assimilation with the particle filter, J. Hydrol., 376, 428–442, 2009.

Salaman, P. and Feyen, L.: Disentangling uncertainties in distributed hydrological modeling using multiplicative error models and sequential data assimilation, Water Resour. Res., 46, W12501, doi:10.1029/2009WR009022, 2010.

Schaap, M. G., Leij, F. J., and van Genuchten, M. T.: ROSETTA: a computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions, J. Hydrol., 251, 163–176, 2001.

Seo, D.-J., Koren, V., and Cajina, N.: Real-time variational assimilation of hydrologic and hydrometeorological data into operational hydrologic forecasting, J. Hydrometeorol., 4, 627–641, 2003.

Seo, D.-J., Cajina, L., Corby, R., and Howieson, T.: Automatic state updating for operational streamflow forecasting via variational data assimilation, J. Hydrol., 367, 255–275, 2009.

Smith, P. J., Beven, K. J., and Tawn, J. A.: Detection of structural inadequacy in process-based hydrological models: a particle-filtering approach, Water Resour. Res., 44, W01410, doi:10.1029/2006WR005205, 2008.

Thyer, M., Renard, B., Kavetski, D., Kuczera, G., Franks, S. W., and Srikanthan, S.: Critical evaluation of parameter consistency and predictive uncertainty in hydrologic modeling: a case study using Bayesian total error analysis, Water Resour. Res., 45, W00B14, doi:10.1029/2008WR006825, 2009.

van Delft, G., El Serafy, G. Y., and Heemink, A. W.: The ensemble particle filter (EnPF) in rainfall-runoff models, Sotch. Environ. Risk Assess., 23, 1203–1211, 2009.

van Leeuwen, P. J.: Particle filtering in geophysical systems, Mon. Weather Rev., 137, 4089–4114, 2009.

Vrugt, J. A., Gupta, H. V., Nuallain, B. O., and Bouten, W.: Real-time data assimilation for operational ensemble streamflow forecasting, J. Hydrolmeteorol., 7, 548–565, 2006.

