IMPROVING FLOOD FORECASTING USING CONDITIONAL BIAS-AWARE ASSIMILATION OF STREAMFLOW OBSERVATIONS AND DYNAMIC ASSESSMENT OF FLOW-DEPENDENT INFORMATION CONTENT

BY

HAOJING SHEN

PRESENTED TO THE FACULTY OF THE GRADUATE SCHOOL OF

THE UNIVERSITY OF TEXAS AT ARLINGTON

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT ARLINGTON

DECEMBER 2020

ACKNOWLEDGEMENTS

I am extremely grateful to my advisor Dr. Dong-Jun Seo for his support during my PhD program. I would also like to thank my dissertation committee members, Drs. Yu Zhang, Michelle Hummel and Haksu Lee for their valuable comments. In addition, I am grateful to Drs. Sunghee Kim and Babak Alizadeh, Messrs. Ali Jozaghi and Vaghef Ghazvinian and all other members and friends at UTA for their support.

This material is based on work supported in part by the National Science Foundation (Grant CyberSEES-1442735) and by the National Oceanic and Atmospheric Administration's Joint Technology Transfer Initiative Program (Grants NA16OAR4590232, NA17OAR4590174 and NA17OAR4590184). These supports are gratefully acknowledged.

Finally, I must express my gratitude to my wonderful parents, family, and friends for all their support and encouragement.

Nov, 2020

ABSTRACT

Accurate forecasting of floods is a long-standing challenge in hydrology and water management. Data assimilation (DA) is a popular technique used to improve forecast accuracy by updating the model states in real time using the uncertainty-quantified actual and model-simulated observations. A particular challenge in DA concerns the ability to improve the prediction of hydrologic extremes, such as floods, which have particularly large impacts on society. Almost all DA methods used today are based on least squares minimization. As such, they are subject to conditional bias (CB) in the presence of observational uncertainties which often leads to under- and over-prediction of the predictand over the upper and lower tails, respectively. To address the adverse impact of CB in DA, conditional bias penalized Kalman filter (CBPKF) and conditional bias penalized ensemble Kalman filter (CBEnKF) have recently been proposed which minimize a weighted sum of the error variance and expectation of the CB squared. Whereas CBPKF and CBEnKF significantly improve the accuracy of the estimates over the tails, they deteriorate performance near the median due to the added penalty. To address the above, this work introduces CB-aware DA, which adaptively weights the CB penalty term in real time, and assesses the flow-dependent information content in observation and model prediction using the degrees of freedom for signal (DFS), which serves as a skill score for information fusion. CB-aware DA is then comparatively evaluated with ensemble Kalman filter in which the marginal information content of observations and its flow dependence are assessed given the hydrologic model used. The findings indicate that CB-aware DA with information content analysis offers an objective framework for improving DA performance for prediction of extremes and dynamically balancing the predictive skill of hydrologic models, quality and frequency of hydrologic observations, and scheduling of DA cycles for improving operational flood forecasting cost-effectively.

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

Chapter 1. General Introduction

Accurate estimation of the dynamically-varying state of the hydrologic system is a requisite for accurate hydrologic prediction and advancing understanding and modeling of the physical processes at work. Toward that end, it is necessary to co-utilize real-time observations and hydrologic model predictions via data assimilation (DA) or model-data fusion (Liu and Gupta, 2007; Liu et al., 2012; Seo et al., 2014). Information fusion for hydrologic state estimation, however, has proven to be a significant challenge.

To achieve optimality in model-data fusion in some well-defined sense, it is necessary to quantify the information content in the observations and model predictions objectively and accurately. Because the rainfall-runoff processes are nonlinear and the primary hydrologic variables of interest such as rainfall, soil moisture, runoff and discharge are generally skewed particularly at high resolutions, the hydrologic state variables tend to be highly heteroscedastic, often encompassing several orders of magnitude. In addition, the information content assessed at one spatiotemporal scale does not transfer to other scales (NRC, 2012). These factors contribute to strong flow dependence in information content in the observations and model predictions, such that the content may vary greatly according to the dynamically-varying state of the hydrologic system. For example, in flood forecasting, one may readily hypothesize that the information content in observed streamflow is larger in conditions of fast-increasing flow than in steady, normal flow, and that the additional information content in observations at new locations, or at the same locations but made more frequently, is also flow-dependent.

