



Hydrological data assimilation with the Ensemble Square-Root-Filter: Use of streamflow observations to update model states for real-time flash flood forecasting



He Chen^a, Dawen Yang^{a,*}, Yang Hong^{b,c}, Jonathan J. Gourley^d, Yu Zhang^{b,c}

^a State Key Laboratory of Hydrosience and Engineering, Department of Hydraulic Engineering, Tsinghua University, Beijing 100084, China

^b School of Civil Engineering and Environmental Sciences, University of Oklahoma, Norman, OK 73072, USA

^c Hydrometeorology and Remote Sensing Laboratory, National Weather Center Atmospheric Radar Research Center, Norman, OK 73072, USA

^d NOAA/National Severe Storms Laboratory, Norman, OK 73072, USA

ARTICLE INFO

Article history:

Received 28 February 2012

Received in revised form 28 February 2013

Accepted 17 June 2013

Available online 27 June 2013

Keywords:

Data assimilation

Ensemble Kalman filter

Ensemble Square-Root-Filter

Flash flood forecast

Rainfall–runoff model

ABSTRACT

The objective of the study is to evaluate the potential of a data assimilation system for real-time flash flood forecasting over small watersheds by updating model states. To this end, the Ensemble Square-Root-Filter (EnSRF) based on the Ensemble Kalman Filter (EnKF) technique was coupled to a widely used conceptual rainfall–runoff model called HyMOD. Two small watersheds susceptible to flash flooding from America and China were selected in this study. The modeling and observational errors were considered in the framework of data assimilation, followed by an ensemble size sensitivity experiment. Once the appropriate model error and ensemble size was determined, a simulation study focused on the performance of a data assimilation system, based on the correlation between streamflow observation and model states, was conducted. The EnSRF method was implemented within HyMOD and results for flash flood forecasting were analyzed, where the calibrated streamflow simulation without state updating was treated as the benchmark or nature run. Results for twenty-four flash-flood events in total from the two watersheds indicated that the data assimilation approach effectively improved the predictions of peak flows and the hydrographs in general. This study demonstrated the benefit and efficiency of implementing data assimilation into a hydrological model to improve flash flood forecasting over small, instrumented basins with potential application to real-time alert systems.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Flooding is one of the most devastating natural hazards across the world, causing tremendous economic losses and casualties. Flood forecasting is one of the crucial issues in hydrology, as well as a challenging problem in operational practice, especially over ungauged small river basins [1]. Two factors affecting flood forecasting are quantitative precipitation estimation and forecasting and streamflow simulation by rainfall–runoff models [2]. A great number of quantitative precipitation estimation methods, numerical weather forecast systems, and hydrologic models have been developed for the purpose of improving streamflow simulation and ultimately, flood forecasting.

However, uncertainties in the forcing data, observed system response, and imperfect model structures are inevitably involved in flood forecasting [3,4]. First, hydrologic models contain errors due to the simplification and parameterization of real world pro-

cesses, despite optimizing parameter settings using automatic or manual calibration methods. Secondly, the difficulty in flood forecasting is largely due to measurement errors of physical quantities, particularly precipitation and streamflow. For flash flood cases, the precipitation forcing data is arguably among the most important factors. Quantitative precipitation estimation and forecast products are subject to the most error during extreme rainfall events over small-sized watersheds. Another limitation in flash flood forecasting is the sensitivity of rainfall–runoff models to initial conditions given short duration, intense storms because of nonlinear threshold effects and runoff responses. In this regard, data assimilation techniques can update estimated model states by jointly taking into account model errors, forcing data uncertainties and output uncertainties, hence to improve streamflow forecasting [5].

Among a range of data assimilation techniques, the Kalman filter (KF) developed by Kalman [6] is the most well-known. The traditional KF method was restricted to linear systems, and was later extended to nonlinear models as an extended Kalman Filter (EKF). EKF method was introduced to hydrologic models for flood forecasting in the 1980s (e.g. [7–9]), but the computational demand

* Corresponding author. Tel.: +86 10 62796976; fax: +86 10 62796971.

E-mail address: yangdw@tsinghua.edu.cn (D. Yang).

resulting from the error covariance integration and the instability in cases of strong model non-linearity limited the applicability of the EKF [10,11]. Ensemble Kalman filter (EnKF) was thus developed to circumvent these problems by performing a Monte Carlo ensemble of model runs instead of reforming the model into state-space form [12]. A series of comparison studies showed the EnKF method had the advantage of easy implementation, flexibility in covariance modeling, robustness and computational efficiency (e.g. [13–15]). The EnKF has great potential in flood forecasting and thus has emerged as the most popular choice for hydrologic model assimilation. Whitaker and Hamill [16] proposed the Ensemble Square-Root-Filter (EnSRF), which is a variant of EnKF that does not require the observation to be perturbed as the standard implementation of EnKF. The EnSRF algorithm has been demonstrated to be as fast as the EnKF, and more accurate than EnKF for a given ensemble size.

One research theme of data assimilation is diagnosing the uncertainties present in the model itself and observations, and then transferring this knowledge to improve the understanding and representation of model concepts describing the physical hydrologic system. Vrugt et al. [17] presented a simultaneous optimization and data assimilation approach to better treat parametric uncertainty by integrating all sources of uncertainties using the EnKF method. Input, output and model structural errors are distinguished and quantified during the model calibration process. Another important use for data assimilation is predicting future conditions by incorporating knowledge of current system states.

Recently, the potential of data assimilation in streamflow forecasting has motivated an increasing number of studies. For example, Aubert et al. [18] studied the possibility for improving streamflow forecasting by coupling soil moisture observations with the routing function of a hydrologic model using the EKF method. Weerts and El Serafy [15] tested the applicability of EnKF for the correction of the model inputs for flood forecasting purposes. Komma et al. [19] implemented the EnKF method for flood forecasting in Austria. Thirel et al. [20] described an assimilation system in France that utilized past discharges in order to obtain a better initial state. Very few operational applications of such assimilation systems exist for real-time flash-flood forecasting over small watersheds. Unlike the situation over large scale or meso-scale basins, flash floods are usually caused by intense, local storms. First, the odds of there being an observation of streamflow are slimmer compared to larger rivers. Second, the quickness of the hydrologic response to rainfall limits the effectiveness to obtain observations of models states and incorporate them in future predictions.

The objective of the study is to examine the potential of a data assimilation system for real-time flash flood forecasting over small watersheds through the applicability of the EnSRF method. A relatively simple but robust conceptual rainfall-runoff Hydrologic MODel called HyMOD [21] was used because it has shown its effectiveness in flood forecasting in an operational context, and is especially well adapted to real time forecasting (e.g. [22,23]). In the following sections, we describe the study watersheds, data, HyMOD structure, EnSRF method, implementation of the proposed data assimilation methodology, evaluation criteria, and model calibration. Then we analyzed the modeling and observation uncertainties and the determined the proper ensemble size, which is a key factor in the EnSRF implementation. The model states-updating experiment was followed by a case study of the application over two small watersheds in America and China.

2. Methodology

2.1. Study watersheds and data

Two small watersheds, Cobb Creek watershed and Chuzhou watershed located in America and China respectively were used in the

study to evaluate the usefulness of data assimilation for flash-flood forecasting. Cobb Creek watershed, the largest sub-basin of Ft. Cobb watershed (see [24, Fig. 1]), is located in southwestern Oklahoma in Caddo, Washita and Custer counties. Chuzhou watershed is one of the sub-basins of Suichuanjiang river basin, which is located in the Yangtze River Basin of Southern China (see [25, Fig. 3]). The characteristics of the two watersheds are provided below in Table 1.

Both watersheds share a similar basin area and temperate climate, and the rainy seasons are during the warm season. Typically, the wettest and warmest month of Cobb Creek watershed is May, which has 150 mm of precipitation, while in Chuzhou watershed, 70% of the annual rainfall occurs from Apr to Sep. However, remarkable differences are shown with the two watersheds, such as annual precipitation, land uses, and hill slope. These differences may cause variation in flood hydrographs between the two watersheds, which can provide more insights in the forthcoming flood-forecasting analysis.

Available data in the Cobb Creek watershed include precipitation data from USDA Agricultural Research Service (ARS), discharge from a USGS stream gauge, meteorology data from Micronet and Mesonet stations. Hourly precipitation data is available from fifteen stations located within or around the Cobb Creek watershed from May 2005 operated by the USDA ARS Grazing lands Research Laboratory. River discharge is measured by the USGS streamflow gauge (07305800, Cobb Creek near Eakly) at 15-min intervals.

Data used in Chuzhou watershed include hourly precipitation interpolated from 7 rain gauges within the watershed, daily pan evaporation data from a nearby meteorological station approximately 60 km away from the center of the watershed, hourly discharge at the outlet of the watershed. Hourly data from 16 large flood events occurring from 1981 to 2002 were used in this study. Of these events, 4 are used for calibration and the remaining 12 are used for validation.

2.2. Hydrologic model

To implement and evaluate the data assimilation approach in rainfall-runoff process simulation and forecast, the conceptual Hydrologic MODel (HyMOD) described by Boyle [21] was used in the study. HyMOD is a rainfall excess model using a nonlinear tank connected with a series of linear tanks (three identical quick-flow tanks) in parallel representing the surface flow to a slow-flow tank representing groundwater flow. Five parameters in the HyMOD system are C_{max} : the maximum storage capacity within the watershed, b_{exp} : the degree of spatial variability of the soil moisture capacity within the watershed, α : a factor for partitioning the flow between two series of tanks, R_q and R_s : the residence time parameters of quick-flow and slow-flow tanks, respectively. The model has five state variables, S : storage in the nonlinear tank representing the soil moisture content in the watershed, x_1 , x_2 and x_3 : the quick-flow tank storages representing the temporary (short-time) detentions, x_4 : the slow-flow tank storage (subsurface storage). Additional details about the model structure are available in Boyle [21] and Wagener [30]. Primarily attributed to the fast computation speed

Table 1
Characteristics of Cobb Creek watershed and Chuzhou watershed.

