



Variational assimilation of streamflow into operational distributed hydrologic models: effect of spatiotemporal scale of adjustment

H. Lee^{1,2}, D.-J. Seo^{1,2,*}, Y. Liu^{1,3,**}, V. Koren¹, P. McKee⁴, and R. Corby⁴

¹NOAA, National Weather Service, Office of Hydrologic Development, Silver Spring, Maryland, USA

²University Corporation for Atmospheric Research, Boulder, Colorado, USA

³Riverside Technology, Inc., Fort Collins, Colorado, USA

⁴NOAA, National Weather Service, West Gulf River Forecast Center, Fort Worth, Texas, USA

* present address: Department of Civil Engineering, The University of Texas at Arlington, Arlington, TX 76019-0308, USA

** present address: Goddard Space Flight Center, National Aeronautics and Space Administration, Greenbelt, MD 20771, USA

Correspondence to: H. Lee (haksu.lee@noaa.gov)

Received: 13 December 2011 – Published in Hydrol. Earth Syst. Sci. Discuss.: 5 January 2012

Revised: 30 May 2012 – Accepted: 14 June 2012 – Published: 23 July 2012

Abstract. State updating of distributed rainfall-runoff models via streamflow assimilation is subject to overfitting because large dimensionality of the state space of the model may render the assimilation problem seriously underdetermined. To examine the issue in the context of operational hydrologic forecasting, we carried out a set of real-world experiments in which streamflow data is assimilated into the gridded Sacramento Soil Moisture Accounting (SAC-SMA) and kinematic-wave routing models of the US National Weather Service (NWS) Research Distributed Hydrologic Model (RDHM) via variational data assimilation (DA). The nine study basins include four in Oklahoma and five in Texas. To assess the sensitivity of the performance of DA to the dimensionality of the control vector, we used nine different spatiotemporal adjustment scales, with which the state variables are adjusted in a lumped, semi-distributed, or distributed fashion and biases in precipitation and PE are adjusted at hourly or 6-hourly scale, or at the scale of the fast response of the basin. For each adjustment scale, three different assimilation scenarios were carried out in which streamflow observations are assumed to be available at basin interior points only, at the basin outlet only, or at all locations. The results for the nine basins show that the optimum spatiotemporal adjustment scale varies from basin to basin and between streamflow analysis and prediction for all three streamflow assimilation scenarios. The most preferred adjustment scale for seven out of the nine basins is found to be distributed and hourly. It was found that basins with highly

correlated flows between interior and outlet locations tend to be less sensitive to the adjustment scale and could benefit more from streamflow assimilation. In comparison with outlet flow assimilation, interior flow assimilation produced streamflow predictions whose spatial correlation structure is more consistent with that of observed flow for all adjustment scales. We also describe diagnosing the complexity of the assimilation problem using spatial correlation of streamflow and discuss the effect of timing errors in hydrograph simulation on the performance of the DA procedure.

1 Introduction

Improving flood forecasting has long been an important research topic for natural hazard mitigation (Droegemeier et al., 2000; NHWC, 2002; NRC, 2010; USACE, 2000). Changes in spatiotemporal patterns of precipitation and occurrences of record-breaking events at unprecedented scales around the globe during the past decades (Knutson et al., 2010; Milly et al., 2008; Min et al., 2011; Trapp et al., 2007; Trenberth et al., 2003) are pressing further the needs for rapid advances in real-time flood forecasting (NRC, 2010). In the US River Forecast Centres (RFCs), the flood forecasting process has often involved manual modifications of the model states by human forecasters (MOD; Seo et al., 2003; Smith et al., 2003) to reconcile any significant differences between the model results and the observations. With distributed models,

such MOD's are a very difficult proposition due to the generally very large dimensionality of the state variables involved. As such, automatic data assimilation (DA) procedures are a necessity.

DA techniques merge information in the real-time hydrologic and hydrometeorological observations into the hydrologic model dynamics by considering uncertainties from different error sources (Liu and Gupta, 2007; McLaughlin, 2002; Moradkhani, 2008; Seo et al., 2003, 2009; Troch et al., 2003). Compared to applying DA to lumped rainfall-runoff models (e.g., Bulygina and Gupta, 2009; Moradkhani et al., 2005a,b; Seo et al., 2003, 2009; Vrugt et al., 2005, 2006; Weerts and El Serafy, 2006), state updating of distributed rainfall-runoff models is subject to overfitting to a much greater extent due to the typically much larger dimensionality of the state space of the model.

In operational streamflow forecasting, other than the atmospheric forcing data, streamflow observation is usually the only source of data available for assimilation, which is often insufficient to reduce the large degrees of freedom (DOF) associated with distributed models. As such, most, if not all, distributed rainfall-runoff models are under-determined, i.e., the information available in the data is not enough to uniquely determine the state variables and/or parameters of the model. Note that, while a vast amount of remote-sensing data, in particular satellite-based, are widely available, they are generally of limited utility for operational river and flash flood forecasting due to large temporal sampling intervals and relatively low information content at the catchment scale. In an under-determined system, streamflow analysis at independent validation locations as well as streamflow prediction at any locations in a basin could be worse than the base model streamflow simulation due to overfitting. This poses an obvious obstacle to advances in DA for distributed models which requires developing appropriate assimilation strategy to constrain large degrees of freedom causing a state and/or parameter identifiability problem.

In the following, we summarize previous studies that address state and/or parameter identifiability when applying DA techniques to distributed hydrologic modelling. Clark et al. (2008) tested the impact of assimilating streamflow at one location on streamflow prediction at other locations in the Wairau River basin in New Zealand by using the ensemble square root Kalman filter (EnSRF) and the distributed model TopNet. They obtained degraded streamflow results from assimilation at independent validation locations in the basin, highlighting the importance of accurately modelling spatial variability, or correlation structure of hydrological processes in order to improve streamflow prediction at ungauged locations by assimilating streamflow observations from elsewhere in the basin. With the limited data available in field operations, the assimilation technique may necessarily adjust state variables at some lumped fashion (e.g., at the sub-basin scale) that reflects or preserves the spatial correlation length or structure of hydrological processes. Lee

et al. (2011) found in a synthetic experiment using the El-don basin (ELDO2) in Oklahoma that assimilating outlet flow into the gridded SAC via variational assimilation degraded streamflow prediction at interior locations. This indicates difficulty of solving the inverse problem in distributed modelling based on the limited information available in the outlet flow data alone, and points out the need for different assimilation strategies to reduce the degrees of freedom associated with the problem. Van Loon and Troch (2002) noted degraded prediction of ground water depth at some locations in a 44-ha catchment in Costa Rica from assimilating soil moisture measured by a Trime time domain reflectometry (TDR) system at multiple sites. The above result was obtained even though discharge predictions were benefited considerably from soil moisture assimilation. Chen et al. (2011) carried out assimilation of 20 different sets of synthetically generated soil moisture observations into the SWAT model with the ensemble Kalman filter (EnKF). They found that, for some cases, analyses of groundwater flow and percolation rate were degraded. Brocca et al. (2010) assimilated the rescaled Soil Wetness Index (SWI) into the semi-distributed model Modello Idrologico SemiDistribuito in continuo (MISDc) in a synthetic experiment. They found that the assimilation results for flood prediction were degraded for some experimental settings. Performance degradation following assimilation from above studies may be due to a combination of factors, including inadequate model physics as mentioned in Van Loon and Troch (2002) and Chen et al. (2011) or inappropriate assimilation strategy. Although it is difficult to trace the causes in real-world applications due to the presence of a number of different error sources and their unknown characteristics, the degraded assimilation results in the above studies warrant exploring different assimilation strategies.

To address the aforementioned issues with DA into distributed models in an operational setting and to develop an effective assimilation strategy in order to limit the degrees of freedom in DA with distributed models, in this study we investigate the effect of the spatiotemporal scale of adjustment on analysis and prediction of streamflow. The analysis and prediction are generated by assimilating streamflow data into the distributed SAC-SMA and kinematic-wave routing models of HL-RDHM with the variational DA procedure (Seo et al., 2010; Lee et al., 2011). We tested nine spatiotemporal adjustment scales based on combinations of three spatial scales of adjustment (lumped, semi-distributed, distributed) to state variables and three temporal scales (hourly, 6-hourly, fast-response time of the basin) of adjustment to mean field bias in the precipitation and PE data. The strategy of adjusting short-term biases in the forcing data in the assimilation procedure is motivated by the use of long-term bias adjustment factors in the SAC-SMA model calibration. Adopting a coarser spatiotemporal adjustment scale would reduce the dimensionality of the control vector, which may help prevent overfitting when solving the inverse problem. For basins

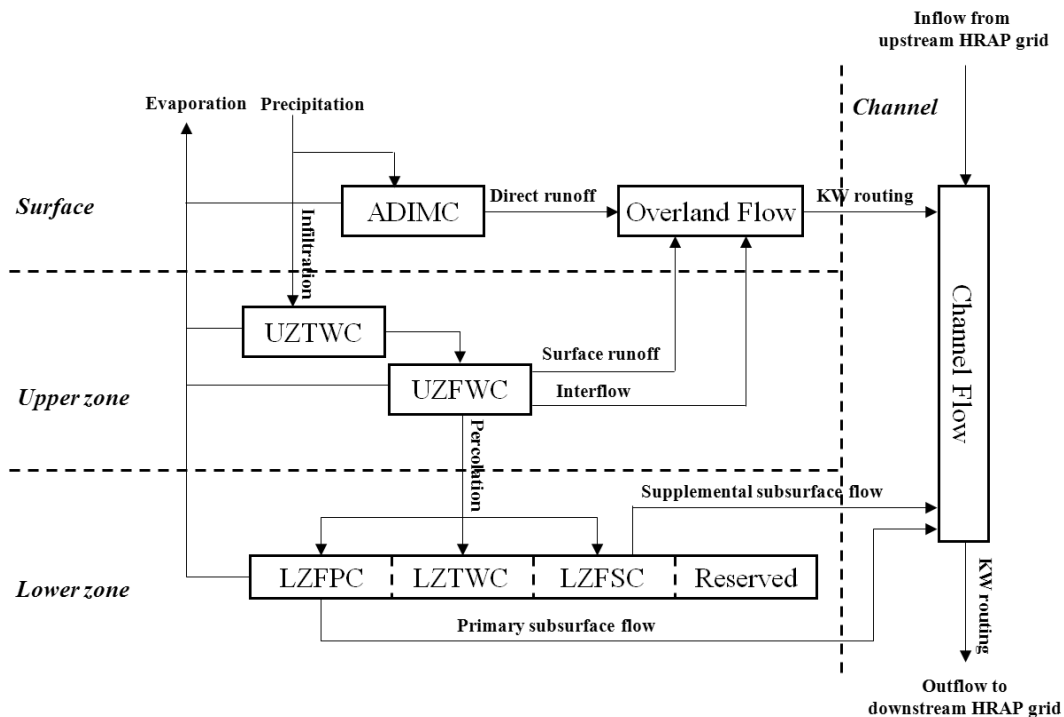


Fig. 1. Schematic of the gridded SAC and kinematic-wave routing models of HL-RDHM (from Lee et al., 2011).

with highly heterogeneous soil and physiographic properties and precipitation patterns, however, a finer adjustment scale may be preferable. In this work, three streamflow assimilation scenarios are considered, i.e., assimilating interior flow observations only, outlet flow observations only, or all flow observations. Given a spatiotemporal adjustment scale, each streamflow assimilation scenario is carried out for four basins in Oklahoma and five basins in Texas, US.

The paper is organised as follows: Sect. 2 describes the methodology including the hydrologic model, the assimilation technique, and the evaluation metrics; Sect. 3 describes the study basins; Sect. 4 describes the multi-basin experiment and presents the results and discussions; and finally, Sect. 5 summarises conclusions and future research recommendations.