State estimation of hydrologic systems at the catchment scale poses additional challenges because most catchments are under-observed. Though in-situ soil moisture sensing has become more popular, they are usually available only from the mesoscale networks such as the Oklahoma Mesonet (Brock et al., 1995; Fiebrich et al., 2006), and the Texas Mesonet (Greene, 2013), and from the generally very sparse national networks such as the Soil Climate Analysis Network (SCAN)(Schaefer et al., 2007). Due to large spatial variability, in-situ soil moisture observations are subject to large representativeness errors which may greatly reduce their information content at the catchment scale (Mohanty et al., 2000; Margulis et al., 2002; Jacobs et al., 2004; Lee et al., 2011; Chen et al., 2014). Although satellite sensing of soil moisture is steadily advancing, the relatively low sampling frequency and generally large uncertainties associated with retrieval significantly limit its utility at the catchment scale (Crow and Ryu, 2009; Brocca et al., 2010; Crow et al., 2011; Chen et al., 2014; Afshar et al., 2019, p. 2; Das et al., 2019; Mao et al., 2019). Though spatially sparse, streamflow observations reflect the spatiotemporally-integrated state of the catchment, and hence are high in information content at that scale. Given the above observations, one recognizes that effective assimilation of streamflow observations is critical to accurate estimation of the hydrologic state at the catchment scale and skillful prediction of streamflow, which is arguably the most important variable in hydrologic prediction (Clark et al., 2008; Noh et al., 2018), and other hydrologic variables.

An important consideration in hydrologic information fusion is that society places a much higher premium on accurate prediction of extremes, such as floods, droughts, and algal blooms, than that of "normal" events. As such, assessment and utilization of flow-dependent information content are particularly important to accurate prediction of large-to-extreme events. Also important, but not very widely recognized in the hydrologic literature, are the compounding observational uncertainties associated with assimilating observations of "non-normal" states. For example, streamflow observations are subject to errors due to imperfect measurement of stage and uncertain estimation of flow from stage observations. In addition, statistical modeling using observations and model predictions such as DA and information fusion is subject to conditional bias (CB) when the system is in non-normal states, as in very high or low flow conditions.

It is well known in statistics and econometrics that, in the presence of significant observation error, least-squares minimization tends to introduce negative and positive biases, or CB, over the upper and lower tails of the predictands, respectively (Fuller, W.A., 1987; Seber and Wild, 1989; Frost and Thompson, 2000). Referred to as the "iron law of econometrics" for its ubiquity in statistical modeling and prediction (Hausman, 2001), the CB is more acute when large observational uncertainties exist or the predictability is limited, a condition frequently encountered in hydrologic prediction particularly of large-to-extreme events for under-observed systems.

There are two types of CB (Jolliffe and Stephenson, 2003; Wilks, 2011). The Type-I CB, defined as $E\left[X\middle|\hat{X}=\hat{x}\right]-\hat{x}$, where X, $\hat{X}$, and $\hat{x}$ denote the unknown truth, the estimate, and the realization of $\hat{X}$, respectively, is associated with false alarm. The Type-II CB, defined as $E\left[\hat{X}\middle|X=x\right]-x$, where x denotes the realization of X, is associated with failure to detect an event. Of the two, the Type-II is of particular interest for hydrologic prediction in that, whereas the Type-I CB may be reduced by calibration (not desirable but not impossible if large enough data exist), the Type-II cannot (Wilks, 2011; Seo et al., 2018a, 2018b). Because the Type-II CB arises over the tails of the distribution, its reduction is particularly important for estimation and prediction of large-to-extreme events. Indeed, Lee et al. (2019) have shown that, in flood forecasting, the larger the event is, the larger the margin of improvement in predictive skill is from addressing the Type-II CB. Figure 1-1 shows that CB-penalized ensemble Kalman filter (CBEnKF)(Lee et al., 2019), which co-minimizes the Type-II CB and error variance, significantly improves over EnKF (Evensen, 1994, 2003) under identical conditions, and that, the larger the verifying flow is, the larger the margin of improvement is.

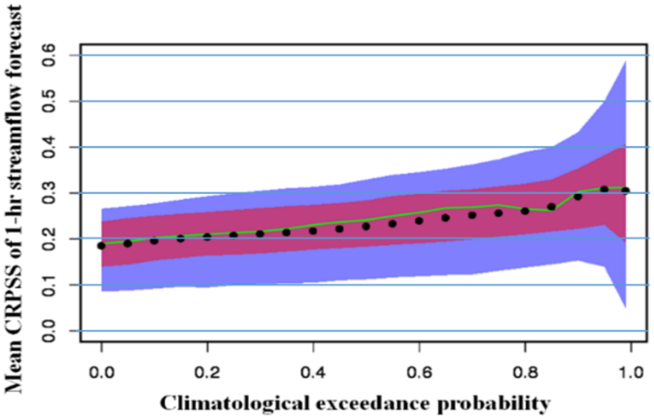


Figure 1-1. Mean Continuous Ranked Probability Skill Score (CRPSS) of the CBEnKF 1-hr streamflow forecast over the EnKF for 10 basins in TX as a function of the conditioning verifying observed flow exceeding the climatological probability shown on the x-axis. The probability levels used range from 0 to 0.99. At each probability level, there are 90 values (10 basins, 9 cases) from which the 5, 25, 50, 75, 95th percentiles and the mean value were calculated. Blue and red shaded areas represent the 5-95th and 25-75th percentile ranges, respectively. Green line and black dot denote the median and mean, respectively (from Lee et al. 2019).