	Cobb Creek watershed	Chuzhou watershed [29]
(Latitude, longitude)	(35.29° N, 98.64° W)	(26.35° N, 114.13° E)
Basin area	342 km ²	289 km ²
Annual precipitation	820 mm [26]	1550 mm
Predominant land use	Cropland/grassland [27]	Forest
Predominant soil type	Silt loam/loam [28]	Loam
Elevation range	379–564 m a.s.l.	379–2090 m a.s.l.

and simple model structure, HyMOD has been used in previous research studies of data assimilation such as in Moradkhani [22] and Blasone [31].

2.3. Ensemble Square-Root-Filter

The data assimilation approach used in this study was based on the ensemble Kalman filter (EnKF) concept introduced by Evensen [12]. The basic idea of EnKF incorporates an ensemble of forecasts to estimate background error covariance, which is a Monte Carlo approximation to the traditional Kalman filter [32]. The update equation of EnKF is

$$\mathbf{x}_i^a = \mathbf{x}_i^b + \mathbf{K}(\mathbf{y}_i - \mathbf{H}\mathbf{x}_i^b), \quad i = 1 \dots m \quad (1a)$$

where

$$\mathbf{K} = \mathbf{P}^b \mathbf{H}^T (\mathbf{H} \mathbf{P}^b \mathbf{H}^T + \mathbf{R})^{-1} \quad (1b)$$

is the Kalman gain. In Eq. (1), \mathbf{x}_i^b is the background vector of all model states for each ensemble member before the update ($n \times 1$ dimension), \mathbf{x}_i^a is the analysis after the update ($n \times 1$ dimension), \mathbf{y}_i is the vector of the observations ($p \times 1$ dimension), \mathbf{P}^b is the background error covariance matrix ($n \times n$ dimension), \mathbf{H} is the operator that converts the model states to the observation space ($p \times n$ dimension), \mathbf{R} is the observation error covariance matrix ($p \times p$ dimension), m is the number of ensemble members, n is the dimension of model states, and p is the dimension of observations.

In the standard implementation of EnKF, each i of the m ensemble members is updated by Eq. (1). For each ensemble member, the p dimensional observation \mathbf{y}_i is sampled from a distribution with the mean equal to the observation and a variance of \mathbf{R} . Burgers et al. [33] demonstrated that the perturbation of the observation is actually necessary to provide the correct analysis error covariance; otherwise, the analysis error covariance will be underestimated.

To simplify the standard implementation of EnKF, Whitaker and Hamill [16] proposed another ensemble filtering algorithm called Ensemble Square-Root-Filter (EnSRF) which does not require the perturbation of observations. They demonstrated the new algorithm is as fast as EnKF and is more accurate than EnKF for a given ensemble size. The basic idea of EnSRF is to update model states by updating the ensemble mean and a deviation from the mean separately with update equations

$$\begin{aligned} \bar{\mathbf{x}}^a &= \bar{\mathbf{x}}^b + \mathbf{K}(\bar{\mathbf{y}} - \mathbf{H}\bar{\mathbf{x}}^b) \\ \mathbf{x}_i^{\prime a} &= \mathbf{x}_i^{\prime b} + \tilde{\mathbf{K}}(\mathbf{y}_i - \mathbf{H}\mathbf{x}_i^{\prime b}) \end{aligned} \quad (2)$$

where the overbar denotes the ensemble mean and the prime denotes the deviation from the ensemble mean, $\tilde{\mathbf{K}}$ is the reduced Kalman gain identified as

$$\tilde{\mathbf{K}} = \mathbf{P}^b \mathbf{H}^T \left[\left(\sqrt{\mathbf{H} \mathbf{P}^b \mathbf{H}^T + \mathbf{R}} \right)^{-1} \right]^T \times \left[\sqrt{(\mathbf{H} \mathbf{P}^b \mathbf{H}^T + \mathbf{R})} + \sqrt{\mathbf{R}} \right]^{-1} \quad (3)$$

Note that no perturbed observation is needed in the update equation ($\mathbf{y}' = 0$), and \mathbf{K} in Eq. (2) is the traditional Kalman gain given by Eq. (1b). The ensemble mean and the deviation of the model states are updated by Eq. (2). The update equation of EnSRF is finally written as

$$\mathbf{x}_i^a = \bar{\mathbf{x}}^a + \mathbf{x}_i^{\prime a} \quad (4)$$

Clark et al. [34] used both standard implementation of EnKF and EnSRF in a distributed hydrologic model. They demonstrated that the model simulations improved by using the latter approach. In this study, we used EnSRF instead of the standard implementation of EnKF, as it has been shown that the algorithm doesn't lose the

accuracy of analysis-error covariance estimation without the perturbation of observations. Detailed algorithm and equations are shown in Whitaker and Hamill [16, Eqs. (1)–(13)].

2.4. Implementation of EnSRF

A key issue in implementing a data assimilation method is quantifying the covariance error matrices. For the ensemble Kalman filter, variance between multiple ensemble members is used to quantify the model error. There are a number of possibilities to generate ensemble members that are related to formulating the model errors. The easy and robust way is to perturb forcing data and model states separately, because the model error is the combination of multiple sources. With this, the physical basis and interpretation of individual error sources remains clear [19]. In this study, we didn't perturb the model parameters as recommended by Moradkhani et al. [22], which implied the model error is attributed to input error and model states estimation error. This simplification makes a compromise between physical representativeness and computational efficiency, and has been widely adopted for real-time data assimilation for streamflow forecasting [35].

Forcing data uncertainty was emulated by perturbing hourly rainfall amounts at each time step. As mentioned in van Delft et al. [23], HyMOD is less sensitive to evapotranspiration compared to precipitation. So, in this study, evapotranspiration was not perturbed. Ensembles of rainfall forcing were generated as [34]:

$$p' = p(1 + \gamma_p), \gamma_p \sim N(0, \omega_p) \quad (5)$$

where γ_p is the random noise with covariance ω_p , such that γ_p is a realization from a normal distribution ranging from $-\omega_p$ to ω_p .

Model structural uncertainty was simulated by adding noise to the model state transition equations [22]:

$$\mathbf{x}_t = \mathbf{M}_x(\mathbf{x}_{t-1}, \mathbf{u}_t, \theta) + \gamma_t; \gamma_t \sim N[0, \omega_x] \quad (6)$$

where \mathbf{x}_t is a vector of model states at given time t , \mathbf{u} is forcing data, and θ is time-invariant model parameters. Random perturbations to the model states were drawn from zero-mean Gaussian distributions with heteroscedastic variance of:

$$\omega_x = I_D 0.05(\mathbf{x}_{t-1}) \quad (7)$$

where I_D is the D dimension identity matrix. Perturbations were sampled independently across states and time steps.

The data assimilation approach implicates an assumption of knowing, or at least accurately estimating, model uncertainty and observation error; both present challenges. The EnSRF dynamically derives such information by Monte Carlo sampling through assigned random noise in the forcing data, model structure, and observation. Crow and Loon [36] pointed out that inaccurate model error and observation error assumptions would result in deterioration of forecasting using data assimilation. Thus, experiments need to be conducted to test the sensitivity of the data assimilation approach to model uncertainty and observation errors in streamflow estimation.

One merit of the EnSRF method compared to EnKF is that only the magnitude of observational uncertainty is needed, and perturbing the observations can be omitted. The observation of streamflow used for updating model states is derived from USGS observations, with an assumed observation error of η . In this experiment, we tested the sensitivity of the data assimilation performance to different magnitudes of η and ω_p . Each assimilation cycle was repeated twenty times with different values of η and ω_p with an ensemble size of 40.

After the appropriate modeling uncertainty and observation errors were determined, another question in implementing the

data assimilation method was the proper ensemble size. Ensemble size represents a balance between sampling error and computational expense. Effects of varying ensemble size were evaluated using ensembles from 10 to 80 members at increments of 10. An appropriate ensemble size is determined when model performance stabilizes and reaches an asymptote with increasing ensemble size.

Theoretically, a data assimilation method is only useful if a significant correlation exists between the model states being updated and the observations being assimilated. The updating process might introduce errors and result in a worsening in the accuracy of the forecasts when the correlation is weak. The HyMOD quick-flow tanks are hypothesized to have strong correlation with streamflow observations, but a simulation study is needed to evaluate the data assimilation performance by updating different combinations of model states. In the experiment, we first quantified the correlation between all the five model states with streamflow observations. Then the states were separated into three groups: soil water tank (S), quick-flow tank (x_1, x_2, x_3), and slow-flow tank (x_4). Each of the model states were updated separately and then updated in unison.

Observed precipitation data was used as model forcing, with perturbations applied based on the simple model in Eq. (5). A much more complex error model would need to be incorporated in the case of using quantitative precipitation forecasts as model forcing rather than gauge-based estimates. This would add additional uncertainty to the forecasts, but the benefit would be longer lead times for the impending floods. The implementation didn't involve additional sources of error to the forcing data, which is particularly useful for assessing the effect of data assimilation. The benchmark run was the open loop flood forecasting without using data assimilation.