2 Methodology

2.1 The gridded SAC and kinematic-wave routing models of HL-RDHM

The models used are the gridded Sacramento Soil Moisture Accounting (SAC-SMA) and kinematic-wave routing models of the National Weather Service (NWS) Hydrology Laboratory’s Research Distributed Hydrologic Model (HL-RDHM, Koren et al., 2004). The SAC-SMA is a conceptual rainfall runoff model (Burnash et al. 1973) which calculates fast and slow runoffs from two subsurface zones, i.e., Upper

Zone (UZ) and Lower Zone (LZ). The UZ is thinner than the LZ and consists of tension and free water storages. The LZ is composed of tension and primary and supplemental free water storages. Soil moisture states at each subsurface storage compartment are named as Upper Zone Tension Water Content (UZTWC), Upper Zone Free Water Content (UZFWC), Lower Zone Tension Water Content (LZTWC), Lower Zone Primary Free Water Content (LZFWC), and Lower Zone Supplemental Free Water Content (LZFSC) (Koren et al., 2004). The sum of the surface and subsurface runoff is then routed through the kinematic-wave routing model to calculate flow at each HRAP grid. The models operate at an hourly time step on the Hydrologic Rainfall Analysis Project (HRAP) grid (~ 16 km²) (Greene and Hudlow, 1982; Reed and Maidment, 1999). The NEXRAD-based multi-sensor precipitation data (Fulton et al., 1998; Seo, 1998; Seo et al., 1999; Young et al., 2000) are available on the HRAP grid, a primary reason for its use by HL-RDHM. If higher-resolution data and model parameters are available, it is possible to run HL-RDHM on a finer grid. For PE, monthly climatology is used (Smith et al., 2004). Figure 1 shows a schematic of the gridded SAC and kinematic-wave routing models of the HL-RDHM (Lee et al., 2011). The a priori estimates of the SAC parameters (Koren et al., 2000) are derived from the soil data, STATSGO2 (NRCS, 2006) and SSURGO (NRCS, 2004). Optimisation of the a priori parameters is carried out via manual or automatic calibration (Koren et al., 2004). For the four Oklahoma basins, we used the manually-optimised

parameters used in the Distributed Model Intercomparison Project (DMIP, Smith et al., 2004). For the five Texas basins, we used the manual calibration results from WGRFC. The routing parameters are estimated from the DEM, channel hydraulic data and observed flow data (Koren et al., 2004). The flow direction from upstream to downstream HRAP grid cells is determined by the Cell Outlet Tracing with an Area Threshold (COTAT) algorithm (Reed, 2003) using the Digital Elevation Model (DEM) data. We note here that, in this study, streamflow predictions were generated assuming perfectly known forcing, i.e., using the historical observed forcing data rather than forecast forcing because our primary interest here is in reducing hydrologic uncertainty (Krzysztofowicz, 1999; Seo et al., 2006).

2.2 DA procedure

The automated DA procedure used in this study is based on the variational DA (VAR) technique. There are a number of reasons for this choice among different DA techniques. While simpler to implement, EnKF is optimum only if the observation equation is linear, which is easily violated when assimilating streamflow for soil moisture updating. On the other hand, the VAR technique is optimum in the least squares sense, even if observation equations are strongly nonlinear (Zhang et al., 2001). Also, since the VAR procedure is smoother than a filter and, hence, equivalent to ensemble Kalman smoother, but with an ability to handle nonlinear observation equations, it can easily account for the time lag due to flow routing (Seo et al., 2003; Clark et al., 2008; Pauwels and De Lannoy, 2006; Weerts and El Serafy, 2006). In theory, one may use the particle filter to overcome the linearity or distributional assumptions (Doucet et al., 2001; Pham, 2001). In reality, however, particle filtering is computationally prohibitively expensive for high-dimensional problems such as the one dealt with in this work.

In the following, we formulate the DA problem for the gridded SAC and kinematic-wave routing models of HL-RDHM, which may be stated as follows: *Given the a priori SAC states at the beginning of the assimilation window and observations/estimates of precipitation, potential evapotranspiration (PE) and streamflow at the outlet and/or interior locations, update the state variables of the gridded SAC and kinematic-wave routing models by adjusting the initial SAC states and multiplicative biases for precipitation and PE over the assimilation window at the predefined spatiotemporal scales of adjustment.*

The VAR technique formulates the above as a least-squares minimisation problem that minimises the objective function J constrained by the model physics (Lewis et al., 2006; Liu and Gupta, 2007):

Minimize

$$J_K(X_{S,K-L}, X_{P,k}, X_{E,k}, X_{W,k}) = \frac{1}{2} \sum_{k=K-L+1}^K [Z_{Q,k} - \mathbf{H}_{Q,k}(X_{S,K-L}, X_{P,K-L+1:k}, X_{E,K-L+1:k}, X_{W,K-L+1:k})]^T \mathbf{R}_{Q,k}^{-1} [Z_{Q,k} - \mathbf{H}_{Q,k}(X_{S,K-L}, X_{P,K-L+1:k}, X_{E,K-L+1:k}, X_{W,K-L+1:k})] + \frac{1}{2} \sum_{k=K-L+1}^K [Z_{P,k} - \mathbf{H}_{P,k}X_{P,k}]^T \mathbf{R}_{P,k}^{-1} [Z_{P,k} - \mathbf{H}_{P,k}X_{P,k}] + \frac{1}{2} \sum_{k=K-L+1}^K [Z_{E,k} - \mathbf{H}_{E,k}X_{E,k}]^T \mathbf{R}_{E,k}^{-1} [Z_{E,k} - \mathbf{H}_{E,k}X_{E,k}] + \frac{1}{2} [Z_{B,K-L} - \mathbf{H}_B X_{S,K-L}]^T \mathbf{R}_{B,K-L}^{-1} [Z_{B,K-L} - \mathbf{H}_B X_{S,K-L}] + \frac{1}{2} \sum_{k=K-L+1}^K X_{W,k}^T \mathbf{R}_{W,k}^{-1} X_{W,k} \quad (1)$$

$$\text{subject to } \begin{cases} X_{S,k} = M(X_{S,k-1}, X_{P,k}, X_{E,k}, X_{W,k}), \\ k = K-L+1, \dots, K \\ X_{S,j,i}^{\min} \leq X_{S,j,i,k} \leq X_{S,j,i}^{\max} \\ k = K-L, \dots, K; j = 1, \dots, n_S; i = 1, \dots, n_C \end{cases} \quad (2)$$

The objective function presented above is based on the following observation equations:

$$Z_{B,K-L} = \mathbf{H}_B X_{S,K-L} + V_{B,K-L} \quad (3)$$

$$Z_{P,k} = \mathbf{H}_{P,k} X_{P,k} + V_{P,k} \quad (4)$$

$$Z_{E,k} = \mathbf{H}_{E,k} X_{E,k} + V_{E,k} \quad (5)$$

$$Z_{Q,k} = \mathbf{H}_{Q,k}(X_{S,K-L}, X_{P,K-L+1:k}, X_{E,K-L+1:k}, X_{W,K-L+1:k}) + V_{Q,k} \quad (6)$$

In Eqs. (1) to (6), $X_{S,k-1}$, $X_{S,k}$, and $X_{S,K-L}$ denote the five SAC states (UZTWC, UZFWC, LZTWC, LZFWC, LZFPWC) at hour $k-1$, k , and $K-L$, respectively; $X_{P,k}$ and $X_{E,k}$ denote the multiplicative adjustment factors for biases in precipitation and PE at hour k , respectively; $Z_{B,K-L}$, $Z_{P,k}$, $Z_{E,k}$, and $Z_{Q,k}$ denote the observations of SAC states at the beginning of the assimilation window, precipitation, PE, and streamflow, respectively; $\mathbf{H}_{P,k}$ and $\mathbf{H}_{E,k}$ are the same as $Z_{P,k}$ and $Z_{E,k}$ (this follows from the fact that $X_{P,k}$ and $X_{E,k}$ are multiplicative adjustment factors), respectively; $\mathbf{H}_{Q,k}$ represents the gridded SAC and kinematic-wave routing models; \mathbf{H}_B is the identity matrix; $X_{W,k}$ denotes the model error; $V_{B,K-L}$, $V_{P,k}$, $V_{E,k}$ and $V_{Q,k}$ denote the measurement error vectors associated with $Z_{B,K-L}$, $Z_{P,k}$, $Z_{E,k}$, and $Z_{Q,k}$, respectively; $X_{P,K-L+1:k}$ denotes $X_{P,K-L+1}$, $X_{P,K-L+2}$, ..., $X_{P,k}$; $X_{E,K-L+1:k}$ denotes $X_{E,K-L+1}$, $X_{E,K-L+2}$, ..., $X_{E,k}$; $X_{W,K-L+1:k}$ denotes $X_{W,K-L+1}$, $X_{W,K-L+2}$, ..., $X_{W,k}$; $\mathbf{R}_{P,k}$, $\mathbf{R}_{E,k}$, $\mathbf{R}_{Q,k}$, $\mathbf{R}_{B,K-L}$, and $\mathbf{R}_{W,k}$ represent the observation error covariance matrices associated with $Z_{P,k}$, $Z_{E,k}$, $Z_{Q,k}$, $Z_{B,K-L}$, and a priori estimates of the model error, respectively.

In simplifying the above minimisation problem, we drop the model error, $X_{W,k}$ because, in reality, little is known about its statistical properties. Seemingly an oversimplification, such strong-constraint formulation (Zupanski, 1997) is still very reasonable for our problem because $V_{P,k}$ and $V_{E,k}$ in Eqs. (4) and (5) act like model errors to a certain extent

(Seo et al., 2003). We then assume that the observation errors are independent and time-invariant (Seo et al., 2003) so that $\mathbf{R}_{P,k}$, $\mathbf{R}_{E,k}$, $\mathbf{R}_{Q,k}$ and $\mathbf{R}_{B,K-L}$ become diagonal and static. This assumption significantly reduces statistical modelling and computational requirements. Eq. (7) shows the resulting objective function used in this work:

Minimize

$$J_K(\lambda_{j,i}, X_{P,k}, X_{E,k}) = \frac{1}{2} \sum_{k=K-L+1}^K \sum_{l=1}^{n_Q} [Z_{Q,l,k} - H_{Q,l,k}(X_{S,K-L}, X_{P,K-L+1:k}, X_{E,K-L+1:k})]^2 \sigma_{Q,l}^{-2} + \frac{1}{2} \sum_{k=K-L+1}^K Z_{P,k}^2 [1 - X_{P,k}]^2 \sigma_P^{-2} + \frac{1}{2} \sum_{k=K-L+1}^K Z_{E,k}^2 [1 - X_{E,k}]^2 \sigma_E^{-2} + \frac{1}{2} \sum_{j=1}^{n_S} \sum_{i=1}^{n_C} [Z_{B,j,i,K-L} - \lambda_{j,i} Z_{B,j,i,K-L}]^2 \sigma_{B,j,i}^{-2} \quad (7)$$

$$\text{subject to } \begin{cases} X_{S,k} = \mathbf{M}(X_{S,k-1}, X_{P,k}, X_{E,k}), \\ k = K - L + 1, \dots, K \\ X_{S,j,i}^{\min} \leq X_{S,j,i,k} \leq X_{S,j,i}^{\max}, \\ k = K - L, \dots, K; j = 1, \dots, n_S; i = 1, \dots, n_C \end{cases} \quad (8)$$

Equations (7) and (8) pose a nonlinear constrained least-squares minimisation problem with the model dynamics as a strong constraint. In Eqs. (7) and (8), n_Q denotes the number of stream gauge stations, $Z_{Q,l,k}$ denotes the streamflow observation at the l -th gauge station at hour k , and $Z_{B,j,i,K-L}$ denotes the background (i.e., the a priori or before-DA) model soil moisture state associated with the j -th state variable and i -th cell at the beginning of the assimilation window, $H_{Q,l,k}(\cdot)$ denotes the observation operator that maps $X_{S,K-L}$ to streamflow at the l -th gauge station and hour k , $X_{S,K-L}$ denotes the SAC states at hour $K - L$, $\sigma_{Q,l}$ denotes the standard deviation of the streamflow observation error at the l -th stream gauge location, σ_P and σ_E denote the error standard deviations of observed precipitation and PE, respectively, $\sigma_{B,j,i}$ denotes the standard deviation of the error associated with the j -th background model state at the i -th grid, and $\lambda_{j,i}$ denotes the multiplicative adjustment factor to $Z_{B,j,i,K-L}$. The vector $X_{S,K-L}$ consists of $\lambda_{j,i} Z_{B,j,i,K-L}$.