One of the most important societal functions of hydrologic prediction is to minimize or reduce surprises, i.e., the Type-II errors. With urbanization and climate change, calibrating hydrologic models, including uncertainty models, based on long periods of records is becoming an increasingly tenuous practice (Wi et al., 2015; Alizadeh et al., 2020). In parts of TX, there are also growing evidence that the uncertainty in the initial conditions (IC) of operational hydrologic models may be increasing (Cotter, 2015). The record-breaking flooding of the Blanco River in central TX in May 2015 (Furl et al., 2018), which was attributed to rapid urbanization whose potential impact was assessed and predicted as early as in 2007 (Curran, 2007), is one of many recent examples that encapsulate the challenges in hydrologic state estimation and prediction in changing conditions.

(Jervis 2015 USA TODAY) On May 23, 2015, the Blanco River in Central Texas rose more than 33 feet in 3 hours, cresting at about 40 ft, or 27 ft above flood stage. It broke the record crest in 1929 by about 7 ft before the gauge stopped reporting (see Figure 1-2). “Basically, no warning,” said Jim Harris, a Blanco resident, whose three-bedroom home filled with 5 ft of water. Local officials said they were caught off guard by the floods, due in part to a lack of urgency from federal forecasters who were predicting much lower flood levels than what actually occurred in town. Hydrologists at the West Gulf River Forecast Center (WGRFC) forecasted at 7:15 p.m. that the Blanco River at Wimberley would crest at around 12 ft. Forty-five minutes later, they issued another forecast saying Wimberley could see "minor flooding," with the river cresting at 16 ft. The center only called for "major flooding" at 11:19 pm, hours into the flood and as residents in both Blanco and Wimberley were being rescued from rooftops.”

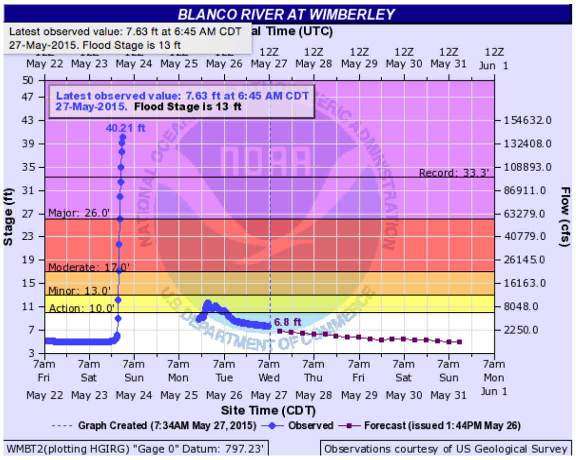


Figure 1-2. Blanco River at Wemberly, TX, on May 22 through 27, 2015 (from

The predictive skill for events such as flash floods is predicated on the skillful prediction of rainfall, a particularly large challenge for convective events (Clark et al., 2016; Herman and Schumacher, 2016). To gain as much hydrologic lead time as possible toward the “golden time”, i.e., the minimum time necessary to secure safety from flash flooding or flooding (Moon et al. 2017), it is critically important to recognize and discern significant departures of the model predictions from the unfolding reality as quickly as possible, and to update the transient state of the hydrologic models objectively and accurately for skillful prediction. It is argued that, to utilize fully the flow-dependent information content in the observations and model predictions, it is necessary to address the compounding observational uncertainties, and that advancing understanding and utilization of flow-dependent information is essential to identifying and addressing the science and technology gaps in the observation-prediction chain toward step-change improvement in water forecasting in a changing world.

The overarching hypothesis of this research is that there exists significant flow-dependent information content in hydrologic observations that is currently not utilized fully in estimation and prediction of large-to-extreme events. It is also hypothesized that, by unraveling the compounding observational uncertainties, it is possible to realize significant improvement in predictive skill and to reduce predictive uncertainty. To test these hypotheses, this research proposes to develop and comparatively evaluate CB-aware DA, and objectively assess the flow-dependent information content in observations and model predictions under a wide range of hydrologic conditions. The degrees of freedom for signal (DFS)(Rodgers, 2000) is used as the primary measure of information content, and CBEnKF (Lee et al., 2019; Shen et al., 2019) is used as the primary building block for CB-aware DA. Simulation experiments for comprehensive comparative evaluation with EnKF using multiple hydrologic models are designed and carried out. The observations of the prognostic variable of streamflow, and the forcing variables of precipitation and potential evapotranspiration (PE) are assimilated. A diverse set of basins in TX, including those in semi-arid regions, is used for hypothesis testing, comparative evaluation of CB-aware DA, and assessment of flow-dependent information content.

The new and significant contribution of this research are:

1) Advances in understanding of CB in optimal estimation and hydrologic prediction,
2) Development of variance-inflated KF (VIKF), algorithmically simpler and computationally less expensive approximation for CB-penalized KF (CBPKF),
3) Comparative evaluation of CBEnKF for flood forecasting,
4) Development of adaptive CBPKF and CBEnKF,

5) Comparative evaluation of CB-aware DA based on adaptive CBEnKF for flood forecasting,
6) Advances in understanding of flow-dependent information content in hydrologic observations in catchment's streamflow response,
7) Assessment of value of DA for event-based flash flood forecasting using one way-nested high-resolution hydrologic model.
8) CB-aware DA is suitable add-on to any existing streamflow prediction systems to improve forecast skill of the extreme events, hence protect the flood prone areas, such as Texas.