To test the performance of data assimilation in flash flood forecasting, twelve events from each basin were chosen to be examined in more detail in this study (see Tables 2 and 3). Among those twenty-four events in total, most events happened during the warm season from Jun to Aug, but there were also some events that happened in spring or fall. The duration of the events was commonly less than one week. The typical time delay between precipitation peak and flood peak, or basin response time, was 6 hours and 3 hours for Cobb Creek watershed and Chuzhou watershed, respectively. One interesting observation between the two watersheds was the precipitation in Cobb Creek was more peaked with less total accumulation compared to that in Chuzhou watershed. Convective rainfall is more common in Oklahoma, whereas the rainfall in the Chuzhou basin is of moderate intensity but longer duration. Hence the data set had large variability, which was particularly well suited for evaluating whether the data assimilation approach would actually improve the forecast of flash floods.

2.5. Evaluation criteria

Model assessment consisted of the standard statistical evaluations including percent bias (Bias), mean absolute error in percentage (MAE), root mean square error in percentage (RMSE), and Nash–Sutcliffe coefficient of efficiency (NSCE), which is typically used in evaluating model performance, as:

$$NSCE = 1 - \frac{\sum_{i=i_1}^{i_2} (Q_i - \hat{Q}_{ij})^2}{\sum_{i=i_1}^{i_2} (\bar{Q} - Q_i)^2} \quad (8)$$

where Q is the observed streamflow, the hat symbol denotes the estimation of streamflow, i_1 and i_2 denote the start and end time of the simulation/forecast, subscript j denotes the forecast lead time (1 as in simulation mode), and the overbar is the mean value during the whole simulation/forecast period. In order to evaluate the improvement of the data assimilation approach more intuitively, we used an effectiveness criterion (Eff) and a normalized error reduction index (NER) in this study, as:

$$Eff(\%) = 100 \cdot \left(1 - \frac{\sum_{i=i_1}^{i_2} (Q_{ij}^u - Q_i)^2}{\sum_{i=i_1}^{i_2} (Q_{ij}^b - Q_i)^2} \right) \quad (9)$$

$$NER_E(\%) = 100 \cdot \left(1 - \frac{E_u}{E_b} \right) \quad (10)$$

where u and b represent the update after the data assimilation and benchmark run, respectively. E in Eq. (10) represents the error statistical index (MAE and RMSE in this study). Both Eff and NER range between negative infinity and 100%. For Eff , a value larger than 0 corresponds to a positive impact in skill resulting from the data assimilation. Conversely, a value lower than 0 corresponds to a negative impact. The greater value of Eff indicates better performance following data assimilation compared to the benchmark run, and the ideal value is 100%. For NER , negative values mean the assimilation results in a deterioration compared to the benchmark run. Similarly, values of NER closest to 100% indicate a greater improvement of assimilation relative to the benchmark run.

2.6. Model calibration

HyMOD was first calibrated and validated to provide a benchmark performance for assessing the proposed data assimilation results. The Differential Evolution Adaptive Metropolis (DREAM) method [37] was used for parameter optimization during the calibration period from Jun 1, 2005 to May 31, 2006 for Cobb Creek watershed and for 4 flood events on the Chuzhou watershed. The validation period started from Jun 1, 2006 and ends on Sep 30, 2009 for Cobb Creek watershed and for 12 flood events for

Table 2
Information of the twelve largest flood events on record at the Cobb Creek watershed.

	Aug 2007	Jun 2005	Jun 2007	May 2008 a	Aug 2005	Mar 2008
Observed flood peak (m ³ /s)	210	139	138	72	50	38
Peak time	19 Aug, 11 h	13 Jun, 7 h	14 Jun, 14 h	27 May, 13 h	21 Aug, 11 h	18 Mar, 10 h
Precipitation peak time	19 Aug, 5 h	13 Jun, 1 h	14 Jun, 5 h	27 May, 6 h	21 Aug, 3 h	17 Mar 12 h
Precipitation peak (mm/hr)	57	41	54	27	14	8
Precipitation amount (mm)	175	100	134	60	115	69
	Apr 2008	May 2008 b	Sep 2005	Jun 2008	Mar 2007	Aug 2009
Observed flood peak (m ³ /s)	36	32	30	29	27	26
Peak time	10 Apr, 11 h	8 May, 2 h	15 Sep, 9 h	9 Jun, 20 h	31 Mar, 4 h	18 Aug, 15 h
Precipitation peak time	9 Apr, 17 h	7 May, 19 h	15 Sep, 2 h	9 Jun, 8 h	29 Mar, 11 h	18 May, 10 h
Precipitation peak (mm/hr)	12	15	31	21	13	25
Precipitation amount (mm)	65	59	66	65	65	110

Table 3
Information of the twelve largest flood events on record at the Chuzhou watershed.

	Jun 1981	Sep 1981	Aug 1984	Sep 1991	Jun 1994	Aug 1996
Observed flood peak (m ³ /s)	292	468	377	523	317	236
Peak time	11 Jun, 3 h	22 Sep, 21 h	31 Aug, 2 h	8 Sep, 6 h	15 Jun, 13 h	2 Aug, 21 h
Precipitation peak time	11 Jun, 0 h	Sep 22, 18 h	30 Aug, 23 h	7 Sep, 17 h	15 Jun, 9 h	2 Aug, 18 h
Precipitation peak (mm/hr)	29	15	18	13	18	12
Precipitation amount (mm)	123	177	184	275	229	130
	Jun 1998	May 1999	Sep 1999	Jul 2001	Jun 2002	Aug 2002
Observed flood peak (m ³ /s)	222	203	438	714	209	289
Peak time	2 Jun, 8 h	26 May, 5 h	17 Sep, 12 h	7 Jul, 1 h	16 Jun, 12 h	19 Aug, 8 h
Precipitation peak time	2 Jun, 4 h	25 May, 9 h	17 Sep, 9 h	6 Jul, 22 h	16 Jun, 9 h	19 Aug, 5 h
Precipitation peak (mm/hr)	16	15	10	14	12	14
Precipitation amount (mm)	117	157	335	229	204	220

Chuzhou watershed. Although HyMOD is developed for a recommended time step of daily, in our study basin the typical basin concentration time is only several hours. So, in this study the model was operated at an hourly time step. There was no warm-up period prior to the calibration period, and all five model states were initialized to 0 at the beginning. During the validation period, the model parameters were fixed and the initial states were not reinitialized. The parameter ranges and values are provided in Table 4. The Biases during the calibration period and validation period were -0.76% and -17.31% for Cobb Creek, and 16.91% and 15.79% for Chuzhou watershed, respectively. Although the calibration result was not perfect, considering the simple conceptual model with basin-averaged forcing and processes, NSCE values greater than 0.7 indicated reasonable performance suitable for data assimilation experiments.

3. Results

3.1. Modeling and observation uncertainty experiment

Fig. 1 shows the *Eff* values with respect to different background errors in precipitation and streamflow observation errors for a fixed ensemble size of 40. The precipitation error (ω_p) and discharge error (η) both range from 2% to 50%. Recall, the goal of this experiment is to find appropriate values to represent both errors. The result at each bin in the 2D space represents is a mean value computed from 20 runs, in order to minimize random sampling errors stemming from a single run. The contour plot clearly shows that the prescribed errors had a strong impact on the effectiveness of the assimilation procedure. If the background error of precipitation was larger than 35%, the data assimilation led to worse performance than the benchmark run with no assimilation. The effectiveness of data assimilation showed a diminishing trend with increasing streamflow observation error, but this trend was not as

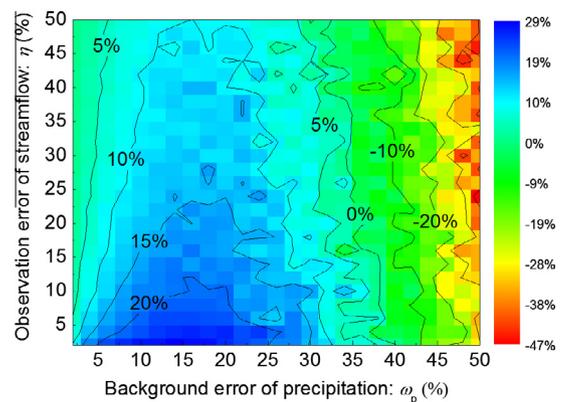


Fig. 1. *Eff* space with respect to forcing and observation error with 40 ensemble members for 20 times run.

obvious when the background error was high (>35%). Typically, standard errors for individual discharge measurements of USGS gauges have been estimated from about 3–6% for direct measurements [38] and 10% for automated measurements [39]. However, during peak flows with flooding events, the discharge errors may be much greater due to uncertainty in the rating curve and additional problems with automated discharge measurements. Thus, we selected twelve flood events from Table 2 to further analyze the impact of discharge observation errors on the effectiveness of the data assimilation. Table 5 shows statistical results by changing the discharge error from 5% to 25% with a fixed ω_p of 20%. This latter value is associated with a relative maximum in the *Eff* values in Fig. 1. Results indicate that the data assimilation didn't perform better by assuming larger errors in the discharge observations for the flooding cases. Hereafter, the observation error η of Cobb Creek watershed was assumed to be 5% according to our results combined with those reported in the literature [38], while that of Chuzhou watershed was assumed to 10% according to the simulation study by Li et al. [29]. The data assimilation result was very sensitive to the assumed error covariance of precipitation, which indicates that it is necessary to identify the appropriate background error. In this study, we set $\omega_p = 20%$, which is reasonable for gauge observations of rainfall (which have their own errors) interpolated to yield a basin-wide mean.