At the beginning of the minimisation, the control variables, $X_{P,k}$, $X_{E,k}$, and $\lambda_{j,i}$, are set to unity for all i , j , and k . During the minimisation, we allow $X_{P,k}$ and $X_{E,k}$ to vary hourly or 6-hourly or keep them constant over the entire assimilation window and $\lambda_{j,i}$ to be adjusted at each cell, uniformly over each sub-catchment or over the entire basin. The computation time for the model simulation was not very sensitive to the spatiotemporal scale of adjustment. Equations (7) and (8) are solved using the Fletcher-Reeves-Polak-Ribiere minimisation (FRPRMN) algorithm (Press et al., 1992), a conjugate gradient method. Gradients of the objective function with respect to the control vector were

calculated using the adjoint code generated from Tapenade (<http://tapenade.inria.fr:8080/tapenade/index.jsp>).

2.3 Evaluation metrics

The performance of DA procedure is evaluated using correlation coefficient (r), skill score (SS), root-mean-square-error (RMSE), and timing error (TE). We developed two types of correlation-based matrices ($\mathbf{r1}$, and $\mathbf{r2}$) as defined in Eqs. (9) to (10) below. The $\mathbf{r1}$ -matrix defines spatial (i.e., interstation) correlation of streamflow, either observed or simulated, at paired gauge locations. The $\mathbf{r2}$ -matrix compares differences in spatial correlation between observed and simulated streamflow in off-diagonal entries, and defines correlation between the two at the same locations in diagonal entities.

$$\mathbf{r1}(Q) = R(Q_i, Q_j) \text{ for all } i \text{ and } j \quad (9)$$

$$\mathbf{r2}(Q_s^-, Q_o) = \begin{cases} R(Q_{s,i}^-, Q_{s,j}^-) - R(Q_{o,i}, Q_{o,j}) & \text{if } i \neq j \\ R(Q_{s,i}^-, Q_{o,j}) & \text{if } i = j \end{cases} \quad (10)$$

where R denotes the operator for the Pearson's correlation coefficient between the two streamflow time series; $Q_{s,i}^-$ and $Q_{s,j}^-$ denote the simulated flow (without assimilation) at gauges i and j , respectively; $Q_{o,i}$ and $Q_{o,j}$ denote the observed flow at gauges i and j , respectively; Q in Eq. (9) can be either Q_s^- or Q_o ; subscripts i and j denote the indices for the stream gauges at interior or outlet locations.

The Skill Score (SS; Murphy, 1996) is calculated based on the summed squared errors of simulated streamflow before and after assimilation:

$$SS = 1 - \frac{\sum_{k=k1}^{k2} (Q_{s,k}^+ - Q_{o,k})^2}{\sum_{k=k1}^{k2} (Q_{s,k}^- - Q_{o,k})^2} \quad (11)$$

In the above, k denotes the time index, $Q_{s,k}^-$ and $Q_{s,k}^+$ denote the simulated streamflow valid at time k before and after assimilation, respectively; $Q_{o,k}$ denotes the streamflow observation valid at time k . A positive SS means improvement after assimilation and the opposite for a negative SS. The SS value is 1 if DA is perfect and 0 if DA adds nothing.

Root-Mean-Square-Error (RMSE) of streamflow is calculated by Eq. (12) where $Q_{s,k}$ denotes either $Q_{s,k}^-$ or $Q_{s,k}^+$.

$$RMSE = \sqrt{\frac{1}{k2 - k1 + 1} \sum_{k=k1}^{k2} (Q_{s,k} - Q_{o,k})^2} \quad (12)$$

Timing Error (TE) in streamflow simulation is represented by the phase difference between observed and simulated hydrographs as computed by a wavelet-based technique (Liu et al., 2011).

$$TE = \frac{T}{2\pi} \tan^{-1} \left(\frac{\Im \left(\left(s^{-1} W_n^{XY}(s) \right) \right)}{\Re \left(\left(s^{-1} W_n^{XY}(s) \right) \right)} \right) \quad (13)$$

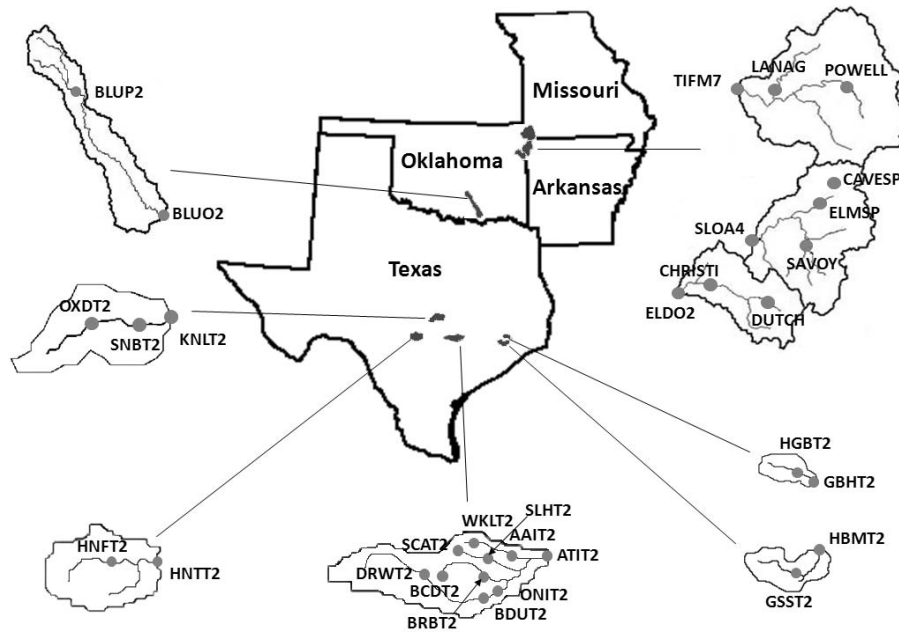


Fig. 2. Map showing the locations of study basins, channel network and stream gauges.

where T denotes the equivalent Fourier period of the wavelet; s and n denote the scale and location parameter of the wavelet, respectively; $W_n^{XY}(s)$ denotes the cross wavelet spectrum of the two time series X and Y ; $\Im(\cdot)$ and $\Re(\cdot)$ denote the imaginary and real parts of the variable bracketed, respectively; $\langle \cdot \rangle$ denotes the smoothing operation in both time and frequency domains (Torrence and Compo, 1998). The TE has a unit of time, e.g., h. A smaller TE means better model performance. In our study, a positive/negative TE means that the simulated hydrograph leads/trails the observed hydrograph. Compared to small basins, large basins may produce large TE because of longer travel time. Further details on the wavelet-based timing error estimation technique are found in Liu et al. (2011).

3 Study basins

Figure 2 shows the nine basins used in this study, and Table 1 provides additional details on the data used for each basin. In Fig. 2, ELDO2 and SLOA4 are nested in the Illinois River basin located near the border of Oklahoma (OK) and Arkansas (AR); TIFM7 is a part of the Elk River basin near the border of Missouri (MO) and Arkansas (AR); BLUO2 is a headwater basin to the Blue River in southern Oklahoma (OK). These four basins are located in the service area of the Arkansas-Red Basin River Forecast Centre (ABRFC). The other five basins, GBHT2, HBMT2, ATIT2, KNLT2, HNTT2, are located in Texas (TX) in the service area of the West Gulf River Forecast Centre (WGRFC). Topography of the ABRFC basins ranges from gently rolling to hilly with the maximum elevation difference between the basin

outlet and the interior exceeding 200 m (Smith et al., 2004). In contrast, topography of the WGRFC basins is generally characterised as flat to very flat (Vieux, 2001). Very large runoff coefficients for HBMT2 and GBHT2 are due mainly to the large urbanised areas around Houston, TX (Liscum, 2001). In particular, HBMT2 has an extremely large runoff coefficient due to the combined effect of 85 % of the watershed area being highly developed, clayey soils with low infiltration rates, and the lower 42 km of the channel being lined with concrete (Vieux, 2001). The basins, BLUO2, KNLT2, HNTT2 and ATIT2 are relatively dry with annual precipitation of less than 850 mm and runoff coefficients of less than 0.14 (Table 1). As with HBMT2, these four basins are also largely covered by clayey soils. Morphologically, BLUO2 is very elongated. SLOA4 and TIFM7 have a radial channel network with tributaries with similarly-sized drainage areas. Figure 3 shows the maps of delineated sub-basins, the soil type and mean event precipitation on the HRAP grid for each basin; selected flood events summarised in the Table 2 are used to calculate mean event precipitation. In Fig. 3, sub-basins were delineated based on the channel connectivity information derived from the COTAT algorithm (Reed, 2003) and an area threshold for channel cell identification, which delineates the channel network the most similar to the actual channel network. Inter-grid variability of mean event precipitation ranges from 12 (HBMT2) to 90 mm (HNTT2). The basins BLUO2 and HNTT2 show a clearer pattern of spatial variability of precipitation than the other basins. Mean event precipitation in the upper half of the BLUO2 basin is approximately 25 mm smaller than that in the lower half. Each basin has one or more interior stream

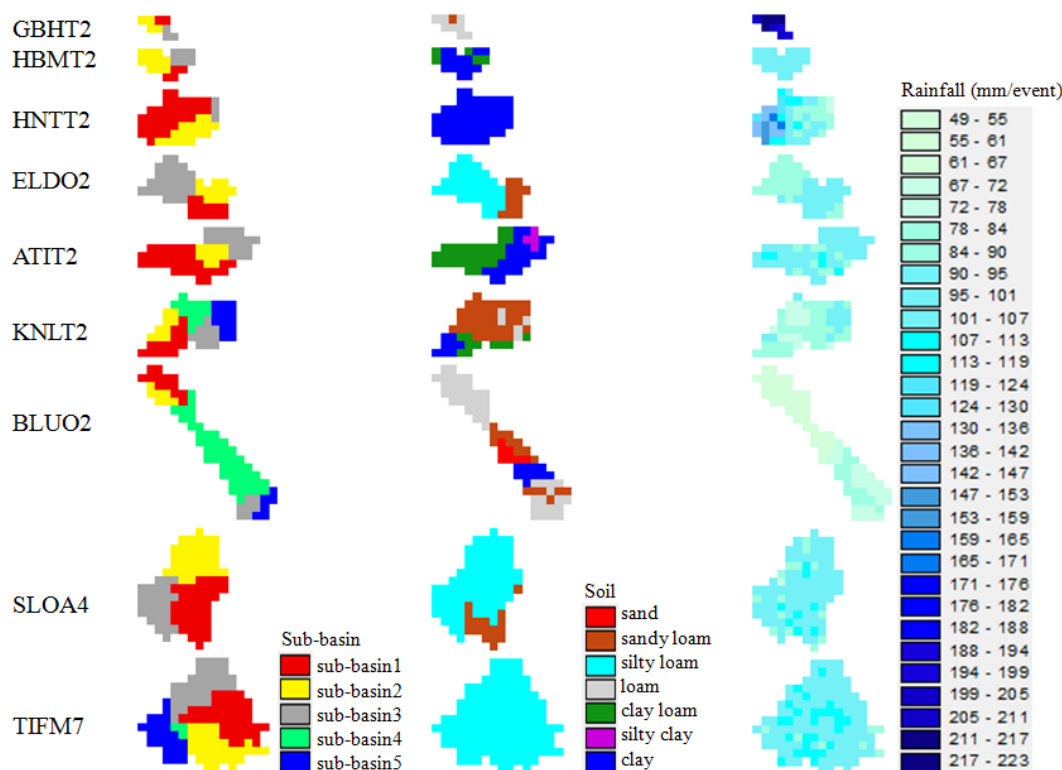


Fig. 3. Map of delineated sub-basins, soil type and mean accumulated rainfall per event.

gauges (nine for ATIT2). The drainage area ranges from 137 (GBHT2) to 2258 (TIFM7) km².