Weert, 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,

Applying sequential Monte Carlo methods into a DHM

S. J. Noh et al.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◀◀ | ▶▶

◀ | ▶

Back | Close

Full Screen / Esc
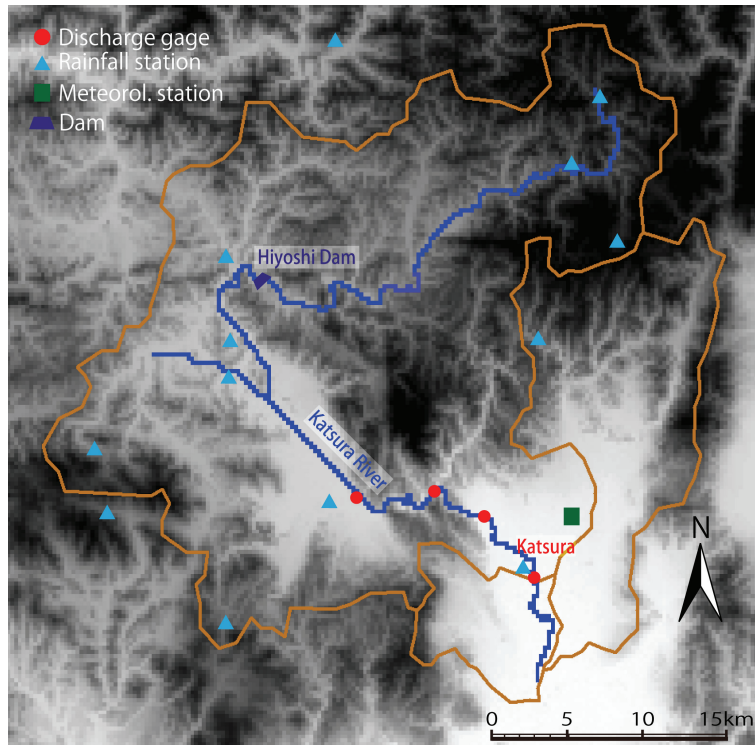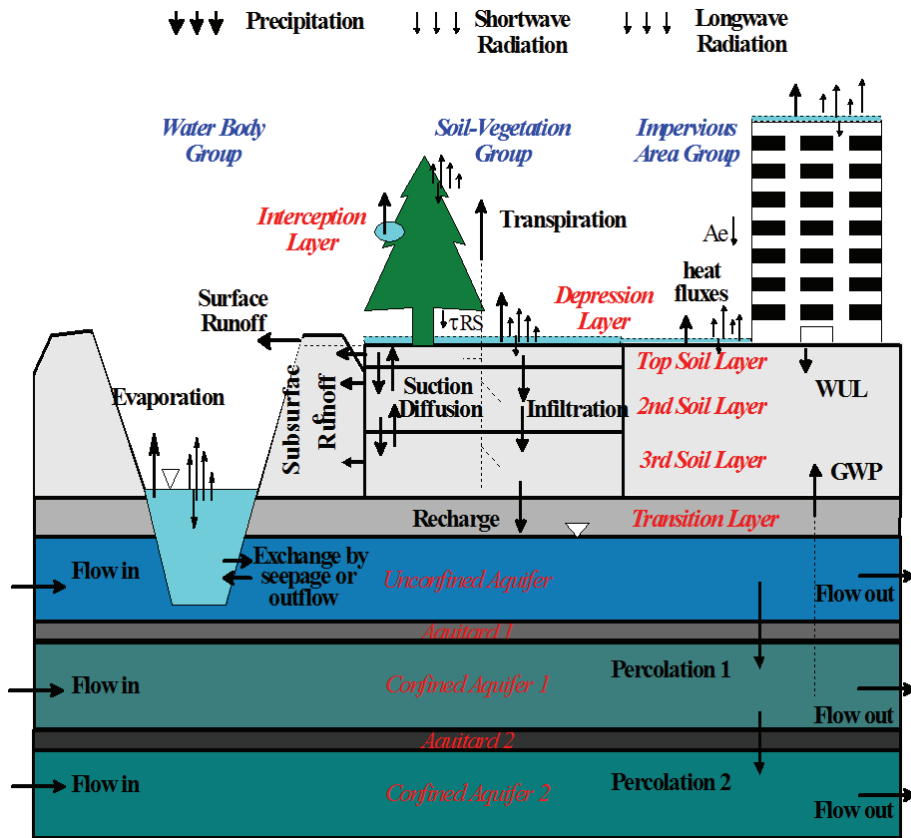
Printer-friendly Version

Interactive Discussion

doi:10.1029/2005WR004093, 2006.

Zhou, Y., McLaughlin, D., and Entkhabi, D.: Assessing the performance of the ensemble Kalman filter for land surface data assimilation, Mon. Weather Rev., 134, 2128–2142, 2006.