This dissertation is organized as follows. In Chapter 2, CBPKF, CB-penalized fisher-like solution, VIKF and adaptive CBPKF are described. Chapter 3 describes CBEnKF and its evaluation. In Chapter 4, CB-aware DA is described and evaluated, including adaptive CBEnKF and assessment of information content using DFS. Chapter 5 describes the application of CB-aware DA in WRF-Hydro. Lastly, Chapter 6 provides the general conclusions and future research recommendations.

Chapter 2. CBPKF: Theory, Adaptive Extension and Approximation

2.1 Introduction

Kalman filter (KF) and its variants and extensions are widely used to fuse observations with model predictions in a wide range of applications (Reichle et al., 2002; Yu et al., 2005; Dong and You, 2006; Antoniou et al., 2007; de Wit and van Diepen, 2007; Kurtz et al., 2012; Wallace et al., 2012; Bhotto and Bajić, 2015; Muñoz-Sabater, 2015; Gao et al., 2016, p. 20; Houtekamer and Zhang, 2016; Jain and Krishnamurthy, 2016; Jiang et al., 2016; Lu et al., 2016; Lv et al., 2016; Nair and Indu, 2016; Yan et al., 2016; Chen et al., 2017; Ma et al., 2017; Zhou et al., 2017; Bocher et al., 2018). In geophysics and environmental science and engineering, often the main objective of information fusion is to improve estimation and prediction of states in their extremes rather than in normal ranges. In hydrologic forecasting, for example, accurate prediction of floods and droughts is far more important than that of streamflow and soil moisture in normal conditions. Because KF minimizes unconditional error variance, its solution tends to improve estimation near median where the state of the dynamic system resides most of the times while often leaving significant biases in the extremes. Such conditional biases (CB) (Ciach et al., 2000) generally result in consistent under- and overestimation of the true states in the upper and lower tails of the distribution, respectively. To address CB, CB-penalized Fisher-like estimation and CB-penalized KF (CBPKF) (Seo et al., 2018a, 2018b) have recently been developed which jointly minimize error variance and expectation of the Type-II CB squared for improved estimation and prediction of extremes. The Type-II CB, defined as $E\left[\hat{X}|X=x\right]-x$, is associated with failure to detect the event where x denotes the realization of X where X, $\hat{X}$ and $\hat{x}$ denote the unknown truth, the estimate, and the realization of $\hat{X}$, respectively (Jolliffe and Stephenson, 2003). The original formulation of CBPKF, however, is computationally extremely expensive for high-dimensional problems.

Also, whereas CBPKF improves performance in the tails, it deteriorates performance in the normal ranges. In this work, we approximate CBPKF with forecast error covariance-inflated KF, referred to hereafter as the variance-inflated KF (VIKF) formulation, as a computationally less expensive and algorithmically simpler alternative, and implement adaptive CBPKF to improve unconditional performance.

Elements of CB-penalized Fisher-like estimation has been described in the forms of CB-penalized indicator cokriging for fusion of predicted streamflow from multiple models and observed streamflow (Brown et al., 2010), CB-penalized kriging for spatial estimation (Seo, 2012) and rainfall estimation (Seo et al., 2014), and CB-penalized cokriging for fusion of radar rainfall and rain gauge data (Kim et al., 2018). The original formulation of CBPKF have been described in (Seo et al., 2018a) and (Seo et al., 2018b), respectively. Its ensemble extension, CB-penalized ensemble KF, or CEnKF, is described in (Lee et al., 2018) in the context of ensemble data assimilation for flood forecasting. Whereas CBPKF was initially motivated for environmental and geophysical state estimation and prediction, it is broadly applicable to a wide range of applications for which improved performance in the extremes is desired.

This chapter is an adaptation of Shen et al. (2019) and is organized as follows. Sections 2.2 and 2.3 describe CB-penalized Fisher-like solution and CBPKF, respectively. Section 2.4 describes approximation of CBPKF. Sections 2.5 describe the evaluation experiments and results, respectively. Section 2.6 describes adaptive CBPKF. Section 2.7 provides conclusions.