4 Streamflow DA experiments

4.1 Experimental design and procedure

Simulation experiments were carried out in which streamflow data were assimilated into the distributed SAC-SMA at the pre-specified spatiotemporal scale of adjustment. Three streamflow assimilation scenarios are considered: outlet flow assimilation, interior flow assimilation, and outlet and interior flow assimilation. The experiment is designed to investigate: (1) the effect of spatiotemporal adjustment scale on streamflow analysis and prediction, (2) the sensitivity of the optimum spatiotemporal adjustment scale to the streamflow assimilation scenario, and (3) the performance of the DA procedure at the optimum spatiotemporal adjustment scale. The experiment is composed of the following four steps:

- Step 1: carry out the base model simulation (i.e., without assimilation) and evaluate its performance on streamflow simulation.
- Step 2: estimate the observational error variances.
- Step 3: given a spatiotemporal adjustment scale, assimilate streamflow observations into the model for each of

the three assimilation scenarios, i.e., assimilation of outlet flow, of interior flow, and of both outlet and interior flows.

- Step 4: repeat Step 3 for each of the nine spatiotemporal adjustment scales.

In Step 2, the sensitivity of the performance of DA on streamflow observational error variance (σ_Q^2) is examined to obtain an optimum σ_Q^2 . In these sensitivity runs, seven different values for σ_Q^2 (0.01, 0.1, 1, 10, 100, 1000, 10000 ($\text{m}^3 \text{s}^{-1}$)²) were used for each of three streamflow assimilation scenarios. The results show that $\sigma_Q^2 = 10$ ($\text{m}^3 \text{s}^{-1}$)² yields the best results for streamflow analysis and prediction in terms of RMSE for all basins except TIFM7, for which $\sigma_Q^2 = 100$ ($\text{m}^3 \text{s}^{-1}$)² was better. For each basin, the optimum σ_Q^2 showed largely insensitive to the assimilation scenario possibly because similar properties associated with flow processes and channel geometry at upstream and downstream locations in the same basin result in the similar amount of error in the estimation of the rating curve and the rating curve-to-flow conversion at those locations. Observational error variances for precipitation and PE are taken directly from Seo et al. (2003). Sample variances calculated from the base model simulation for the entire period of record were used as error variances for background model states (Lee et al., 2011). We assumed that the streamflow observation errors

Table 1. Study basins where A denotes drainage area, N_G the number of interior stream gauges in a basin, \bar{P} mean annual precipitation, \bar{Q} mean annual runoff, C runoff coefficient.

Location of stream gauge at the basin outlet	Basin and sub-basin name	A (km ²)	USGS ID	N_G	Period of record	\bar{P} (mm yr ⁻¹)	\bar{Q} (mm yr ⁻¹)	C
Baron Fork at Eldon, OK	ELDO2	795	7197000	2	Jan 1996–	1163	371	0.32
	DUTCH	105	7196900					
	CHRISTI	65	7196973		Jan 2004			
Illinois River South of Siloam Springs, AR	SLOA4	1489	7195430	3	Apr 2000– Jan 2002	1324	383	0.29
	SAVOY	433	7194800					
	ELMSP	337	7195000					
Elk river near Tiff City, MO	TIFM7	2258	7189000	2	May 2000– Sep 2006	1117	246	0.22
	LANAG	619	7188885					
	POWELL	365	7188653					
Blue river near Blue, OK	BLUO2	1232	7332500	1	Oct 2003– Sep 2006	846	117	0.14
	BLUP2	419	7332390					
Brays Bayou at Houston, TX	HBMT2	246	08075000	1	Jan 1997– Jul 2009	1202	1124	0.94
	GSST2	136	08074810					
Greens Bayou near Houston, TX	GBHT2	137	08076000	1	Jan 2000– Jul 2009	1467	944	0.64
	HGBT2	95	08075900					
Sandy Creek near Kingsland, TX	KNLT2	904	08152000	2	Oct 1997– Sep 2008	767	68	0.09
	SNBT2	401	*					
	OXDT2	381	*					
Guadalupe River at Hunt, TX	HNTT2	769	08165500	1	Jan 1998– Jun 2009	697	82	0.12
	HNFT2	438	08165300					
Onion Creek at US Hwy 183, Austin, TX	ATIT2	844	08159000	9	Jan 1997– Jun 2009	752	96	0.13
	ONIT2	469	08158827					
	BDUT2	437	*					
	DRWT2	321	08158700					
	BRBT2	62	08158819					
	SLHT2	60	08158860					
	AAIT2	49	08158930					
	BCDT2	32	08158810					
	SCAT2	21	08158840					
WHLT2	16	08158920						

* denotes stream gauges operated by Lower Colorado River Authority.

are homoscedastic and that the observation errors for precipitation and PE are homogeneous in space. These assumptions may be lifted in the future in order to more effectively constrain the assimilation problem, relying on advances in uncertainty techniques that properly parameterise and quantify uncertainty associated with stage measurement, stage to discharge conversion, and spatial correlation of forcing error (Clark et al., 2008; Mandapaka et al., 2009).

4.2 Results and discussion

In this subsection, the experiment results are comparatively evaluated. We focus on analysis vs. prediction and dependent vs. independent validation to address the questions associated with overfitting due to large degrees of freedom in distributed modelling.

Table 2. The length of assimilation window, the number of sub-basins delineated from the channel connectivity map, the number of flood events denoted as N_F and the threshold of streamflow (Q_T) used to identify flood events.

Basin name	ELDO2	SLOA4	TIFM7	BLUO2	HBMT2	GBHT2	KNLT2	HNTT2	ATIT2
Assimilation window length (h)	36	48	60	60	42	48	36	30	36
No. of sub-basins	3	3	5	5	3	3	5	3	3
Q_T ($m^3 s^{-1}$)	200	200	200	100	400	150	200	200	100
N_F	17	7	15	7	20	16	15	9	23

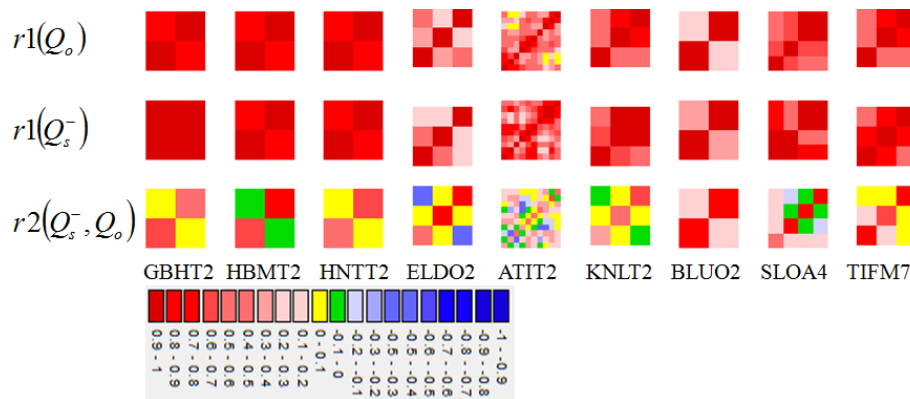


Fig. 4. Spatial correlation structure of the streamflow processes.

4.2.1 Analysis of the assimilation problem

Prior to assimilation, we assess for each basin the level of complexity of the assimilation problem by examining the spatial correlation structure of observed and base-simulated streamflow, and the basin characteristics such as spatial heterogeneity of soil and precipitation. Figure 4 presents the correlation-based matrices of streamflow. In Fig. 4, the correlation coefficients were calculated using streamflow data at any paired gauges (i.e., interior and outlet as well as interior and interior). The data were paired at concurrent time steps due to the difficulty of correctly estimating travel time for all paired gauges. Correlations of time-lagged simulated interior and outlet flow as a function of a lag time closely followed those based on streamflow observations. This supports the idea of using correlation matrices in Fig. 4 for analysing spatial correlation structure of streamflow. The 1st row of Fig. 4 presents the **r1**-matrices (see Eq. 9) showing the spatial correlation of observed streamflow. In all correlation matrices, the stream gauges are sorted in the increasing order of the drainage area starting from the bottom-left corner. Note in Fig. 4 that, for most basins, observed streamflow at the outlet is highly correlated with that at interior locations. For BLUO2, the low correlation between the interior and outlet flows may be due to the distance between the two and large variability in precipitation. For ELDO2,

the weak spatial correlation in flow between the outlet and DUTCH may be contributed by the different soil types. For ATIT2, the upstream flows at some interior gauges, particularly SCAT2 and BCDT2, are weakly correlated with downstream flows at BDUT2 and at the outlet. This may be due to the small drainage areas involved and the locations of SCAT2 and BCDT2 being on minor tributaries. The 2nd row in Fig. 4 shows the **r1**-matrices of streamflow from base model simulation, and the 3rd row in Fig. 4 presents **r2**-matrices (Eq. 10) of observed and simulated flows prior to assimilation, respectively. Both **r1**- and **r2**-matrices in the 2nd and 3rd rows in Fig. 4 indicate that the model simulation generally well reproduces the spatial correlation of streamflow at two locations in a basin, particularly for GBHT2, HBMT2, HNTT2 and KNLT2 for which the differences in correlation between observed and simulated streamflow (off-diagonal terms in **r2**) are less than 0.1. In addition to the absolute value of **r2** off-diagonal terms, the unity of their signature is treated as another information associated with the degree of complexity of the assimilation problem; that is, overall overestimation or underestimation of spatial correlation structure of the streamflow is considered less ill-posed than combination of over- and under-estimation. In the latter case, independent validation results posterior to the assimilation may benefit less at any adjustment scales than the former due to

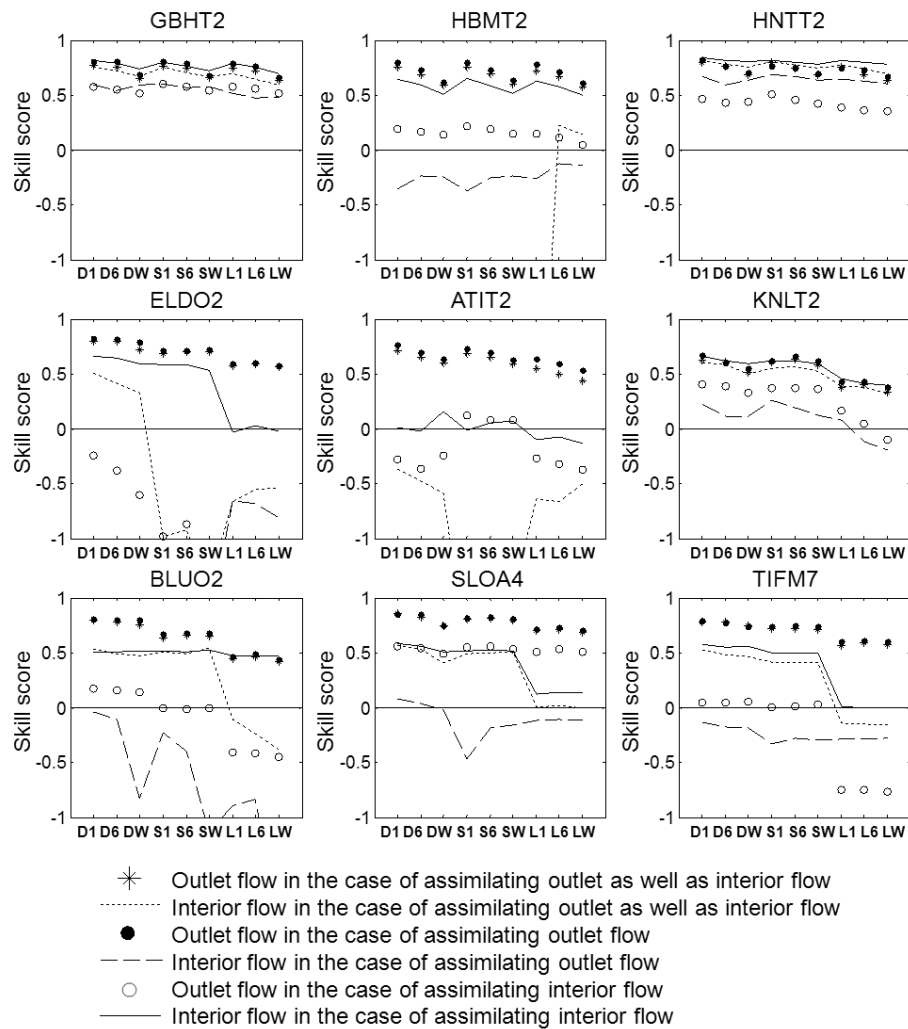


Fig. 5. Mean skill score of streamflow analysis where mean skill score is obtained by averaging mean squared error-based skill score calculated for individual event (D: distributed, S: semi-distributed, L: lumped, 1: 1-h, 6: 6-h, W: the length of time equal to that of the assimilation window).

the interference of the correlation to the assimilation procedure in a complicated way. In this regard, ELDO2, ATIT2, KNLT2 and SLOA4 can be viewed as more ill-posed than the other basins.