Applying sequential Monte Carlo methods into a DHM

S. J. Noh et al.

Title Page

Abstract	Introduction
Conclusions	References
Tables	Figures

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Table 1. Nash-Sutcliffe model efficiency.

Methods	1 Jun to 31 Jul 2007	1 Jun to 31 Aug 2003
Lagged RPF	0.971	0.982
Lagged SIR	0.969	0.982
SIR	0.966	0.985
Deteministic	0.875	0.704
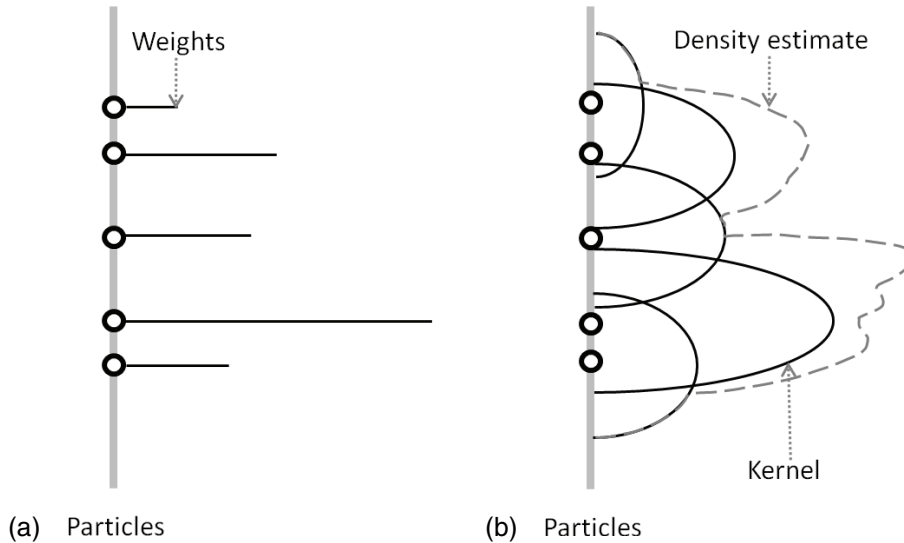
Fig. 1. Regulalized particle filter. **(a)** Weighted empirical measure. **(b)** Regularized measure by kernel. Adapted from Musso et al. (2001).