2.2 Conditional Bias-Penalized Fisher-Like Solution

As in Fisher estimation (Schweppe, 1973), the estimator sought for CB-penalized Fisher-like estimation is $X^* = \boldsymbol{WZ}$ where X^* denotes the $(m \times 1)$ vector of the estimated states, $\boldsymbol{W}$ denotes the $\big(m \times (n+m)\big)$ weight matrix, and $\boldsymbol{Z}$ denotes the $\big((n+m) \times 1\big)$ augmented observation vector. In the above, n denotes the number of observations, m denotes the number of state variables, and $(n+m)$ reflects the dimensionality of the augmented vector of the observations and the model-predicted states to be fused for estimation of the true state X. the purpose of augmentation is to relate directly to CBPKF in Section 2.3. Throughout this chapter, we use regular and bold letters to differentiate the non-augmented and augmented variables, respectively. The linear observation equation is given by:

$$\boldsymbol{Z} = \boldsymbol{H}X + \boldsymbol{V} \tag{2.1}$$

where X denotes the $(m \times 1)$ vector of the true state with $E[X] = Mx$ and $Cov[X, X^T] = \Psi_{\mathrm{XX}}$, $\mathbf{H}$ denotes the $\big((n+m) \times m\big)$ augmented linear observation equation matrix, and $\boldsymbol{V}$ denotes the $\big((n+m) \times 1\big)$ augmented zero-mean observation error vector with $Cov[V, V^T] = \boldsymbol{R}$. Assuming independence between X and $\boldsymbol{V}$, we write the Bayesian estimator for X, or X^*, as (Schweppe 1973):

$$X^* = M_X + \boldsymbol{W}(\boldsymbol{Z} - \boldsymbol{M_Z}) \tag{2.2}$$

The error covariance matrix for X^*, $E[(X - X^*)(X - X^*)^T]$, is given by:

$$\Sigma_{EV} = (\boldsymbol{I} - \boldsymbol{WH})\Psi_{XX}(\boldsymbol{I} - \boldsymbol{WH})^{\boldsymbol{T}} + \boldsymbol{WRW^T} \tag{2.3}$$

With Eq.(2.2), we may write Type-II CB as:

$$X - E[X^*|X] = (X - M_X) - \boldsymbol{W}E[(\boldsymbol{Z} - \boldsymbol{H}M_x)|X] \tag{2.4}$$

Inverting Eq.(2.1), we have for the observation equation for $\mathbf{Z}$:

$$\boldsymbol{X} = \boldsymbol{G}^T\boldsymbol{Z} - \boldsymbol{G}^T\boldsymbol{V} \qquad (2.5)$$

The $(m \times n)$ matrix, $\boldsymbol{G}^T$, in Eq.(2.5) is given by**:**

$$\boldsymbol{G^T} = (\boldsymbol{U^T H})^{-\mathbf{1}}\boldsymbol{U^T} \qquad (2.6)$$

where $\boldsymbol{U^T}$ is some (m×(n+m)) nonzero matrix. Using Eq.(2.5) and the identity, $\boldsymbol{\Psi_{ZZ}} = \boldsymbol{H}\Psi_{XX}\boldsymbol{H^T} + \boldsymbol{R}$, we may write the Bayesian estimate for E[**Z**|X] in Eq.(2.4) as:

$$\hat{E}[\boldsymbol{Z}|X] = \boldsymbol{H}M_X + \boldsymbol{C}(X - M_X) \qquad (2.7)$$

where

$$\boldsymbol{C} = (\boldsymbol{H}\Psi_{XX}\boldsymbol{H^T} + \boldsymbol{R})\boldsymbol{G^T}[\boldsymbol{G}(\boldsymbol{H}\Psi_{XX}\boldsymbol{H^T} + \mathbf{2}\boldsymbol{R})\boldsymbol{G^T}]^{-\mathbf{1}} \qquad (2.8)$$

Eqs.(2.7) and (2.8) state that, if the a priori state error covariance Ψ_{XX} is noninformative or there are no observation errors, the Bayesian estimate of **Z** given X is given by **H**X, but that, if the a priori Ψ_{XX} is perfectly informative or observations are information-less, the Bayesian estimate is given by the average of the a priori mean M_X and the observed true state X.

With Eq.(2.4), we may write the quadratic penalty due to Type-II CB as:

$$\Sigma_{CB} = E[(X - E_{X^*}[X^*|X])(X - E_{X^*}[X^*|X])^T\,] = (I - \boldsymbol{WC})\Psi_{XX}(I - \boldsymbol{WC})^T \qquad (2.9)$$

where I denotes the (m×m) identity matrix. Combining Σ_{EV} in Eq.(2.3) and Σ_{CB} in Eq.(2.9), we have the apparent error covariance, Σ_a, which reflects both the error covariance and Type-II CB:

$$\Sigma_a = (I - \boldsymbol{WH})\boldsymbol{\Psi_{XX}}(I - \boldsymbol{WH})^{\boldsymbol{T}} + \boldsymbol{WRW^T} + \alpha(I - \boldsymbol{WC})\Psi_{XX}(I - \boldsymbol{WC})^{\boldsymbol{T}} \qquad (2.10)$$

where α denotes the scaler weight given to the CB penalty term. Minimizing Eq.(2.10) with respect to **W**, or by direct analogy with the Bayesian solution (Schweppe 1973), we have:

$$\mathbf{W} = \Psi_{XX}\widehat{\boldsymbol{H}}^{\boldsymbol{T}}\left[\widehat{\boldsymbol{H}}\Psi_{XX}\widehat{\boldsymbol{H}}^{\boldsymbol{T}} + \boldsymbol{\Lambda}\right]^{-\mathbf{1}} \tag{2.11}$$