4.2.2 Effect of spatiotemporal adjustment scale on the performance of the DA procedure

Figures 5 and 6 show the mean SS for streamflow analysis and prediction, respectively, for all assimilation scenarios and adjustment scales. As described in Sect. 2.1, streamflow prediction with assimilation is made with updated state variables at the prediction time and historical observed precipitation data and monthly climatology of PE over the forecasting window. For streamflow analysis (Fig. 5), the mean SS is calculated by averaging the SS values calculated at every hour within the assimilation window for each event and for each

gauge location separately (Eq. 14). For streamflow prediction (Fig. 6), the mean SS is calculated by averaging the SS values calculated at every hourly lead time up to 6 hours for each event and for each gauge location separately (Eq. 14).

$$\text{mean SS} = \frac{1}{N_T N_G N_F} \sum_{\tau=1}^{N_T} \sum_{j=1}^{N_G} \sum_{i=1}^{N_F} \text{SS}_{i,j,\tau} \quad (14)$$

In the above, $\text{SS}_{i,j,\tau}$ denotes the skill score (Eq. 11) calculated for the i -th event, j -th gauge, τ -th hour of lead time; N_T denotes the number of lead hours considered, e.g., N_T denotes the length of the assimilation window (Table 2) for Fig. 5, and $N_T=6$ for Fig. 6; N_G denotes the number of gauges involved; N_F denotes the number of selected flood events for each basin (Table 2). Note that the mean SS presented in Figs. 5 and 6 equally weighs SS for each event. Figures 5 and 6 may be summarised as follows. The performance

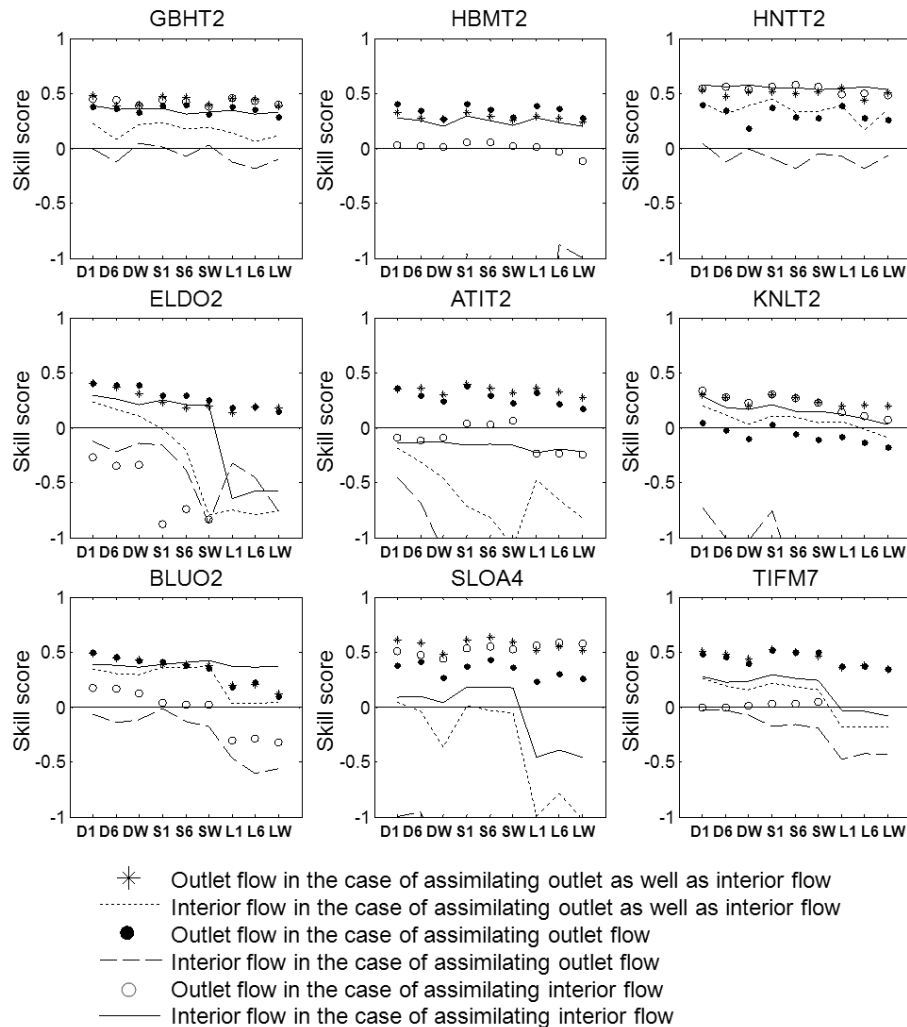


Fig. 6. Same as Fig. 5, but for streamflow prediction for 1- to 6-h lead time.

of DA is less sensitive to the temporal adjustment scale than the spatial adjustment scale. The basins with high spatial correlation between interior and outlet flows (GBHT2, HBMT2, HNTT2), show less sensitive in DA performance to the spatial adjustment scale than those with lower correlation. For BLUO2, lumped adjustment yields less improvement than other assimilation cases due possibly to the low spatial correlation of interior and outlet flows. In a number of independent validation cases (i.e., validating assimilation results with streamflow data not used in the assimilation), the mean SS for streamflow analysis is less than zero, suggesting overfitting. For GBHT2, HNTT2 and KNLT2, assimilating interior flows produced positive mean SS for streamflow prediction for the first 6 h of lead time at both interior and outlet locations. However, assimilating outlet flow generally degrades interior flow prediction for most basins, compared to the base model simulation. This implies assimilating interior flow makes the DA problem less subject to overfitting. Note that some events

are affected by timing errors in the model simulation which are partially responsible for small to negative mean SS for some cases. We further discuss timing errors at the end of this section.

To further examine the sensitivity of DA performance to the adjustment scale, Figs. 7 and 8 show the box-and-whiskers plot of the mean SS shown in Figs. 5 and 6, respectively. In Figs. 7 and 8, each box-and-whiskers plot is constructed with 27 samples resulted from the combinations of nine basins and three (space or time) scales. Figures 7 and 8 can be summarised as follows. The performance of DA is generally higher at finer adjustment scales and is more sensitive to the spatial adjustment scale than the temporal adjustment scale in terms of both the median SS and the interquartile range of the SS. The DA performance greatly depends on the streamflow assimilation scenario, i.e., assimilating outlet and/or interior flow data. Assimilating outlet flow does not improve interior flow simulation in most cases,

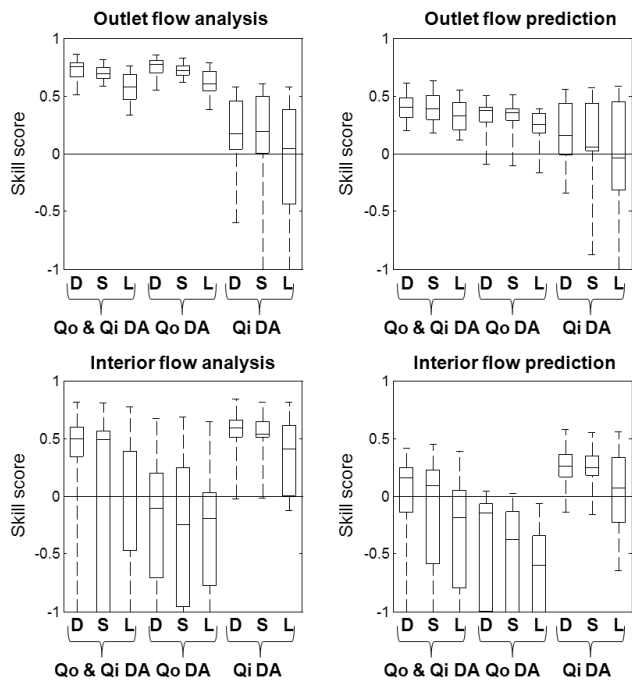


Fig. 7. Mean skill score vs. spatial adjustment scale where the mean skill score is obtained by averaging mean squared error-based skill score calculated for individual event. Mean skill score for the prediction period is calculated using streamflow predicted for 1- to 6-h lead time. In the above, D, S, and L denote distributed, semi-distributed and lumped ways of adjusting the SAC-states, respectively; Qo & Qi DA, Qo DA and Qi DA denote both outlet and interior flow assimilation, outlet flow assimilation and interior flow assimilation, respectively.

whereas, not surprisingly, assimilating interior flows typically improves outlet flow simulation to some degree. This indicates the difficulty of propagating the information contained in outlet flow data backward (i.e., upstream) through the stream network and the hydrologic processes involved to improve prediction of interior flow.

For the three different assimilation scenarios, “optimum” spatiotemporal scales are selected for interior and outlet flow predictions (Fig. 9). The selection is based on the mean SS for streamflow analysis or prediction shown in Figs. 5 and 6. Not surprisingly, a number of cases in streamflow analysis showed the largest improvement with the finest spatiotemporal scale of adjustment. For streamflow prediction, on the other hand, the optimum scale of adjustment is spread over a broader range. This indicates the possible large over-adjustment of state variables in the cases of distributed, hourly adjustment. Despite the issue associated with the over-fitting problem, the cases of distributed, hourly adjustment generally produce the best assimilation results in comparison to other adjustment scales.

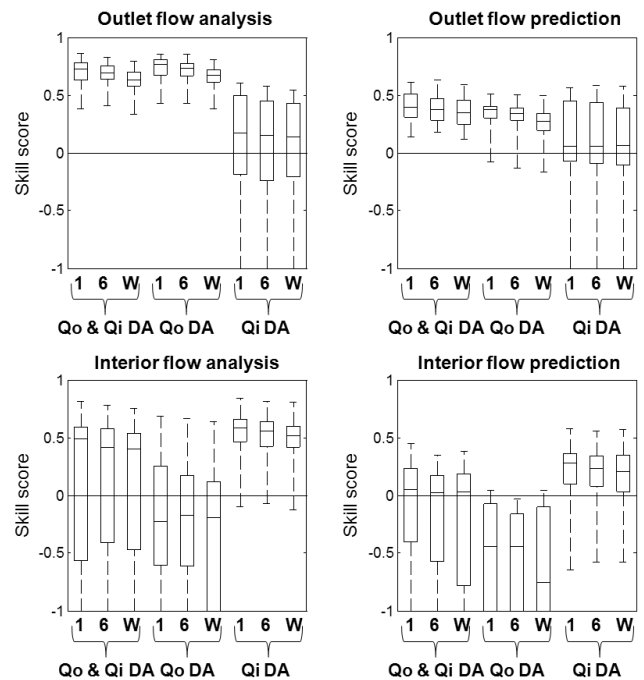


Fig. 8. Same as Fig. 7, but for the temporal adjustment scale. Here 1, 6 and W denote adjusting mean field bias in the precipitation and potential evaporation data on an 1-h or 6-h basis, or uniformly over the entire assimilation window, respectively.

4.2.3 Performance of the DA procedure at the optimum spatiotemporal adjustment scale

For more detailed quantitative analysis of the assimilation results, we chose a single optimum adjustment scale for each basin which produces reasonable assimilation results, based on mean SS in Figs. 5 and 6, for analysis and prediction of interior and outlet flows. The selected adjustment scales are semi-distributed and hourly for GBHT2, lumped and hourly for HBMT2 and distributed and hourly for all the other basins. Figure 10 shows the RMSE of streamflow analysis and prediction evaluated at every hour of lead time for each basin and for each assimilation scenario. Figure 11 shows the amount of reduction in the RMSE of streamflow analysis and prediction for all basins collectively. Figure 12 shows the timing error estimates from streamflow analysis and prediction for each assimilation scenario.