Fig. 2. The concept of a lagged particle filtering approach.
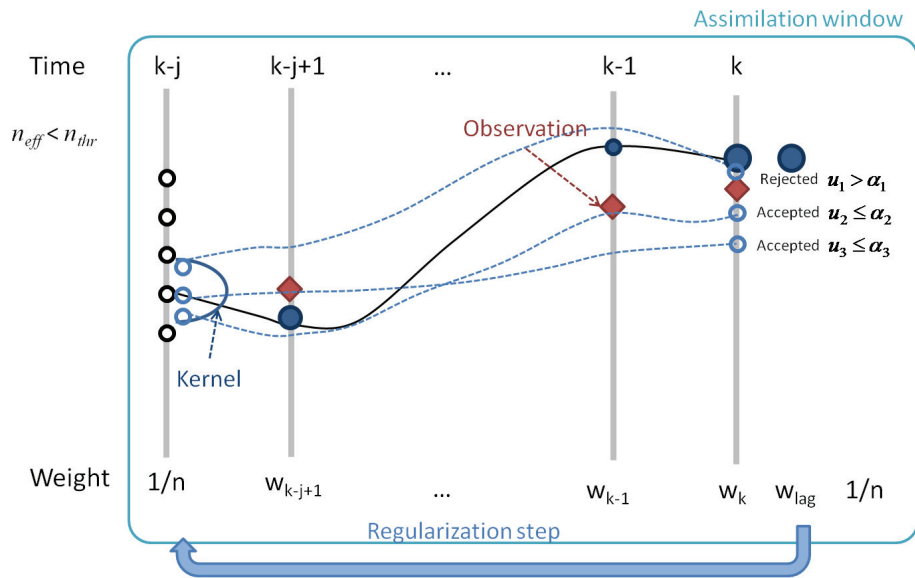
Fig. 3. Particle traces in the regularized step with the MCMC move step under the lagged filtering approach.
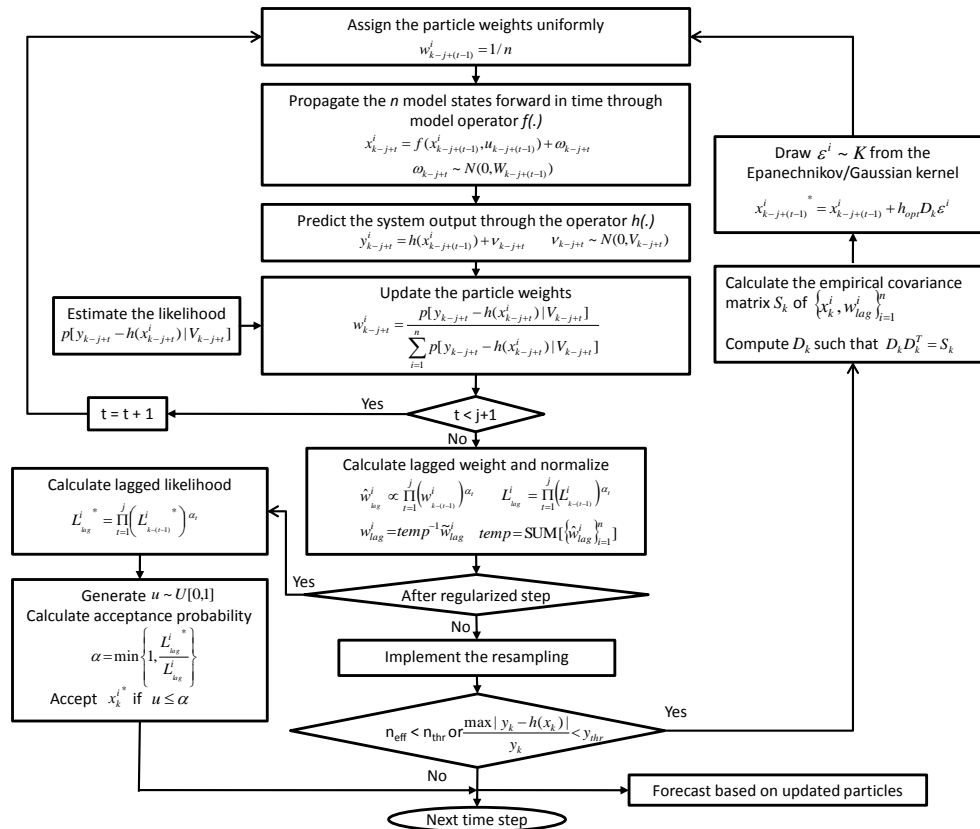
Fig. 4. The flowchart of the regularized particle filter with the MCMC move step in the lagged filtering approach.