The modified structure matrix $\widehat{\boldsymbol{H}}^{\boldsymbol{T}}$ and observation error covariance matrix $\boldsymbol{\Lambda}$ in Eq.(2.11) are given by:

$$\widehat{\boldsymbol{H}}^{\boldsymbol{T}} = \boldsymbol{H}^{\boldsymbol{T}} + \alpha\boldsymbol{C}^{\boldsymbol{T}} \tag{2.12}$$

$$\boldsymbol{\Lambda} = \boldsymbol{R} + \alpha(1-\alpha)\boldsymbol{C}\Psi_{XX}\boldsymbol{C}^{\boldsymbol{T}} - \alpha\boldsymbol{H}\Psi_{XX}\boldsymbol{C}^{\boldsymbol{T}} - \alpha\boldsymbol{C}\Psi_{XX}\boldsymbol{H}^{\boldsymbol{T}} \tag{2.13}$$

Using Eq.(2.11) and the matrix inversion lemma (Woodbury 1950), we may write Σ_a and X^* in Eqs.(2.10) and (2.2), respectively, as:

$$\Sigma_a = \alpha\Psi_{XX} + \left[\widehat{\boldsymbol{H}}\boldsymbol{\Lambda}^{-\mathbf{1}}\widehat{\boldsymbol{H}}^{\boldsymbol{T}} + \Psi_{XX}^{-1}\right]^{-\mathbf{1}} \tag{2.14}$$

$$X^* = \left[\widehat{\boldsymbol{H}}^{\boldsymbol{T}}\boldsymbol{\Lambda}^{-\mathbf{1}}\widehat{\boldsymbol{H}} + \Psi_{XX}^{-1}\right]^{-\mathbf{1}}\left\{\widehat{\boldsymbol{H}}^{\boldsymbol{T}}\boldsymbol{\Lambda}^{-\mathbf{1}}\boldsymbol{Z} + \Psi_{XX}^{-1}M_X\right\} + \Delta \tag{2.15}$$

where $\Delta = \alpha\Psi_{XX}\widehat{\boldsymbol{H}}^{\boldsymbol{T}}[\widehat{\boldsymbol{H}}\Psi_{XX}\widehat{\boldsymbol{H}}^{\boldsymbol{T}} + \boldsymbol{\Lambda}]^{-\mathbf{1}}\boldsymbol{C}M_X$. To render the above Bayesian solution to a Fisher-like solution, we assume no a priori information in X by letting Ψ_{XX}^{-1}, which is associated with error covariance in Eq.(2.3), vanish in the brackets in Eqs.(2.14) and (2.15) to arrive at:

$$\Sigma_a = B\left[\widehat{\boldsymbol{H}}\boldsymbol{\Lambda}^{-\mathbf{1}}\widehat{\boldsymbol{H}}^{\boldsymbol{T}}\right]^{-\mathbf{1}} \tag{2.16}$$

$$X^* = \left[\widehat{\boldsymbol{H}}^{\boldsymbol{T}}\boldsymbol{\Lambda}^{-\mathbf{1}}\widehat{\boldsymbol{H}}\right]^{-\mathbf{1}}\widehat{\boldsymbol{H}}^{\boldsymbol{T}}\boldsymbol{\Lambda}^{-\mathbf{1}}\boldsymbol{Z} + \Delta \tag{2.17}$$

where the scaling matrix B is given by $B = \alpha\Psi_{XX}\widehat{\boldsymbol{H}}^{\boldsymbol{T}}\boldsymbol{\Lambda}^{-\mathbf{1}}\widehat{\boldsymbol{H}} + I$. To obtain the estimator of the form, $X^* = \boldsymbol{W}\boldsymbol{Z}$, we impose the unbiasedness condition, $E[X^*] = X$, or equivalently, $\boldsymbol{W}\boldsymbol{H} = I$. The unbiasedness condition is satisfied by replacing $[\widehat{\boldsymbol{H}}^{\boldsymbol{T}}\boldsymbol{\Lambda}^{-\mathbf{1}}\widehat{\boldsymbol{H}}]^{-\mathbf{1}}$ with $[\widehat{\boldsymbol{H}}^{\boldsymbol{T}}\boldsymbol{\Lambda}^{-\mathbf{1}}\boldsymbol{H}]^{-\mathbf{1}}$ and dropping Δ in Eq.(2.17), which yields:

$$\Sigma_a = \boldsymbol{B}\left[\widehat{\boldsymbol{H}}\boldsymbol{\Lambda}^{-\mathbf{1}}\boldsymbol{H}^{\boldsymbol{T}}\right]^{-\mathbf{1}} \tag{2.18}$$

$$X^* = \left[\boldsymbol{\hat{H}^T \Lambda^{-1} H}\right]^{-1} \boldsymbol{\hat{H}^T \Lambda^{-1} Z} \tag{2.19}$$

Finally, we obtain from Eq.(2.3) the error covariance, Σ_{EV}, associated with X^* in Eq.(2.19):

$$\Sigma_{EV} = \boldsymbol{WRW^T} = \left[\boldsymbol{\hat{H}^T \Lambda^{-1} H}\right]^{-1} \boldsymbol{\hat{H}^T \Lambda^{-1} R \Lambda^{-1} \hat{H}} \left[\boldsymbol{\hat{H}^T \Lambda^{-1} H}\right]^{-1} \tag{2.20}$$

It is readily seen that, if α=0, we have $\boldsymbol{\hat{H}^T} = \boldsymbol{H}$ and $\boldsymbol{\Lambda} = \boldsymbol{R}$, and hence the CB-penalized Fisher-like estimator, Eqs.(2.19) and (2.20), is reduced to the Fisher estimator (Schweppe, 1973).