To diagnose streamflow analysis similarly to Fig. 4, we examined the spatial correlation structure of streamflow analysis at the optimum spatiotemporal adjustment scale (not shown). The spatial correlation between observed and simulated flows at both interior and outlet locations are generally improved for all basins by assimilating streamflow data, but at the expense of slightly adjusting spatial correlation structure of streamflow. Especially, for ELDO2, the spatial correlation between CHRISTIE and the outlet is improved noticeably after streamflow assimilation whereas, for ATIT2 and

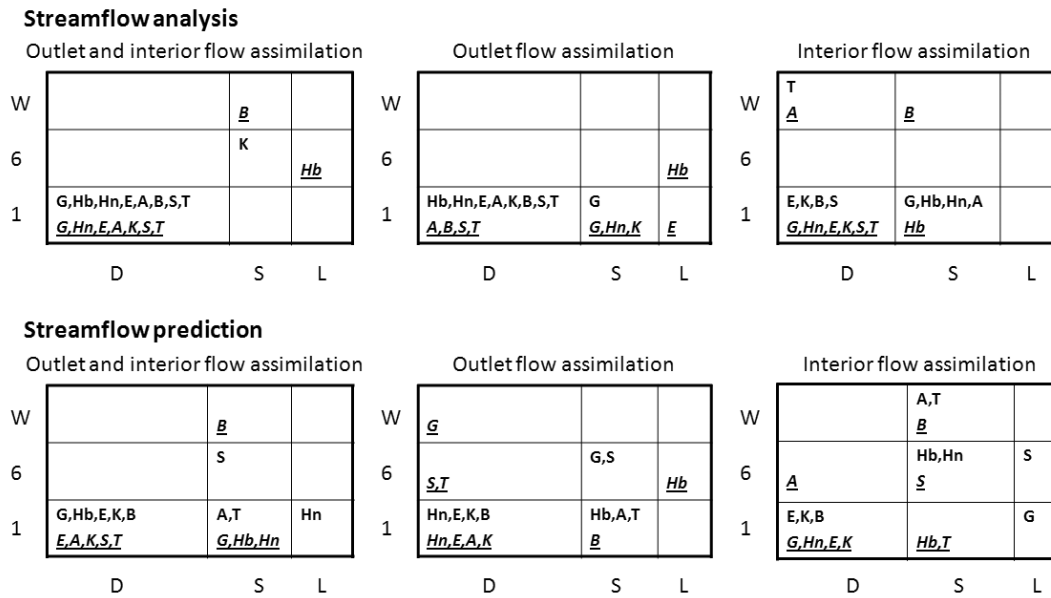


Fig. 9. Optimum spatiotemporal scales of adjustment for streamflow analysis and prediction for each basin and assimilation scenario. Under-scored italic letters represent interior flow results and the others represent outlet flow results (A: ATIT2, B: BLUO2, E: ELDO2, G: GBHT2, Hb: HBMT2, Hn: HNTT2, K: KNLT2, S: SLOA4, T: TIFM7).

KNLT2, the spatial correlation at some paired gauges was reduced after streamflow assimilation compared to that of base model simulation. This indicates that the performance gains from the DA do not always lead to improving the spatial correlation structure of the streamflow, a possible symptom of over-adjustment. In addition, examining the spatial correlation structure of streamflow at all adjustment scales indicated that, compared to outlet flow assimilation, interior flow assimilation produces correlation structure that is more consistent with that of observed streamflow. This may be explained by the local information available in interior flow observations which is diluted at the outlet location due to the various hydrologic and hydraulic processes involved.

Figure 10 shows the RMSE of streamflow as a function of lead time. The lead time is negative over the assimilation window and positive over the prediction horizon. For GBHT2, HBMT2 and HNTT2, all three assimilation scenarios improved both streamflow analysis and prediction. These basins show high spatial correlation between interior and outlet flows and the base model simulation reproduces the spatial correlation structure very well (see Fig. 4). For ill-posed basins ELDO2, ATIT2, KNLT2 and SLOA4, there is an indication of over-adjustment for streamflow analysis and prediction. For ELDO2, over-adjustment is not conspicuous possibly due to the smaller basin size and the relatively better base model simulation than the other ill-posed basins. For BLUO2, weak spatial correlation between interior and outlet flow may explain little improvement in streamflow at gauge locations where the data were not assimilated. The basin TIFM7 also shows similar results as BLUO2. It is noted that

for TIFM7 we use an observational error variance ten times larger than that for the other basins. This may have considerably reduced the amount of adjustment to state variables at most cells. Overall, assimilating streamflow data generally produces, expectedly, improved streamflow analysis and prediction at that gauge location. The margin of improvement at other locations varies, depending on the level of under-determinedness and basin characteristics.

Figure 11 shows the margin of reduction in the RMSE of streamflow analysis and prediction due to the assimilation versus the peak flow of selected events. To evaluate the overall performance of the VAR procedure, each plot is constructed with simulations from all nine basins. Assimilating interior flow yielded similar improvement (14 % reduction in RMSE with assimilation) to outlet flow assimilation in outlet flow prediction for the first 6 h of lead time. In contrast, in the case of outlet flow assimilation, gains in interior flow analysis (19 % reduction in RMSE after the assimilation) did not lead to improvement in multi-basin averaged skill in interior flow prediction over the first 6 lead hours, even though breakdown into each basin showed RMSE reduction by assimilation ranging from -31 % (ATIT2) to 14 % (GBHT2). This indicates that the outlet flow assimilation case is more vulnerable to overfitting than the interior flow assimilation case. Assimilating both outlet and interior flows outperforms the outlet flow assimilation case in terms of outlet flow prediction (22 % vs. 14 % reduction in RMSE with assimilation), although outlet flow analysis is less improved (55 % vs. 59 % reduction in RMSE with assimilation). This shows the value

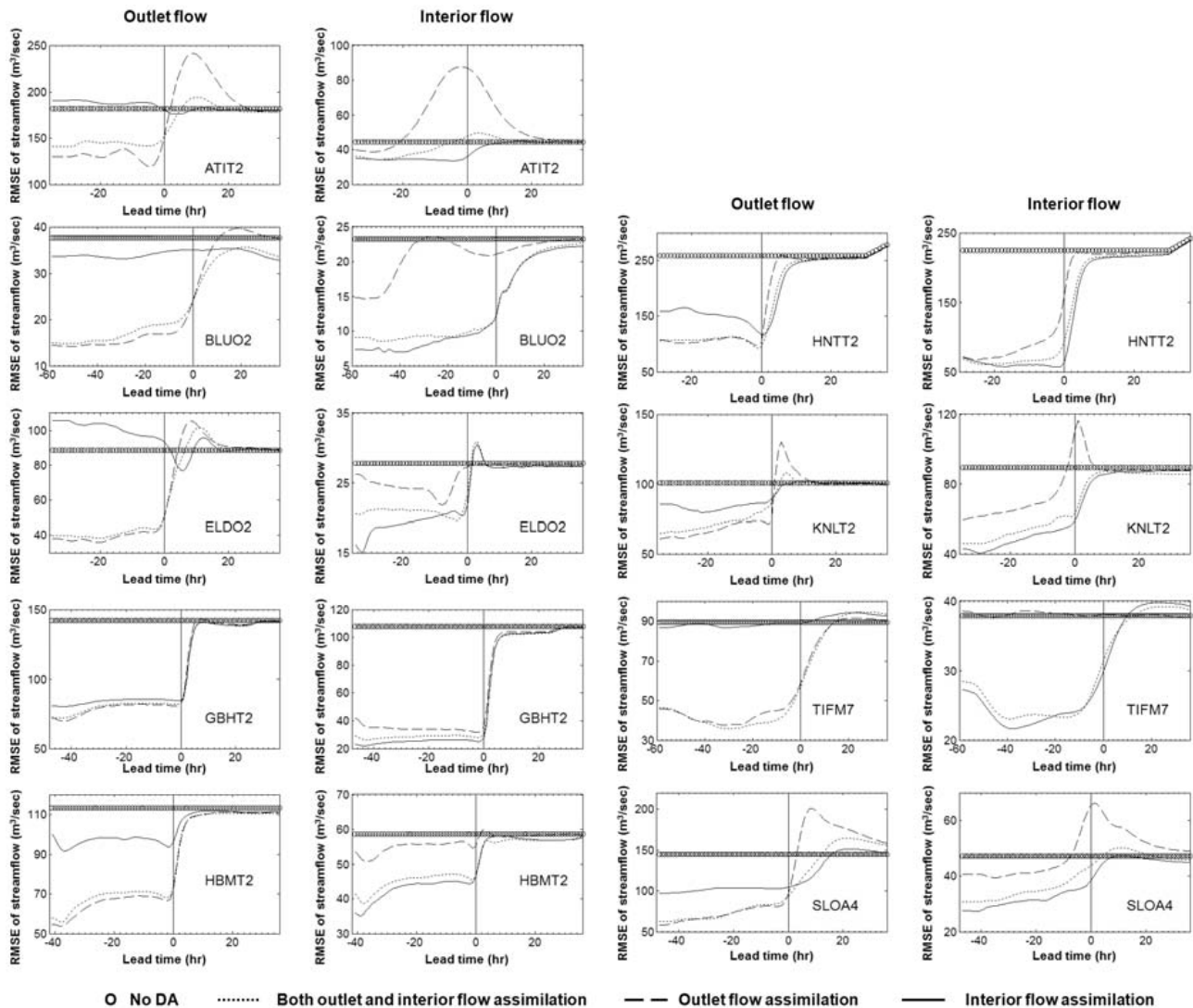


Fig. 10. RMSE of streamflow vs. lead hour where the lead hour is negative within the assimilation window.

of additionally assimilating interior flow for streamflow prediction at the basin outlet.

Phase (or timing) and flow magnitude are the two distinctive attributes in hydrograph evaluation (Liu et al., 2011). To examine the performance of DA, we also examine the timing error of a hydrograph within the assimilation and prediction windows separately as estimated via wavelet analysis (Liu et al., 2011) (see Eq. 13). Note that our timing error analysis is somewhat exploratory because of the objective equation, Eq. (7), used in this study, which includes no explicit timing error modelling component. Figure 12 shows the box-and-whiskers plots of the timing error estimates of simulated hydrographs that characterise inter-basin and inter-event variability. In Fig. 12, the reference is the event-scale timing error in the base simulation. On the whole, timing errors in the simulated hydrographs following assimilation at both outlet and interior stream gauge locations for

the assimilation period are generally smaller than the reference. While flow timing errors for the prediction period are less improved via streamflow assimilation, their medians are mostly free of timing error especially in the case of outlet flow. For events with significant timing errors in the rising limb, the assimilation procedure slightly improved the timing of streamflow analysis, but yielded significant magnitude errors in predicted flows. Examples of this are illustrated in Fig. 13. Note in the figure that the base model simulation for Event A shows significant timing errors in the rising limb, peak flow and the overall shape of the hydrograph, whereas Event B has smaller timing errors than Event A. The above situation arises due to lack of timing error modelling in the DA formulation used in this study (see Eq. 7). As a result, the VAR procedure over-adjusts control variables to compensate for timing errors in streamflow analysis which result in magnitude errors in predicted flows. Further analysis