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Fig. 5. The Katsura River catchment.

Fig. 6. Vertical structure of WEP model. Adapted from Jia et al. (2009).

Title Page

Abstract | Introduction

Conclusions | References

Tables | Figures

◄◄ | ►►

◄ | ►

Back | Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

Fig. 7. Results of the lagged regularized PF using the WEP model by assimilating hourly streamflow of the Kastura gauaing station during the validation period (1 June to 31 August 2003). **(a)** Observed versus 2-h-lead forecast at the Katsura station. The black dots represent observed discharge. The blue line and area represent mean value and 90% confidential interval, respectively. A dashed line represents a deterministic modeling case. **(b)** Traces of volumetric soil moisture estimated from total soil moisture depth within the catchment. The blue line and area represent mean value and 90% confidential interval, respectively.
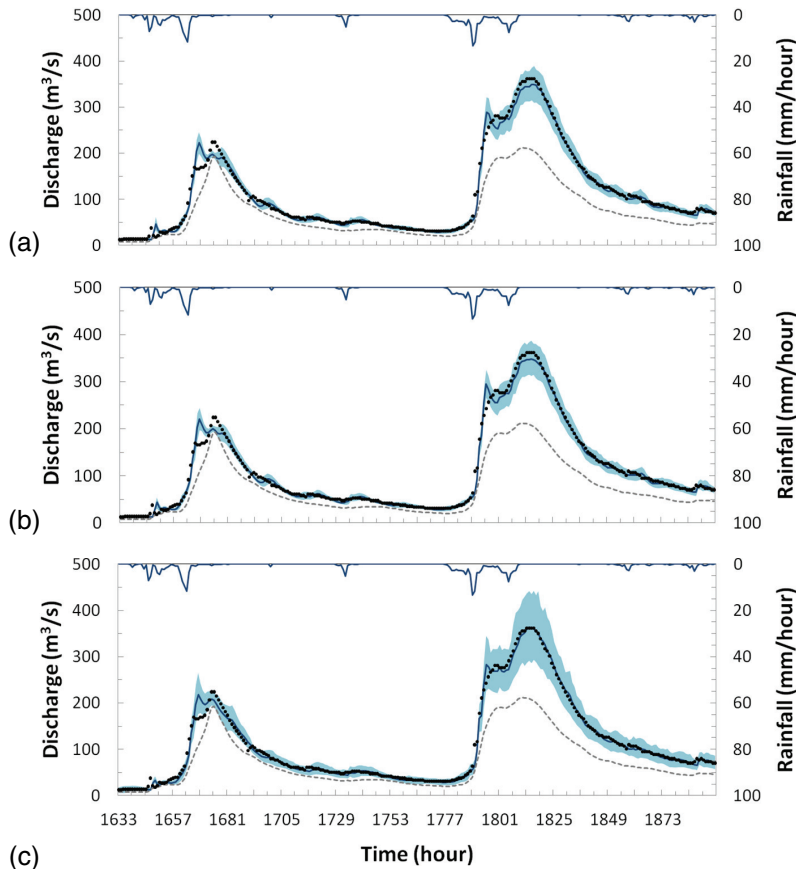
Fig. 8. Observed versus 2-h-lead forecast at the Katsura station via several PFs (8 to 19 August 2003). **(a)** The lagged RPF. **(b)** The lagged SIR. **(c)** The SIR. A dashed line represents a deterministic modeling case. The blue line and area represent mean value and 90% confidential interval, respectively.