3.2. Ensemble size experiment

After selecting the appropriate modeling and observation errors, we designed an experiment using data from Cobb Creek watershed to evaluate the sensitivity of results to varying the ensemble size. Fig. 2 shows a box plot of the *Eff* statistics for streamflow simulations using ensemble sizes from 10 to 80

Table 4
Calibration and simulation results of the HyMOD.

Parameter	Unit	Range ^a	Cobb Creek	Chuzhou	
C_{max}	(mm)	1–500	49.1790	64.5150	
b_{exp}	(-)	0.1–2	0.5535	0.6090	
α	(-)	0–0.99	0.1599	0.1345	
R_t	(h)	0.1–0.99	0.3925	0.2378	
R_s	(h)	0–0.1	$1.1287e^{-5}$	0.0873	
		Cobb Creek		Chuzhou	
		Bias (%)	NSCE (-)	Bias (%)	NSCE (-)
Calibration		-0.76	0.73	16.91	0.84
Validation		-17.31	0.75	15.79	0.74

^a The range of the parameters is from Blasone et al. [31].

Table 5
Data assimilation performance with respect to discharge error over peak flow period.

Discharge error		Eff (%)					NSCE (-)				
		5%	10%	15%	20%	25%	5%	10%	15%	20%	25%
Flood period	Aug-07	67.66	59.53	54.84	47.05	50.69	0.95	0.95	0.94	0.93	0.94
	Jun-05	33.11	32.99	27.38	32.48	25.46	0.84	0.84	0.81	0.83	0.79
	Jun-07	59.50	45.95	42.37	36.18	29.78	0.85	0.78	0.75	0.73	0.71
	May-08	0.37	0.13	0.65	0.47	0.48	0.00	0.00	0.00	0.00	0.00
	Aug-05	50.55	30.58	19.30	13.74	20.42	0.72	0.67	0.65	0.62	0.65
	Mar-08	11.37	7.11	5.41	11.48	6.81	0.65	0.62	0.60	0.64	0.62
	Apr-08	11.44	4.47	8.82	2.27	6.67	0.70	0.66	0.67	0.63	0.66
	May-08	1.55	0.12	-3.93	-1.65	-1.55	0.66	0.67	0.69	0.68	0.68
	Sep-05	23.93	20.78	16.18	1.43	19.57	0.86	0.86	0.86	0.85	0.86
	Jun-08	25.61	18.13	17.84	12.67	9.80	0.65	0.66	0.67	0.68	0.68
Mar-07	22.7	-0.30	7.02	1.09	11.3	0.81	0.79	0.80	0.80	0.81	
Aug-09	73.99	64.38	45.04	40.63	45.97	0.57	0.53	0.48	0.47	0.48	

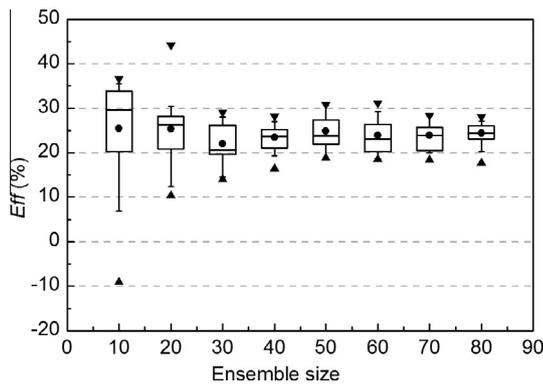


Fig. 2. Box plot of *Eff* with different ensemble members for 20 times run. The black dots show mean value of the *Eff*, the borders of the boxes show 25 and 75 percentiles, the long lines in the boxes show the median of the *Eff*, the whiskers show the 5 and 95 percentiles, and the triangles show the maximum and minimum of the *Eff*.

members. Mean and median values of *Eff* were within 20–30% for all ensemble sizes. However, the spread of the sample was quite large when the ensemble sizes were less than 40 members. The spread of the *Eff* values were approximately equal for ensembles with more than 40 members. Another factor in the consideration of ensemble size is the computational requirements. So, in this study we chose an ensemble size of 50, which is 10 times the

dimension of the model states as recommended by Zhou et al. [40] and Pan et al. [41]. This selection ensured the sampling error was minimized while maintaining acceptable computational speed for practical considerations.

3.3. Model states updating experiment

Table 6 shows the correlation between the observed streamflow and model states for the continuous time series of data during the validation period of Cobb Creek watershed. Strong correlations are shown between streamflow and quick-flow tanks. The x_3 tank conceptually situated closest to the basin outlet had the highest correlation of 0.82, which follows expectations. Observed streamflow had a correlation of 0.32 with the soil moisture tank, which indicates that the water content stored in the soils also had an influence basin outflow. Poor correlations were found between the x_4 slow-flow tank and streamflow and all other model states, which indicates x_4 was not a sensitive state variable for the time period considered.

An analysis was conducted to examine the sensitivity of updating different model states. Five model states were separated into three groups: *S* for the soil moisture tank; x_1, x_2, x_3 combined and denoted hereafter as the quick-flow tank; and x_4 as the slow-flow tank. The entire validation period for the Cobb Creek watershed was used in this experiment, and statistical results are presented in Table 7. Compared with the statistical results in the benchmark run, MAE and RMSE were reduced by updating all states and the quick-flow tank only, and the NSCE improved from

Table 6
Correlation between streamflow observation and model states.

	Observation	Soil tank	Quick-flow tank			Slow-flow tank
	Streamflow	<i>S</i>	x_1	x_2	x_3	x_4
Streamflow	1	0.32	0.27	0.60	0.82	0.027
<i>S</i>	0.32	1	0.19	0.23	0.25	0.01
x_1	0.27	0.19	1	0.73	0.41	-0.00
x_2	0.60	0.23	0.73	1	0.84	-0.00
x_3	0.82	0.25	0.41	0.84	1	-0.00
x_4	0.027	0.01	-0.00	-0.00	-0.00	1

Table 7
Statistical results of updating different model states.

	Benchmark	All states	Soil tank	Quick-flow tank	Slow-flow tank
Bias (%)	-14.3	-21.0	-15.0	-16.6	-23.0
MAE (%)	59.0	55.4	58.9	55.2	59.7
RMSE (%)	201.9	164.1	198.2	163.4	200.7
NSCE (-)	0.75	0.79	0.75	0.79	0.75

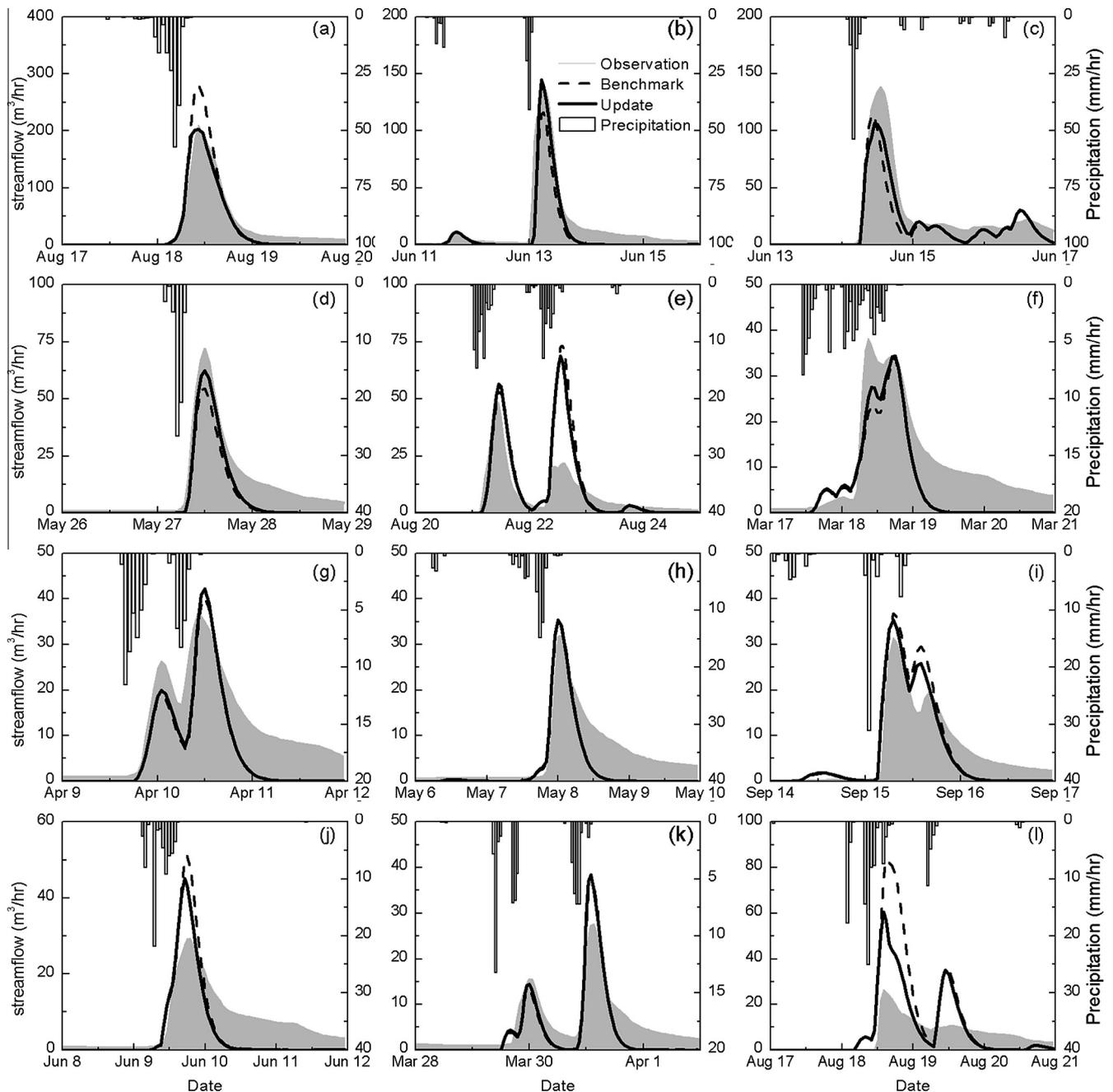


Fig. 3. Hydrographs of one-hour-ahead streamflow forecasting at Cobb Creek watershed. (a) – (l) represents each event listed in Table 2.