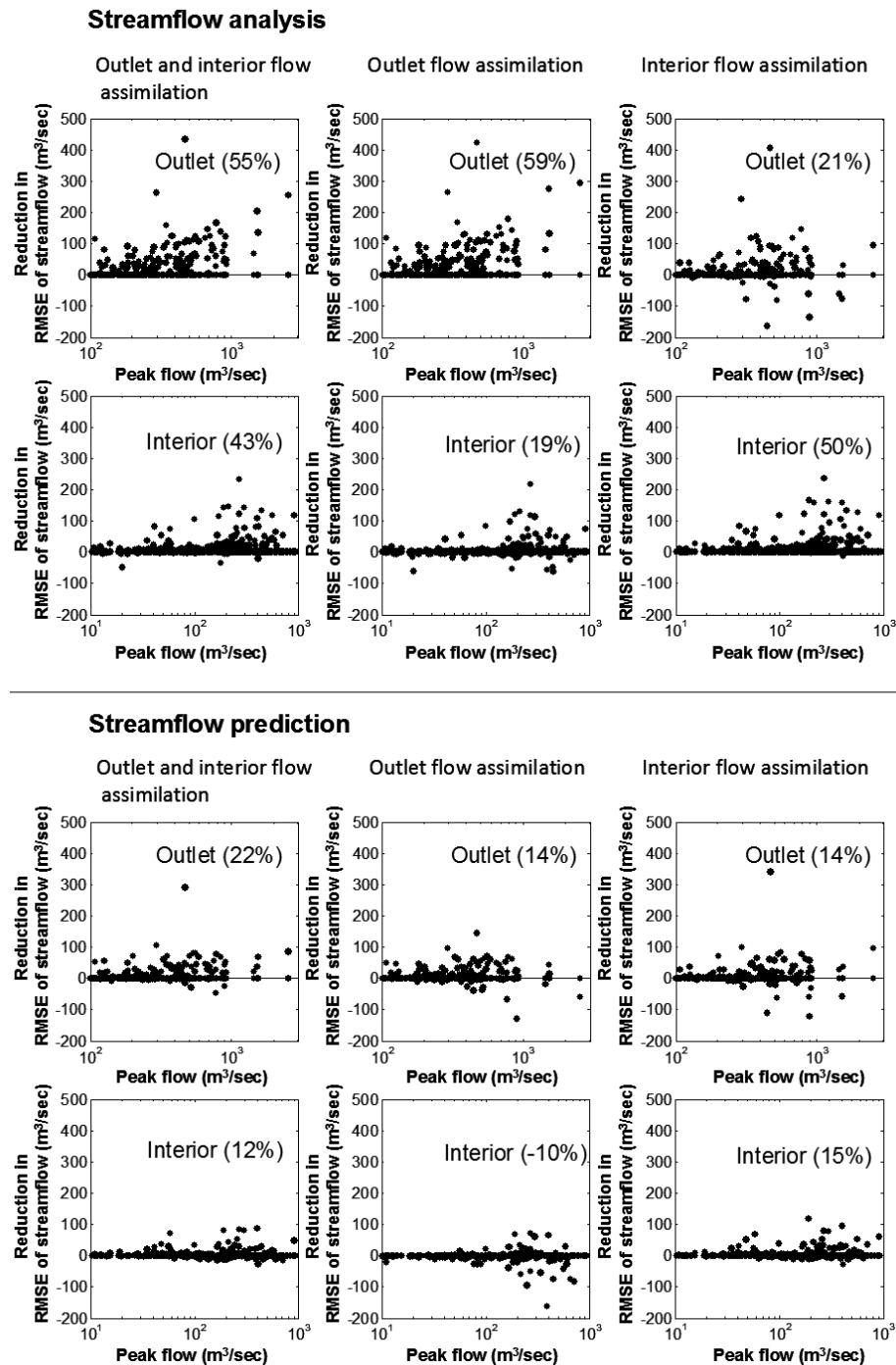


Fig. 11. Reduction in the RMSE of streamflow analysis due to the assimilation vs. peak flows of selected events summarised in Table 2 (top panel); the bottom panel shows the same but for streamflow prediction for 1- to 6-h lead time. The RMSE is calculated for each event and individual lead hour separately. The figure in the parenthesis denotes the percentage reduction in RMSE after the assimilation.

based on r_2 -matrices indicates that the assimilation problem for events with timing errors of 3 h or bigger in the rising limb or peak flow simulation is more ill-posed than the other events, and that the spatial correlation structure of streamflow from the entire simulation appear to be very similar to that of

events with timing errors. To address the above issues, timing errors must be dealt with explicitly, a topic left for future endeavours.

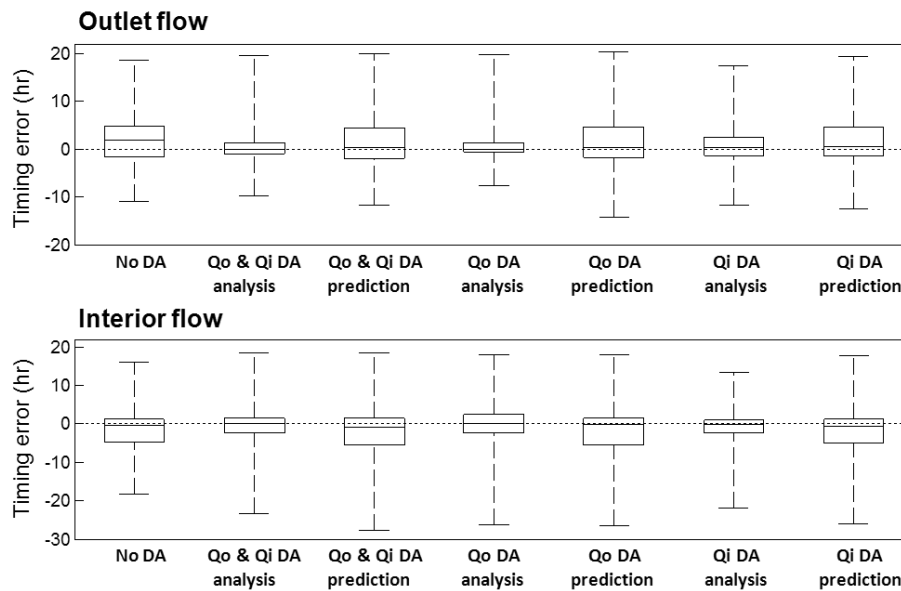


Fig. 12. Timing error estimates in the simulation of outlet and interior flows. The box-plot characterises both inter-basin and event-to-event variability.

5 Conclusion and future work

The importance of hydrologic DA has been emphasised by many researchers as a unifying approach to accounting for different error sources in hydrologic model simulations in a cohesive manner and improving skill in streamflow prediction (Aubert et al., 2003; Liu and Gupta, 2007; Seo et al., 2003, 2009; Clark et al., 2008; Vrugt et al., 2006). Compared to lumped models, distributed rainfall-runoff models are subject to overfitting to a much greater extent due to large dimensionality of the inverse problem involved. In this work, we investigated the effects of the spatiotemporal scale of adjustment in assimilating streamflow data at outlet and/or interior locations into the NWS's Hydrology Laboratory Research Distributed Hydrologic Model (HL-RDHM, Koren et al., 2004). The assimilation technique used is variational assimilation similar to those used in Seo et al. (2009) with lumped models and Lee et al. (2011) with distributed models. For large sample evaluation, we used 4 basins in Oklahoma and 5 basins in Texas in the US.

The main conclusions from this study are as follows:

- The optimum spatiotemporal scale of adjustment varies from basin to basin and between streamflow analysis and prediction. The latter indicates over-adjustment of state variables. The performance of the assimilation procedure is more sensitive to the spatial scale of adjustment than the temporal scale. The preferred strategy identified in this study is to adjust the state variables in a spatially distributed manner and precipitation and PE on an hourly basis, despite the fact that validation with

streamflow at interior and outlet gauge locations at this adjustment scale may indicate overfitting in some cases.

- The quality of streamflow analysis and prediction is highly dependent on the availability of streamflow data at interior locations. At the optimum spatiotemporal adjustment scale, assimilating interior flow and assimilating outlet flow yielded comparable improvement (14 % reduction in RMSE after the assimilation) in outlet flow prediction for the first 6 h of lead time. However, outlet flow assimilation produced degraded interior flow prediction for the first 6-h lead time (10 % increase in RMSE after assimilation), but 15 % reduction in RMSE in the case of assimilating interior flow observations. This indicates that, as one might expect, outlet flow assimilation is more susceptible to overfitting than interior flow assimilation. Assimilating both outlet and interior flows outperforms assimilating outlet flow only for streamflow prediction at the outlet (22 % vs. 14 % reduction in RMSE with assimilation), indicating the importance of additionally assimilating interior flow.
- Basins with highly correlated interior and outlet flows tend to benefit more from streamflow assimilation and be less sensitive to the adjustment scale. Streamflow assimilation at most adjustment scales generally improves the match in the interstation correlation pattern between the observed and the simulated flows. Compared to outlet flow assimilation, interior flow assimilation reproduces better the spatial correlation structure of observed flow. This may be explained by the local information available in interior flow observations, whereas at the

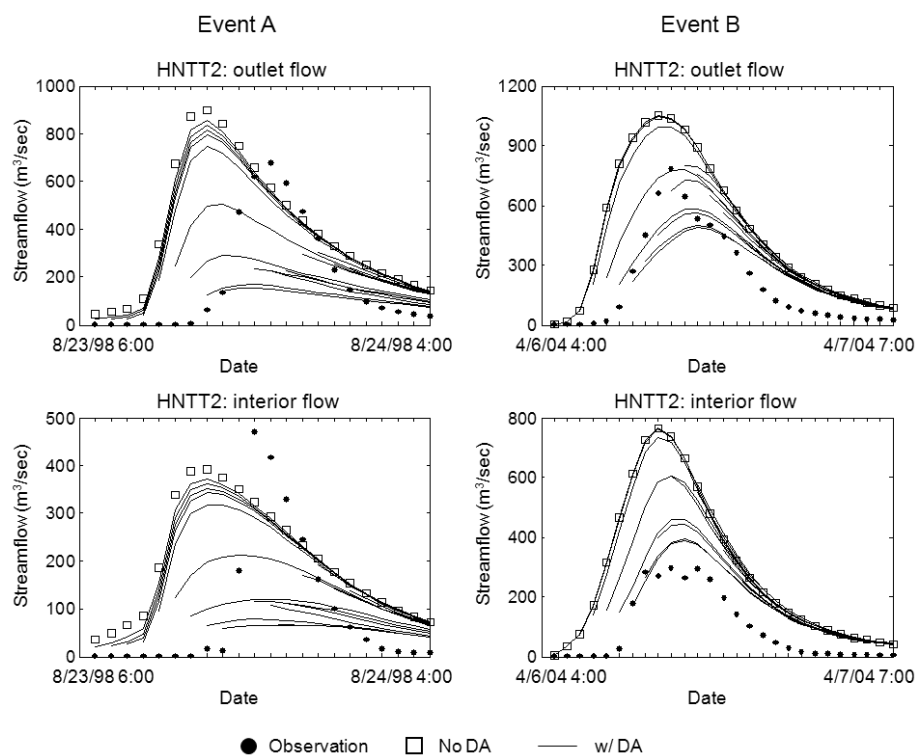


Fig. 13. Streamflow evaluated at the outlet and interior gauge locations for two events in HNTT2. The adjustment scale used is distributed and hourly. The data assimilated is outlet flow. Each curve represents analysis (at the prediction time) and prediction of hourly streamflow generated at different prediction time.

outlet location the information is diluted, or fuzzed up, due to the various intervening hydrologic and hydraulic processes.

- Timing errors in streamflow analysis and prediction are found to be largely related to the ill-posedness of the assimilation problem, which was diagnosed using the information associated with the spatial correlation structure of streamflow. In the cases of events with significant timing errors in rising limb, the assimilation procedure yielded large magnitude errors in streamflow prediction followed by slight improvement in the timing of streamflow analysis. This indicates error compensation with over-adjusting state variables due partly to a lack of timing error modelling component in the objective function used in this study.

The future work should include improving the DA methodology to account for timing errors explicitly, accounting for the model structural error (Van Loon and Troch, 2002; Chen et al., 2011) by applying the model as a weak constraint (Zupanski, 1997), and generalising the procedure in an ensemble framework via, e.g., maximum likelihood ensemble filter (Zupanski, 2005).

Acknowledgements. This work is supported by the Advanced Hydrologic Prediction Service (AHPS) programme of the National Weather Service (NWS). This support is gratefully acknowledged. The first author is grateful to Pedro Restrepo and Minxue He for helpful comments on the manuscript. We are also grateful to Guillaume Thirel and the anonymous reviewer for many very helpful comments.