2.3 Conditional Bias-Penalized Kalman Filter

CBPKF results directly from decomposing the augmented matrices and vectors in Eq.(2.19) and Eq.(2.20) as KF does from the Fisher solution (Schweppe, 1973). The CBPKF solution, however, is not very simple because the modified observation error covariance matrix, **Λ**, is no longer diagonal. An important consideration in casting the CB-penalized Fisher-like solution into CBPKF is to recognize that CB arises from the error-in-variable effects associated with uncertain observations (Hausman, 2001), and that the a priori state, represented by the dynamical model forecast, is not subject to CB. We therefore apply the CB penalty to the observations only, and reduce **C** in (8) to $\boldsymbol{C^T} = \left(C_{1,k}^T C_{2,k}^T\right) = \left(C_{1,k}^T\ 0\right)$. Separating the observation and dynamical model components in $\boldsymbol{\hat{H}^T}$ and **Λ** via the matrix inversion lemma, we have**:**

$$\boldsymbol{\hat{H}^T} = \left(\hat{H}_{1,k}^T\ I\right) \tag{2.21}$$

$$\boldsymbol{\Lambda} = \begin{bmatrix} \Lambda_{11,k} & \Lambda_{12,k} \\ \Lambda_{21,k} & \Lambda_{22,k} \end{bmatrix} \tag{2.22}$$

where

$$\hat{H}_{1,k}^T = H_k^T + \alpha C_{1,k}^T \tag{2.23}$$

$$\Lambda_{11,k} = R_k + \alpha(1-\alpha)C_{1,k}\Psi_{XX}C_{1,k}^T - \alpha H_k\Psi_{XX}C_{1,k}^T - \alpha C_{1,k}\Psi_{XX}H_k^T \tag{2.24}$$

$$\Lambda_{12,k} = -\alpha C_{1,k}\Psi_{XX} \tag{2.25}$$

$$\Lambda_{21,k} = \Lambda_{12,k}^T \tag{2.26}$$

$$\Lambda_{22,k} = \Sigma_{k|k-1} \tag{2.27}$$

In the above, H_k denotes the (n × m) observation matrix, and R_k denotes the (n×n) observation error covariance matrix. To evaluate the (m × n) matrix, $C_{1,k}$, it is necessary to specify $\boldsymbol{U^T}$ in (2.6). We use $\mathbf{U^T}$=$\mathbf{H^T}$ which ensures invertibility of $\mathbf{U^T H}$, but other choices are also possible. We then have for $C_{1,k}^T$:

$$C_{1,k} = \left[(H_k\Psi_{XX}H_k^T + R_k)G_{1,k} + H_k\Psi_{XX}G_{2,k}\right]L_k^{-1} \tag{2.28}$$

where

$$G_{2,k}^T = (H_k^T H_k + I)^{-1} \tag{2.29}$$

$$G_{1,k}^T = G_{2,k}^T H_k^T \tag{2.30}$$

$$L_k = G_{2,k}^T\left[H_k^T(H_k\Psi_{XX}H_k^T + 2R_k)H_k + H_k^T H_k\Psi_{XX} + \Psi_{XX}H_k^T H_k + \Psi_{XX} + 2\Sigma_{k|k-1}\right]G_{2,k} \tag{2.31}$$

Expanding $\mathbf{W}$ in (2.11) with $\boldsymbol{\Lambda}^{-1} = \boldsymbol{\Gamma} = \begin{bmatrix}\Gamma_{11,k} & \Gamma_{12,k}\\ \Gamma_{21,k} & \Gamma_{22,k}\end{bmatrix}$, we have;

$$\boldsymbol{W} = \left[\boldsymbol{\widehat{H}^T\Lambda^{-1}H}\right]^{-1}\boldsymbol{\widehat{H}^T\Lambda^{-1}} = \left(\varpi_{1,k}H_k + \varpi_{2,k}\right)^{-1}\left(\varpi_{1,k}\varpi_{2,k}\right) \tag{2.32}$$