0.75 to 0.79. There was little impact by updating the soil tank only, but the Bias statistic deteriorated from 14.3% to -23% by updating the slow-flow tank only. No clear evidence indicated that updating those model states that were only weakly correlated to streamflow would introduce errors, so we updated all five states simultaneously in the experiments hereafter.

3.4. Data assimilation performance experiments in flood forecasting

Figs. 3 and 4 show the one-hour-ahead forecast hydrographs of the twelve flash flood events listed in Table 2 and Table 3 for the two study basins. Generally speaking, data assimilation improved the simulations for most events, especially for the forecast of the peak flow magnitude. HyMOD had better skill on Cobb Creek compared with the Chuzhou watershed. But, after data assimilation,

the Chuzhou watershed had a great improvement in forecasting the peak flow and simulating the entire hydrograph. Simulations on Cobb Creek had improvements in the flood peak after data assimilation, but there were some errors in simulating the entire hydrograph, especially for the recession period. The reason was because the duration of rainfall in Cobb Creek was much shorter than that in Chuzhou watershed. This didn't permit enough time to appropriately perturb the precipitation of the ensemble members, which was a particularity of the specific algorithm used for data assimilation. Although the model states were perturbed by adding random noise, we noted that the model state of soil water content dropped to 0 just a few hours after the flood peak. Thus, the data assimilation approach had a positive effect in the rising limbs in the hydrographs in Cobb Creek, but didn't have much effect during the recession periods. Considering all events, the data assimilation

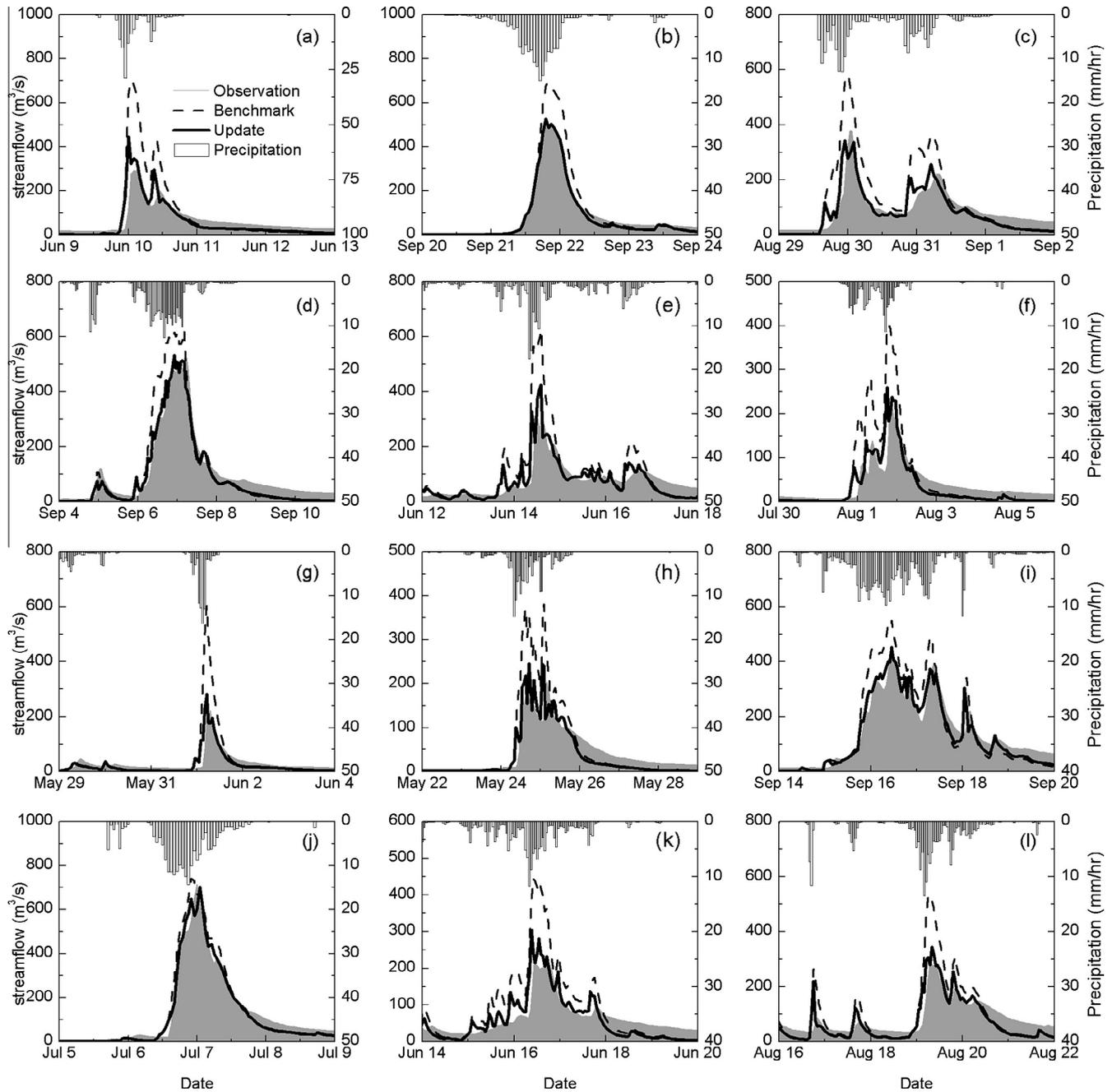


Fig. 4. Hydrographs of one-hour-ahead streamflow forecasting at Chuzhou watershed. (a) – (l) represents each event listed in Table 3.

approach had the most significant improvement with the events caused by rainfall with longer duration. Nonetheless, flood peaks and peak times were improved even for the flash floods, which is of great importance in flood early warning.

We noted that the model performed relatively poorly for double-peaked events. The peakflow and the overall hydrograph were frequently overestimated for the second peak (Figs. 3(e) and (l) and 4(c)). Moreover, the impact of the state updating was minimal. The reason for this was because when the second band of precipitation fell, the water in the routing tanks hadn't been totally released, so the runoff from the soil tank was accumulated in the second flood peak. This problem might be attributed to the simplified structure of the model, but the data assimilation approach could correct the results to some extent. In this example, the advantage of updating during the high flow period was obvious. However, the effective-

ness of data assimilation was still limited because it could not change the model structure nor the model states beyond reasonable ranges, else there would be an imbalance in the total water amount in the soil water or routing tanks.

In real operations, a one-hour-ahead forecast is typically insufficient for early warning, so it is important to analyze the feasibility of data assimilation in the flood forecasting at longer lead times. Here, we tested the ability of the updating procedure to improve the forecast of the flood peaks listed in Tables 2 and 3. A 3-hour-ahead forecast was considered given the small sizes and quick responses of the basins, which meant that the data assimilation approach was implemented 3 hours before the flood peak occurred. The results of the comparison over both watersheds are shown in Figs. 5 and 6. For Cobb Creek, the flood peaks were improved in ten out of twelve of the flood events. One of the weaker events was slightly deteriorated

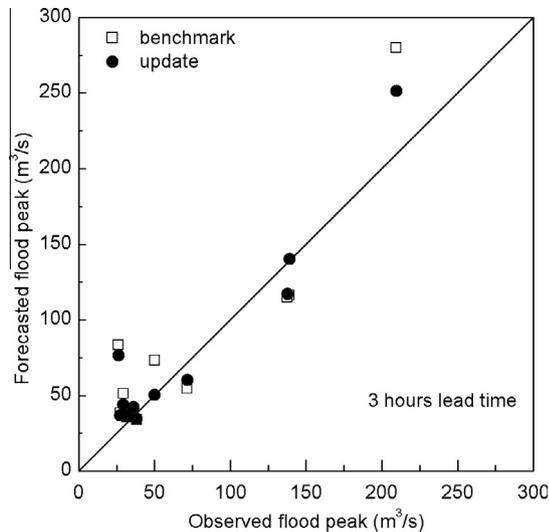


Fig. 5. Comparison of the 3 h-ahead forecast of the flood peaks with and without update for the twelve largest flood events at Cobb Creek watershed on record as of Table 2.