Edited by: F. Pappenberger

References

- Aubert, D., Loumagne, C., and Oudin, L.: Sequential assimilation of soil moisture and streamflow data in a conceptual rainfall-runoff model, *J. Hydrol.*, 280, 145–161, 2003.
- Brocca, L., Melone, F., Moramarco, T., Wagner, W., Naeimi, V., Bartalis, Z., and Hasenauer, S.: Improving runoff prediction through the assimilation of the ASCAT soil moisture product, *Hydrol. Earth Syst. Sci.*, 14, 1881–1893, doi:10.5194/hess-14-1881-2010, 2010.
- Bulygina, N. and Gupta, H.: Estimating the uncertain mathematical structure of a water balance model via Bayesian data assimilation, *Water Resour. Res.*, 45, W00B13, doi:10.1029/2007WR006749, 2009.

- Burnash, R. J., Ferral, R. L., and McGuire, R. A.: A generalized streamflow simulation system: conceptual modelling for digital computers, US Department of Commerce National Weather Service and State of California, Department of Water Resources, 1973.
- Chen, F., Crow, W. T., Starks, P. J., and Moriasi, D. N.: Improving hydrologic predictions of a catchment model via assimilation of surface soil moisture, *Adv. Water Resour.*, 34, 526–536, doi:10.1016/j.advwatres.2011.01.011, 2011.
- Clark, M. P., 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.
- Doucet, A., De Freitas, N., and Gordon, N.: *Sequential Monte Carlo Methods in Practice*, Springer, New York, USA, 2001.
- Droegemeier, K. K., Smith, J. D., Businger, S., Dosell III, C., Doyle, J., Duffy, C., Foufoula-Georgiou, E., Graziano, T., James, L. D., Krajewski, V., LeMone, M., Lettenmaier, D., Mass, C., Pielke Sr., R., Ray, P., Rutledge, S., Schaake, J., and Zipser, E.: Hydrological aspects of weather prediction and flood warnings: Report of the 9th Prospectus Development Team of the U.S. Weather Research Program, *B. Am. Meteorol. Soc.*, 81, 2665–2680, 2000.
- Fulton, R. A., Breidenbach, J. P., Seo, D.-J., and Miller, D. A.: WSR-88D rainfall algorithm, *Weather Forecast.*, 13, 377–395, 1998.
- Greene, D. R. and Hudlow, M. D.: Hydrometeorologic grid mapping procedures, AWRA International Symposium on Hydrometeor, Denver, CO, 1982.
- Knutson, T. R., McBride, J. L., Chan, J., Emanuel, K., Holland, G., Landsea, C., Held, I., Kossin, J. P., Srivastava, A. K., and Sugi, M.: Tropical cyclones and climate change, *Nat. Geosci.*, 3, 157–163, 2010.
- Koren, V., Smith, M., Wang, D., and Zhang, Z.: Use of soil property data in the derivation of conceptual rainfall-runoff model parameters, *Proceedings of the 15th Conference on Hydrology*, AMS, Long Beach, CA, 103–106, 2000.
- Koren, V., Reed, S., Smith, M., Zhang, Z., and Seo, D. J.: Hydrology Laboratory Research Modelling System (HL-RMS) of the U.S. National Weather Service, *J. Hydrol.*, 291, 297–318, 2004.
- Krzysztofowicz, R.: Bayesian theory of probabilistic forecasting via deterministic hydrologic model, *Water Resour. Res.*, 35, 2739–2750, 1999.
- Lee, H., Seo, D.-J., and Koren, V.: Assimilation of streamflow and in-situ soil moisture data into operational distributed hydrologic models: Effects of uncertainties in the data and initial model soil moisture states, *Adv. Water Resour.*, 34, 1597–1615, 2011.
- Lewis, J. M., Lakshmivarahan, S., and Dhall, S. K.: *Dynamic Data Assimilation: A Least Squares Approach*, Cambridge University Press, 2006.
- Liscum, F.: Effects of urban development on stormwater runoff characteristics for the Houston, Texas, metropolitan area, USGS Water-Resources Investigations Report 01-4071, U.S. Dept. of the Interior, US Geological Survey, Austin, Texas, 2001.
- Liu, Y. and Gupta, H. V.: Uncertainty in hydrologic modelling: Toward an integrated data assimilation framework, *Water Resour. Res.*, 43, W07401, doi:10.1029/2006WR005756, 2007.
- Liu, Y., Brown, J., Demargne, J., and Seo, D.-J.: A wavelet-based approach to assessing timing errors in hydrologic predictions, *J. Hydrol.*, 397, 210–224, 2011.
- Mandapaka, P. V., Krajewski, W. F., Ciach, G. J., Villarini, G., and Smith, J. A.: Estimation of radar-rainfall error spatial correlation, *Adv. Water Resour.*, 32, 1020–1030, 2009.
- McLaughlin, D.: An integrated approach to hydrologic data assimilation: interpolation, smoothing, and filtering, *Adv. Water Resour.*, 25, 1275–1286, 2002.
- Milly, P. C. D., Betancourt, J., Falkenmark, M., Hirsch, R. M., Kundzewicz, Z. W., Lettenmaier, D. P., and Stouffer, R. J.: Stationarity is dead: Whither water management?, *Science*, 319, 573–574, 2008.
- Min, S.-K., Zhang, X., Zwiers, F. W., and Hegerl, G. C.: Human contribution to more-intense precipitation extremes, *Nature*, 470, 378–381, 2011.
- Moradkhani, H.: Hydrologic remote sensing and land surface data assimilation, *Sensors*, 8, 2986–3004, 2008.
- Moradkhani, H., Sorooshian, S., Gupta, H. V., and Hauser, P. R.: Dual state-parameter estimation of hydrological models using ensemble Kalman filter, *Adv. Water Resour.*, 28, 135–147, 2005a.
- Moradkhani, H., Hsu, K., Gupta, H. V., and Sorooshian, S.: Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using particle filter, *Water Resour. Res.*, 41, W05012, doi:10.1029/2004WR003604, 2005b.
- Murphy, A. H.: General decompositions of MSE-based skill scores: Measures of some basic aspects of forecast quality, *Mon. Weather Rev.*, 124, 2353–2369, 1996.
- NHWC – National Hydrologic Warning Council, 2002: Use and benefits of the National Weather Service River and Flood Forecasts, available on line at <http://www.nws.noaa.gov/oh/aAHPS/AHPS%20Benefits.pdf> (last access: 3 January 2012), 2002.
- NRC – National Research Council: When weather matters: Science and service to meet critical societal needs, the National Academies Press, Washington, D.C., 2010.
- NRCS – Natural Resources Conservation Service, United States Department of Agriculture: Soil Survey Geographic (SSURGO) Database, available on line at <http://soildatamart.nrcs.usda.gov> (last access: 19 July 2012), 2004.
- NRCS – Natural Resources Conservation Service, United States Department of Agriculture: US General Soil Map (STATSGO2), available on line at <http://soildatamart.nrcs.usda.gov> (last access: 19 July 2012), 2006.
- Pauwels, V. R. N. and De Lannoy, G. J. M.: Improvement of modeled soil wetness conditions and turbulent fluxes through the assimilation of observed discharge, *J. Hydrometeorol.*, 7, 458–477, 2006.
- Pham, D. T.: Stochastic methods for sequential data assimilation in strongly nonlinear systems, *Mon. Weather Rev.*, 129, 1194–1207, 2001.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P.: *Numerical recipes in fortran*, 2nd Edn., Cambridge University Press, New York, USA, 1992.
- Reed, S. M.: Deriving flow directions for coarse-resolution (1–4 km) gridded hydrologic modelling, *Water Resour. Res.*, 39, 1238, doi:10.1029/2003WR001989, 2003.

- Reed, S. M. and Maidment, D. R.: Coordinate transformations for using NEXRAD data in GIS-based hydrologic modelling, *J. Hydrol. Eng.*, 4, 174–183, 1999.
- Seo, D.-J.: Real-time estimation of rainfall fields using radar rainfall and rain gauge data, *J. Hydrol.*, 208, 37–52, 1998.
- Seo, D.-J., Breidenbach, J. P., and Johnson, E. R.: Real-time estimation of mean field bias in radar rainfall data, *J. Hydrol.*, 223, 131–147, 1999.
- 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., Herr, H. D., and Schaake, J. C.: A statistical post-processor for accounting of hydrologic uncertainty in short-range ensemble streamflow prediction, *Hydrol. Earth Syst. Sci. Discuss.*, 3, 1987–2035, doi:10.5194/hessd-3-1987-2006, 2006.
- 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.
- Seo, D.-J., Demargne, J., Wu, L., Liu, Y., Brown, J. D., Regonda, S., and Lee, H.: Hydrologic ensemble prediction for risk-based water resources management and hazard mitigation, Joint Federal Interagency Conference on Sedimentation and Hydrologic Modelling (JFIC2010), 27 June–1 July 2010, Las Vegas, Nevada, USA, 2010.
- Smith, M. B., Laurine, D. P., Koren, V. I., Reed, S. M., and Zhang, Z.: Hydrologic Model Calibration in the National Weather Service, in: Calibration of Watershed Models, Water Science and Application 6, edited by: Duan, Q., Gupta, H., Sorooshian, S., Rousseau, A., and Turcotte, R., AGU Press, Washington, D.C., 133–152, 2003.
- Smith, M. B., Seo, D.-J., Koren, V. I., Reed, S. M., Zhang, Z., Duan, Q., Moreda, F., and Cong, S.: The distributed model intercomparison project (DMIP): motivation and experiment design, *J. Hydrol.*, 298, 4–26, 2004.
- Torrence, C. and Compo, G. P.: A practical guide to wavelet analysis, *B. Am. Meteorol. Soc.*, 79, 61–78, 1998.
- Trapp, R. J., Diffenbaugh, N. S., Brooks, H. E., Baldwin, M. E., Robinson, E. D., and Pal, J. S.: Changes in severe thunderstorm environment frequency during the 21st century caused by anthropogenically enhanced global radiative forcing, *P. Natl. Acad. Sci. USA*, 104, 19719–19723, 2007.
- Trenberth, K. E., Dai, A., Rasmussen, R. M., and Parsons, D. B.: The changing character of precipitation, *B. Am. Meteorol. Soc.*, 84, 1205–1217, 2003.
- Troch, P. A., Paniconi, C., and McLaughlin, D.: Catchment-scale hydrological modelling and data assimilation, *Adv. Water Resour.*, 26, 131–135, 2003.
- USACE – US Army Corps of Engineers: Annual Flood Damage Report to Congress for Fiscal Year 2000, USACE, Washington, D.C., 2000.
- van Loon, E. E. and Troch, P. A.: Tikhonov regularization as a toll for assimilating soil moisture data in distributed hydrological models, *Hydrol. Process.*, 16, 531–556, 2002.
- Vieux, B. E.: Distributed hydrologic modelling using GIS, Kluwer Academic Publishers, 2001.
- Vrugt, J. A., Diks, C. G. H., Gupta, H. V., Bouten, W., and Verstegen, J. M.: Improved treatment of uncertainty in hydrologic modelling: Combining the strengths of global optimization and data assimilation, *Water Resour. Res.*, 41, W01017, doi:10.1029/2004WR003059, 2005.
- Vrugt, J. A., Gupta, H. V., Nualláin, B., and Bouten, W.: Real-Time Data Assimilation for Operational Ensemble Streamflow Forecasting, *J. Hydrometeorol.*, 7, 548–565, 2006.
- Weerts, A. H. and El Serafy, G. Y. H.: Particle filtering and ensemble Kalman filtering for state updating with hydrological conceptual rainfall-runoff models, *Water Resour. Res.*, 42, W09403, doi:10.1029/2005WR004093, 2006.
- Young, C. B., Bradley, A. A., Krajewski, W. F., and Kruger, A.: Evaluating NEXRAD multisensor precipitation estimates for operational hydrologic forecasting, *J. Hydrometeorol.*, 1, 241–254, 2000.
- Zhang, S., Zou, X., and Ahlquist, J. E.: Examination of numerical results from tangent linear and adjoint of discontinuous nonlinear models, *Mon. Weather Rev.*, 129, 2791–2804, 2001.
- Zupanski, D.: A general weak constraint applicable to operational 4DVAR data assimilation systems, *Mon. Weather Rev.*, 125, 2274–2292, 1997.
- Zupanski, M.: Maximum Likelihood Ensemble Filter: Theoretical aspects, *Mon. Weather Rev.*, 133, 1710–1726, 2005.