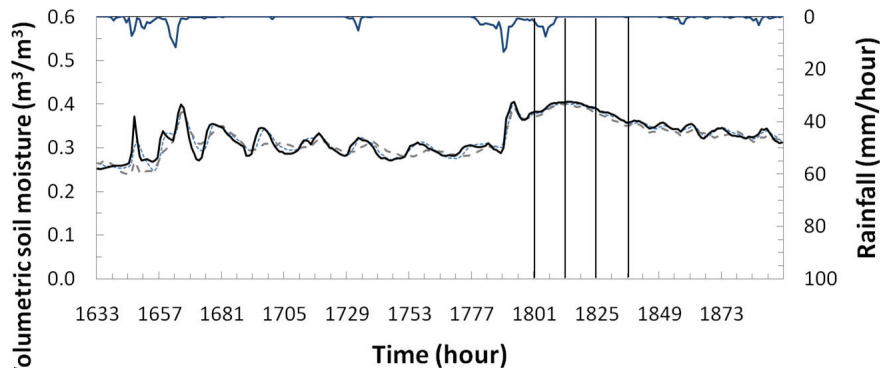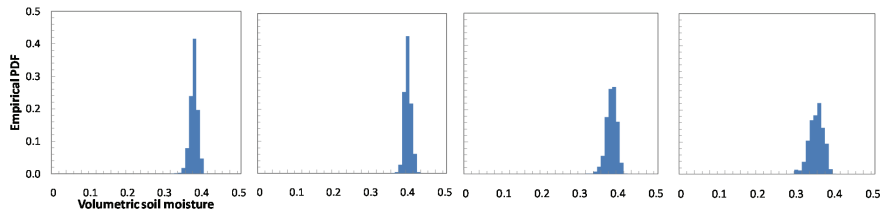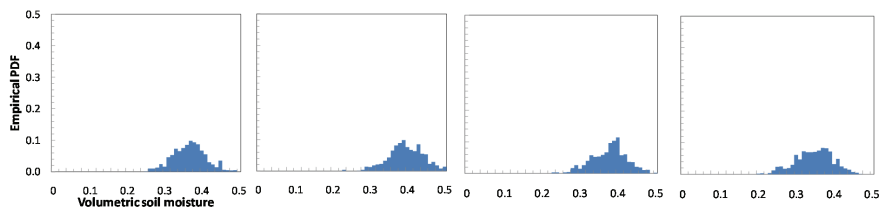
Fig. 9. Results of several PF (8 to 19 August 2003). **(a)** Traces of volumetric soil moisture estimated from total soil moisture depth within the catchment. The black, grey dashed and blue dalshed lines represent mean value of lagged RPF, lagged SIR and SIR, respectively. **(b)** and **(c)** represent posterior distributions of volumetric soil moisture of lagged RPF **(b)** and SIR **(c)** at four time point of **(a)** (black bars).
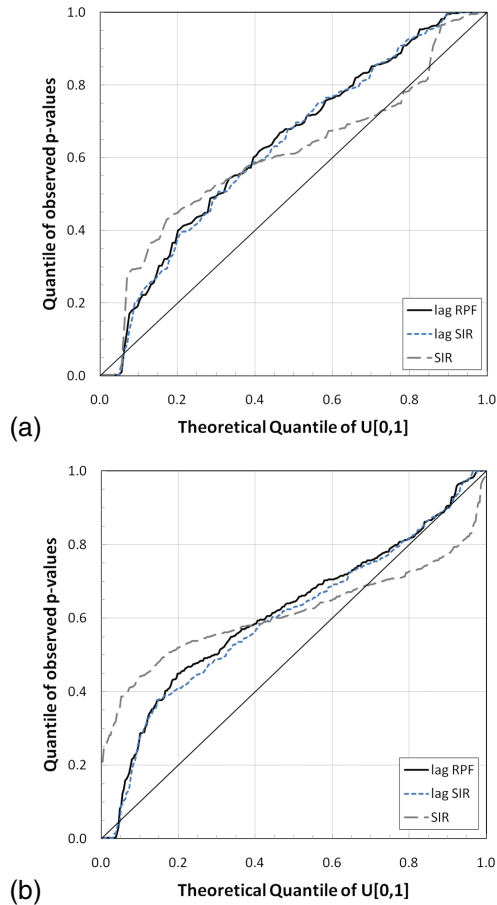
Fig. 10. The predictive QQ plot for 2-h-lead forecast via several particle filters. **(a)** 12 to17 July 2007 (within the calibration period). **(b)** 8 to 19 August 2003 (within the validation period).
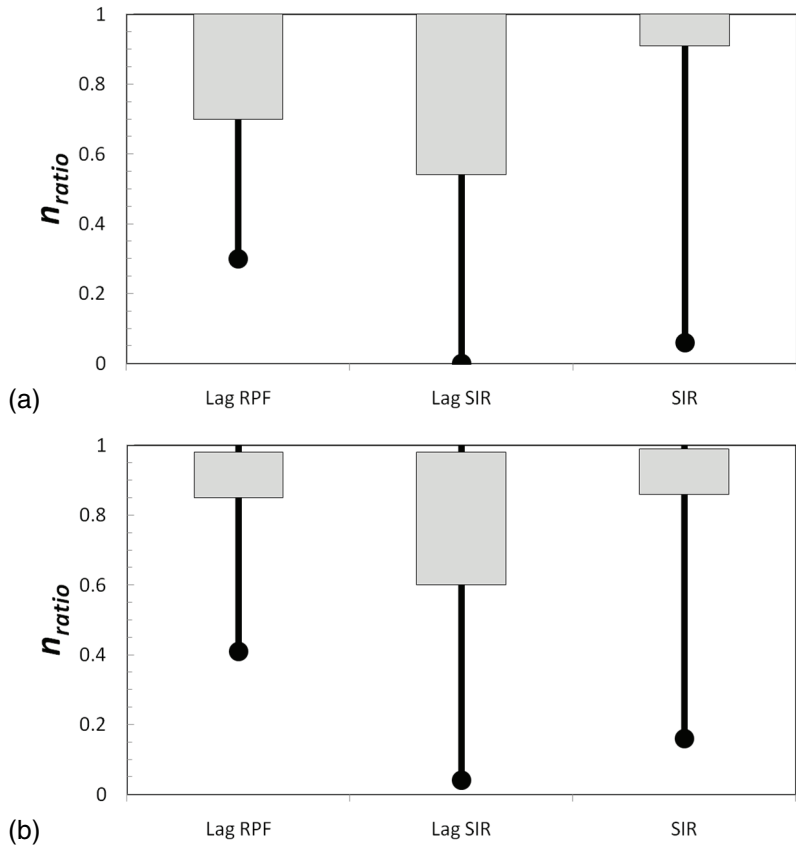
Fig. 11. Distributions of ratio of the effective particle number n_{ratio}. **(a)** The calibration period (1 June to July 2007). **(b)** The validation period (1 June to 31 August 2003). Black lines represent the maximum and minimum bounds of n_{ratio}. The grey boxes represent 90% bounds of n_{ratio}.