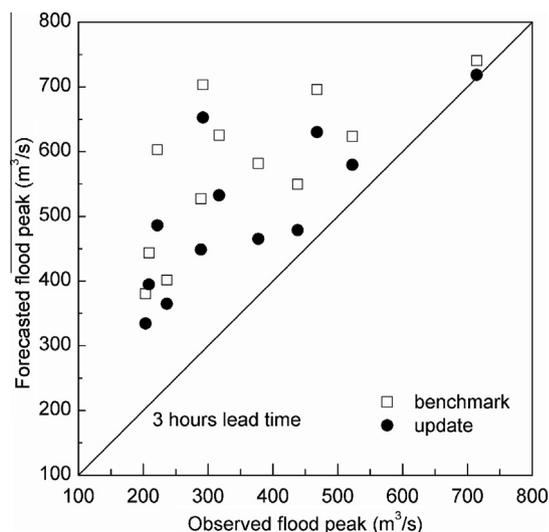


Fig. 6. Comparison of the 3 h-ahead forecast of the flood peaks with and without update for the twelve largest flood events at Chuzhou watershed on record as of Table 3.

for the flood peak, which had an observed value of $36 \text{ m}^3 \text{ s}^{-1}$, $42 \text{ m}^3 \text{ s}^{-1}$ with the state updates, and $40 \text{ m}^3 \text{ s}^{-1}$ without them. For the twelve events on Cobb Creek in Fig. 5, the mean MAE of the peaks was 43% for the benchmark run, which decreased to 32% following the state updates. For the Chuzhou watershed, however, the flood peaks were overestimated by approximately 60% for the benchmark run (Fig. 6), which indicates that the calibrated model had difficulty in capturing the flood peaks in this small, mountainous watershed. However, following data assimilation, all twelve flood events revealed an improvement in flood peak forecasting, with a Bias of 42%. All twelve events were better forecast following data assimilation. It is obvious that the data assimilation approach has significant merits in terms of forecasting flood peaks with a 3-hour lead time.

The forecast accuracy for a series of lead times was evaluated for all flood events on both study basins. Four error measures, normalized error reduction (NER) of MAE and RMSE, NSCE, and *Eff*, were all considered. The results of the evaluation metrics are plotted as a function of forecast lead time in Figs. 7 and 8 for Cobb

Creek and Chuzhou watershed, respectively. Each box incorporates all twelve events in order to show the general performance of the data assimilation approach. The results show similar patterns for all four metrics on Cobb Creek (Fig. 7). In general, the data assimilation approach had a positive impact on the forecast of peak flows for lead times less than 6 hour. The impact of data assimilation decreased with longer lead times. This result was understandable because the typical flow routing time in this watershed was approximately 6 hour. If the lead time exceeds the basin concentration time, then the state updating had already been completed before the rainfall. Adjustments at these long lead times were quite minor and ineffective because the update excluded the perturbations to precipitation, the largest uncertainty in the streamflow forecast. In this case, the data assimilation approach had little effect on the streamflow forecast. For the Chuzhou watershed, the data assimilation approach had a definite positive impact on flood-event forecasting for a broader range of lead times. The different performance between Cobb Creek watershed and Chuzhou watershed is primarily attributed to the differences in the precipitation processes and geomorphologies in the two watersheds. First, the rainfall in Cobb Creek was commonly from short-duration, intense convective thunderstorms, while that in Chuzhou watershed was mostly moderate rainfall with long duration. Thus, the ensembles in Cobb Creek offered shorter durations and thus less opportunity to get fully perturbed. Another reason was that the mechanism of runoff generation was different in the two watersheds. Intense storms in Cobb Creek caused infiltration excess runoff while that in Chuzhou watershed was more saturation excess. This latter mechanism is better conceptualized in the HyMOD structure, more so than the former, which provides some explanation as to the effectiveness of data assimilation in the in Chuzhou watershed.

To further illustrate the impact of data assimilation with lead time, we selected the largest flood event on record in Table 2 (Aug 2007). Fig. 9 shows forecast hydrographs at different lead times as well as the observed streamflow and benchmark simulation. For one-hour-ahead forecasting, the forecast after updating performed very well. Both the hydrograph and the peak of the flood were close to the observation. The 3-hour-ahead forecast had less effect compared to the one-hour lead, but still improves upon the result compared with the benchmark run. The forecast with the 6-hour lead time shows little difference compared to the benchmark run, which means that the data assimilation approach has little effect when the lead time is longer than the basin concentration time.

4. Discussion

Real-time flood forecasting remains a challenging problem in hydrology, especially for flash flooding events over small watersheds. In gauged basins, data assimilation provides a potential tool to handle different sources of uncertainty, which are inevitable in forcing data, model structures, parameters and observations. In this study, we applied the EnSRF approach into HyMOD to test the effectiveness of the state-updating procedure in flood forecasting. As a variant of the standard implementation of the EnKF method, the EnSRF method avoids systematic underestimation of the posterior covariance, which leads to the use of perturbed observations in the EnKF method [16], and the feasibility in streamflow simulations has been tested in a distributed hydrologic model [34]. In former sections, the applicability of the EnSRF method was demonstrated through testing the state-updating procedure for flood forecasting.

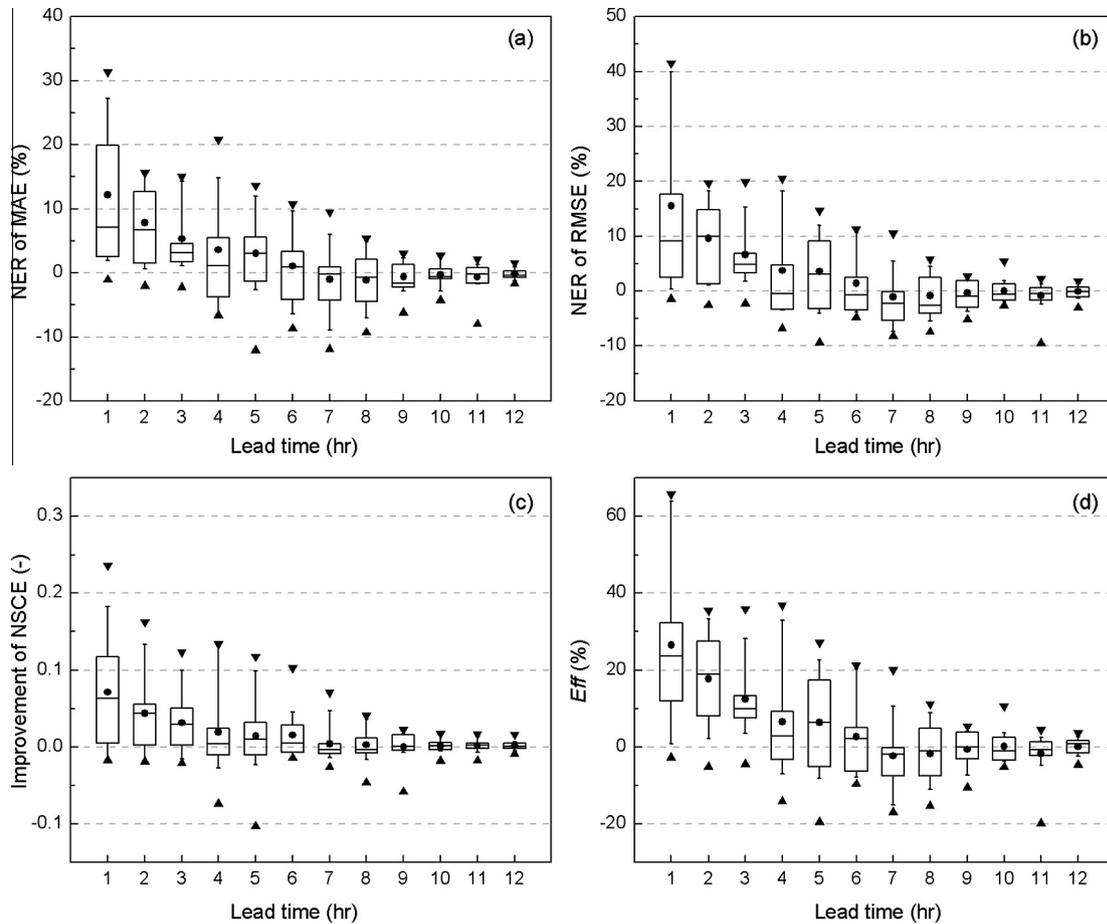


Fig. 7. Box plot of forecast evaluation with different lead time for the twelve largest flood events at Cobb Creek watershed on record of Table 2. (a) Normalized error reduction (NER) of MAE; (b) NER of RMSE; (c) difference of NSCE between update run and benchmark run; (d) Eff of the data assimilation.

The most crucial part in applying the filter on actual events is to get a good estimate of the model errors, since less accurate estimates of the model errors will lead to a suboptimal estimation of the model states. The method for estimating model error in this study was similar to those used in other studies (e.g. [19],[34]), which separated the model error into different error terms of forcing data, model states, and the observations of system behavior. To further investigate the assimilation performance with respect to model error, the effectiveness of state updating was examined for a range of forcing and observation perturbations. Results showed that the effectiveness of data assimilation was more sensitive to the rainfall-forcing perturbation than the observation of streamflow, implying that accurate streamflow simulation using HyMOD was highly dependent on the accuracy and the uncertainty estimation of the rainfall forcing. Moradkhani et al. [22] performed a similar analysis with a different evaluation criterion of Normalized RMSE Ratio (NRR), but their results showed that NRR was more sensitive to the observation perturbations. This conflicting result indicated that a rigorous systematic sensitivity analysis to quantify the forcing and observation perturbation factors was needed, since their impact on data assimilation performance might depend on the hydrologic model used and the characteristics of the studied basin.

The EnSRF method outperformed the benchmark run for forecasting the general hydrographs of the floods, but its effectiveness remained limited. One reason was the simplified treatment of the water distribution in the soil tank represented within HyMOD. After the flood peak occurred, the water storage in the soil tank quickly and unrealistically reduced to zero due to evapotranspiration, and if no rainfall occurred during the low flow period, the

water storage remained at zero. Thus, the ensemble generated by the data assimilation method had little spread because no perturbations were applied to the precipitation or soil moisture. Another reason was the time lag between model states and streamflow. In this study, the state update was based on an instantaneous observation of streamflow, but the simulated streamflow was affected by the soil water storage a couple of hours prior. So the model state of the soil tank was insensitive to the updating procedure, especially during low flow periods. The assimilation of soil moisture data is necessary to obtain a better updated model state of the hydrologic model [42]. As the objective of the paper is dedicated to flood forecasting, the soil moisture data was not coupled to the assimilation of streamflow, and the assimilation technique was proven to be well adapted in the context of flood forecasting in gauged basins.

Unlike flood events in large-scale river basins, a distinguishing characteristic of floods in small-scale watersheds is their quick but potentially intense and catastrophic responses to rainfall forcing. As stated above, the modeled soil water content in HyMOD before the rising limb was found to be approximately zero, which means the streamflow simulation was insensitive to the initial soil moisture states. Therefore, the soil moisture state initialization improved by the data assimilation system has no appreciable effect on flood forecasting. However, this finding is very likely specific to the simplified structure of HyMOD and perhaps to the small basin sizes studied here. Effectiveness of assimilating soil moisture data on flood responses should be evaluated more comprehensively using more complex model structures across basins with different scales and geomorphological characteristics.

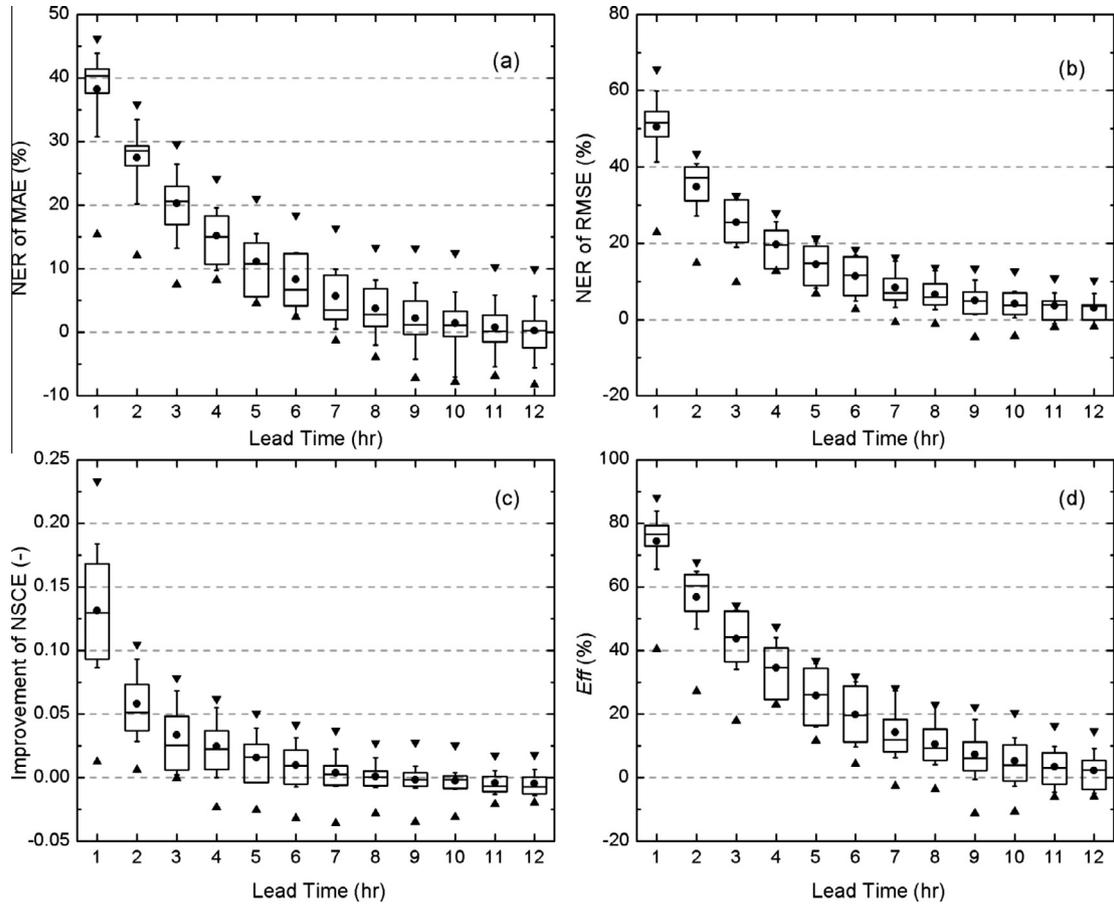


Fig. 8. Box plot of forecast evaluation with different lead time for the twelve largest flood events at Chuzhou watershed on record of Table 3. (a) Normalized error reduction (NER) of MAE; (b) NER of RMSE; (c) difference of NSCE between update run and benchmark run; (d) *Eff* of the data assimilation.

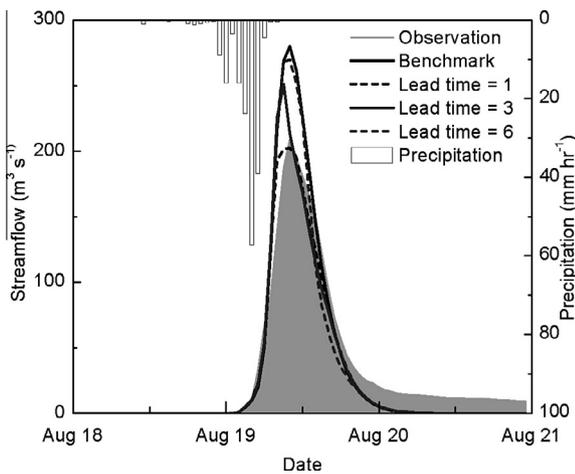


Fig. 9. Precipitation (reversed bars) and streamflow simulations of benchmark runs (black solid line) and forecast of update runs with the lead time of 1-h (black dash line), 3-h (gray solid line), and 6-h (gray dash line) compared with observations (gray shadows) of the largest event on record of Table 2 (Aug 2007).

5. Conclusions

The objectives of the study were to evaluate the feasibility of a data assimilation system for real-time flash flood forecasting over small watersheds, and to examine the applicability of the EnSRF method for updating model states. The following conclusions were reached in this study:

- (1) The sensitivity analysis indicated that the effectiveness of data assimilation is more sensitive to the prescribed background error of the precipitation forcing than the observational error. If the precipitation error was assumed to be greater than 35%, then the data assimilation method was ineffective. More importantly, both errors should be set to a reasonable range, which was found to be 5% for streamflow observations and 20% for basin-averaged, gauge-based rainfall forcing. These results are based on the sensitivity analysis and instrument and sampling considerations. If these errors are not properly accounted for, then the hydrologic simulation may be deteriorated following data assimilation.
- (2) The ensemble size experiment results showed the data assimilation performance was suboptimal until the ensemble size reached 50 members (10 times of the dimension of the model states). Once the ensemble size is sufficient to average out the sampling errors, increasing ensemble members didn't have much of a difference in the data assimilation results, and only increased the computational expense.
- (3) The model states-updating experiment results showed strong correlations between streamflow and quick-flow tanks, but poor correlations between streamflow and the slow-flow tank. However, updating the poorly correlated slow-flow tank didn't introduce extra errors to the streamflow simulation.
- (4) The assimilation of streamflow for flash flood forecasting reduced forecast errors in both flood peak and the hydrograph at both study watersheds. The degree of improvement was limited for Cobb Creek due to the prevalence of short-duration, high-intensity rainfall from convective storms. Also, the improvements decreased with increasing forecast lead time. When the lead time

exceeded the flow routing time of the basin, the data assimilation approach didn't provide any improvements to the streamflow forecast and represents a physical limit.

This study demonstrated that the proposed EnSRF data assimilation technique based on the ensemble Kalman filter concept was suitable to the context of streamflow forecasting even for individual, flash flood events at small watersheds. Although the benefits and efficiency of implementing the EnSRF approach into a conceptual hydrologic model were demonstrated in this study on two small basins, there is a great potential for its use in a real-time flash flood warning system. Future work will evaluate the effectiveness of data assimilation in flood forecasting when considering model resolution, data availability, remotely-sensed input data, and spatially distributed hydrologic models with various complexity.

Acknowledgment

This research was supported by the National Natural Science Funds for Distinguished Yong Scholar (No. 51025931), the National Natural Science Foundation of China (No. 50939004), and the Ministry of Water Resources Funds for Public Welfare Industry (No. 201101004). We thank Prof. Murugesu Sivapalan of University of Illinois at Urbana-Champaign for helpful comments. The authors would like to express their appreciation to three anonymous reviewers, whose comments and suggestions led to significant improvement in the manuscript.

References

- Sivapalan M, Takeuchi K, Franks SW, Gupta VK, Karambiri H, Lakshmi V, et al. IAHS decade on predictions in ungauged basins (PUB), 2003–2012: shaping an exciting future for the hydrological sciences. *Hydrol Sci J* 2003;48:857–80. <http://dx.doi.org/10.1623/hysj.48.6.857.51421>.
- Thirel G, Martin E, Mahfouf JF, Massart S, Ricci S, Habets F. A past discharges assimilation system for ensemble streamflow forecasts over France. Part 1: Description and validation for the assimilation system. *Hydrol Earth Syst Sci* 2010;14:1623–37. <http://dx.doi.org/10.5194/hess-14-1623-2010>.
- Göttinger J, Bårdossy A. Generic error model for calibration and uncertainty estimation of hydrological models. *Water Resour Res* 2008;44:W00B07. <http://dx.doi.org/10.1029/2007WR006691>.
- Moradkhani M, Sorooshian S. General review of rainfall–runoff modeling: model calibration, data assimilation, and uncertainty analysis. In: Sorooshian S, Hsu KL, Coppola E, Tomassetti B, Verdecchia M, Visconti G, editors. *Hydrological modeling and the water cycle*. Springer; 2009. p. 1–24. http://dx.doi.org/10.1007/978-3-540-77843-1_1.
- Salamon P, Feyen L. Assessing parameter, precipitation, and predictive uncertainty in a distributed hydrological model using sequential data assimilation with the particle filter. *J Hydrol* 2009;376:428–42. <http://dx.doi.org/10.1016/j.jhydrol.2009.07.051>.
- Kalman RE. A new approach to linear filtering and prediction problems. *Trans ASME J Basic Eng* 1960;82:35–45. <http://dx.doi.org/10.1115/1.3662552>.
- Kitinadis PK, Bras RL. Real time forecasting with a conceptual hydrologic model. 1. Analysis of uncertainty. *Water Resour Res* 1980;16:1025–33. <http://dx.doi.org/10.1029/WR016i006p01025>.
- Georgakakos KP. A generalized stochastic hydrometeorological model for flood and flash-flood forecasting. 1. Formulation. *Water Resour Res* 1986;22:2083–95. <http://dx.doi.org/10.1029/WR022i013p02083>.
- Georgakakos KP. A generalized stochastic hydrometeorological model for flood and flash-flood forecasting. 2. Case studies. *Water Resour Res* 1986;22:2096–106. <http://dx.doi.org/10.1029/WR022i013p02096>.
- Gelb A. *Applied optimal estimation*. The MIT Press; 1974. ISBN:0-262-20027-9.
- Evensen G. Using the extended Kalman filter with a nonlinear quasigeostrophic ocean model. *J Geophys Res* 1992;97:17905–24. <http://dx.doi.org/10.1029/92JC01972>.
- Evensen G. Sequential data assimilation with a nonlinear quasigeostrophic model using Monte Carlo methods to forecast error statistics. *J Geophys Res* 1994;99:10143–62. <http://dx.doi.org/10.1029/94JC00572>.
- Reichle RH, Walker JP, Koster RD, Houser PR. Extended versus ensemble Kalman filtering for land data assimilation. *J Hydrometeorol* 2002;3:728–40. <http://dx.doi.org/10.1175/1525-7541>.
- Lorenc AC. The potential of the ensemble Kalman filter for NWP – a comparison with 4D-Var. *Q J R Meteor Soc* 2003;129(595):3183–203. <http://dx.doi.org/10.1256/qj.02.132>.
- Weerts AH, El Serafy GYH. Particle filtering and ensemble Kalman filtering for state updating with hydrological conceptual rainfall–runoff models. *Water Resour Res* 2006;4:W09403. <http://dx.doi.org/10.1029/2005WR004093>.
- Whitaker JS, Hamill TM. Ensemble data assimilation without perturbed observations. *Mon Weather Rev* 2002;130(7):1913–24. <http://dx.doi.org/10.1175/1520-0493>.
- Vrugt JA, Diks CGH, Gupta HV, Bouten W, Verstraten JM. Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation. *Water Resour Res* 2005;41:W01017. <http://dx.doi.org/10.1029/2004WR003059>.
- Aubert D, Loumagne C, Oudin L. Sequential assimilation of soil moisture and streamflow data in a conceptual rainfall–runoff model. *J Hydrol* 2003;280:141–61. <http://dx.doi.org/10.1016/S0022-1694>.
- Komma J, Blöschl G, Reszler C. Soil moisture updating by ensemble Kalman filtering in real-time flood forecasting. *J Hydrol* 2008;357:228–42. <http://dx.doi.org/10.1016/j.jhydrol.2008.05.020>.
- Thirel G, Martin E, Mahfouf JF, Massart S, Ricci S, Habets F. A past discharges assimilation system for ensemble streamflow forecasts over France. Part 2: Impact on the ensemble streamflow forecast. *Hydrol Earth Syst Sci* 2010;14:1639–53. <http://dx.doi.org/10.5194/hess-14-1639-2010>.
- Boyle DP. Multicriteria calibration of hydrological models. Ph.D. dissertation, Dep of Hydrol and Water Resour, Univ of Arizona, Tucson; 2000. UMI number: 3023521.
- Moradkhani M, Sorooshian S, Gupta HV, Houser PR. Dual state–parameter estimation of hydrological models using ensemble Kalman filter. *Adv Water Resour* 2005;28:135–47. <http://dx.doi.org/10.1016/j.advwatres.2004.09.002>.
- van Delft G, El Serafy GY, Heemink AW. The ensemble particle filter (EnPF) in rainfall–runoff models. *Stoch Environ Res Risk A* 2009;23:1203–11. <http://dx.doi.org/10.1007/s00477-008-0301-z>.
- Gourley JJ, Hong Y, Flamig ZL, Wang J, Vergara A, Anagnostou EN. Hydrologic evaluation of rainfall estimates from radar, satellite, gauge, and combinations on Ft. Cobb basin, Oklahoma. *J Hydrometeorol* 2011;12:973–88. <http://dx.doi.org/10.1175/2011JHM1287.1>.
- Li M, Yang D, Chen J, Hubbard SH. Calibration of a distributed flood forecasting model with input uncertainty using a Bayesian framework. *Water Resour Res* 2012;48:W08510. <http://dx.doi.org/10.1029/2010WR010062>.
- Tripp DR, Niemann JD. Evaluating the parameter identifiability and structural validity of a probability-distributed model for soil moisture. *J Hydrol* 2008;353:93–108. <http://dx.doi.org/10.1016/j.jhydrol.2008.01.028>.
- Homer C, Dewitz J, Fry J, Coan M, Hossain N, Larson C, et al. Completion of the 2001 national land cover database for the conterminous United States. *Photogramm Eng Remote Sens* 2007;73:337–41.
- Soil survey staff. State soil geographic database (STATSGO) data users guide, vol. 1492. USDA Natural Resources Conservation Service Miscellaneous Publication; 1994 [p. 88–1036].
- Li M. Uncertainty estimation of hydrological models under a Bayesian framework. Ph.D dissertation. Tsinghua University. 2012; p. 141.
- Wagener T, Boyle DP, Lees MJ, Wheater HS, Gupta HV, Sorooshian S. A framework for development and application of hydrological models. *Hydrol Earth Syst Sci* 2001;5(1):13–6. <http://dx.doi.org/10.5194/hess-5-13-2001>.
- Blasone RS, Vrugt JA, Madsen H, Rosbjerg D, Robinson BA, Zylowski GA. Generalized likelihood uncertainty estimation (GLUE) using adaptive Markov Chain Monte Carlo sampling. *Adv Water Resour* 2008;31:630–48. <http://dx.doi.org/10.1016/j.advwatres.2007.12.003>.
- Kalman R, Bucy R. New results in linear prediction and filtering theory. *Trans AMSE J Basic Eng* 1961;83D:95–108. <http://dx.doi.org/10.1115/1.3658902>.
- Burgers G, van Leeuwen PJ, Evensen G. Analysis scheme in the ensemble Kalman filter. *Mon Weather Rev* 1998;126:1719–24. [http://dx.doi.org/10.1175/1520-0493\(1998\)126<1719:ASITEK>2.0.CO;2](http://dx.doi.org/10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2).
- Clark MP, Rupp DE, Woods RA, Zheng X, Ibbitt RP, Slater AG, et al. Hydrological data assimilation with the ensemble Kalman filter: use of streamflow observations to update states in a distributed hydrological model. *Adv Water Res* 2008;31:1309–24. <http://dx.doi.org/10.1016/j.advwatres.2008.06.005>.
- Vrugt JA, Gupta HV, Nuallain BO, Bouten W. Real-time data assimilation for operational ensemble streamflow forecasting. *J Hydrometeorol* 2006;7:548–65. <http://dx.doi.org/10.1175/JHM504.1>.
- Crow WT, Loon EV. Impact of incorrect model error assumptions on the sequential assimilation of remotely sensed surface soil moisture. *J Hydrometeorol* 2006;7(3):421–32. <http://dx.doi.org/10.1175/JHM499.1>.
- Vrugt JA, ter Braak CJF, Diks CGH, Higdon D, Robinson BA, Hyman JM. Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling. *Int J Nonlinear Sci Numer Simul* 2009;10(3):273–90. <http://dx.doi.org/10.1515/IJNSNS.2009.10.3.273>.
- Sauer VB, Meyer RW. Determination of error in individual discharge measurements. US geological survey, Norcross, Georgia; 1992.
- Harmel RD, Cooper RJ, Slade RM, Haney RL, Arnold JG. Cumulative uncertainty in measured streamflow and water quality data for small watersheds. *Am Soc Agric Biol Eng* 2006;49(3):689–701.
- Zhou Y, McLaughlin D, Entekhabi D, Chatdarong V. Assessing the performance of the ensemble Kalman filter for land. *Mon Weather Rev* 2006;134(8):2128–42. <http://dx.doi.org/10.1175/MWR3153.1>.
- Pan M, Wood EF, Wójcik R, McCabe MF. Estimation of regional terrestrial water cycle using multi-sensor remote sensing observations and data assimilation. *Remote Sens Environ* 2008;112:1282–94. <http://dx.doi.org/10.1016/j.rse.2007.02.039>.
- Chen F, Crow WT, Starks PJ, Moriasi DN. Improving hydrologic predictions of a catchment model via assimilation of surface soil moisture. *Adv Water Res* 2011;34:526–36. <http://dx.doi.org/10.1016/j.advwatres.2011.01.011>.