In (2.32), the (m×n) and (m×m) weight matrices for the observation and model prediction, $\omega_{1,k}$ and $\omega_{2,k}$, respectively, are given by:

$$\varpi_{1,k} = \widehat{H}_{1,k}^T\Gamma_{11,k} + \Gamma_{21,k} \tag{2.33}$$

$$\varpi_{2,k} = \widehat{H}_{1,k}^{T}\Gamma_{12,k} + \Gamma_{22,k} \tag{2.34}$$

where

$$\Gamma_{22,k} = \left[\Lambda_{22,k} - \Lambda_{21,k}\Lambda_{11,k}^{-1}\Lambda_{12,k}\right]^{-1} \tag{2.35}$$

$$\Gamma_{11,k} = \Lambda_{11,k}^{-1} + \Lambda_{11,k}^{-1}\Lambda_{12,k}\Gamma_{22,k}\Lambda_{21,k}\Lambda_{11,k}^{-1} \tag{2.36}$$

$$\Gamma_{12,k} = -\Lambda_{11,k}^{-1}\Lambda_{12,k}\Gamma_{22,k} \tag{2.37}$$

The apparent CBPKF error covariance, which reflects both Σ_{EV} and Σ_{CB}, is given by (2.18) as:

$$\Sigma_{a,k|k} = \alpha\Sigma_{k|k-1} + \left[\varpi_{1,k}H_k + \varpi_{2,k}\right]^{-1} \tag{2.38}$$

The CBPKF error covariance, which reflects Σ_{EV} only, is given by (2.20) as:

$$\Sigma_{k|k} = \left[\varpi_{1,k}H_k + \varpi_{2,k}\right]^{-1}\left(\varpi_{1,k}R_k\varpi_{1,k}^{T} + \varpi_{2,k}\Sigma_{k|k-1}\varpi_{2,k}^{T}\right)\left[\varpi_{1,k}H_k + \varpi_{2,k}\right]^{-1} \tag{2.39}$$

Because CBPKF minimizes $\Sigma_{a,k|k}$ rather than $\Sigma_{k|k}$, it is not guaranteed that (2.39) satisfies $\Sigma_{k|k} \leq \Sigma_{k|k-1}$ a priori. If the above condition is not met, it is necessary to reduce α and repeat the calculations. If α is reduced all the way to zero, CBPKF collapses to KF. The CBPKF estimate may be rewritten into a more familiar form:

$$\widehat{X}_{k|k} = \left[\varpi_{1,k}H_k + \varpi_{2,k}\right]^{-1}\left[\varpi_{1,k}Z_k + \varpi_{2,k}\widehat{X}_{k|k-1}\right] = \widehat{\mathrm{X}}_{k|k-1} + K_k\left[Z_k - H_k\widehat{\mathrm{X}}_{k|k-1}\right] \tag{2.40}$$

In (2.40), Z_k denotes the (n×1) observation vector, and the (m×n) CB-penalized Kalman gain, K_k, is given by:

$$K_k = \left[\varpi_{1,k}H_k + \varpi_{2,k}\right]^{-1}\varpi_{1,k} \tag{2.41}$$

To operate the above as a sequential filter, it is necessary to prescribe Ψ_{XX} and α. An obvious choice for Ψ_{XX}, i.e., the a priori error covariance of the state, is $\Sigma_{k|k-1}$. Specifying α requires some care. In general, a larger α improves accuracy over the tails but at the expense of increasing unconditional error. Too small an α may not effect large enough CB penalty in which case the CBPKF and KF solutions would differ little. Too large an α, on the other hand, may severely violate the $\Sigma_{k|k} \leq \Sigma_{k|k-1}$ conditionin which case the filter may have to be iterated at additional computational expense with successively reduced α. A reasonable strategy for reducing α is $\alpha_i = c\alpha_{i-1}, i = 1,2,3, \ldots$, with $0 < c < 1$ where α_i denotes the value of α at the i-th iteration (Lee et al., 2018, p. 2; Seo et al., 2018a). For high-dimensional problems, CBPKF can be computationally very expensive. Whereas KF requires solving an (m×n) linear system only once per updating or fusion cycle, CBPKF additionally requires solving two (m×m) linear systems (for $C_{1,k}$ and Γ_{22}), and an (n×n) system (for Λ_{11}), assuming that the structure of the observation equation does not change in time (in which case $G_{2,k}^T$ in (2.29) may be evaluated only once). To reduce computation, below we approximate CBPKF with KF by inflating the forecast error covariance.

2.4 VIKF Approximation of CBPKF

To address CB, CB-penalized Fisher-like estimation and CB-penalized KF (CBPKF) (Seo et al., 2018a, 2018b) have recently been developed which jointly minimize error variance and expectation of the Type-II CB squared for improved estimation and prediction of extremes. The original formulation of CBPKF, however, is computationally very expensive for high-dimensional problems. Also, whereas CBPKF improves performance in the tails, it deteriorates performance in the normal ranges. In this chapter, we approximate CBPKF with forecast error covariance-inflated KF, referred to hereafter as the variance-inflated KF (VIKF) formulation, as a computationally less expensive and algorithmically simpler alternative, and implement adaptive CBPKF to improve performance in the unconditional mean sense.

The main idea behind this simplification is that, if the gain for the CB penalty, **C**, in Eq.(2.10) can be linearly approximated with **H**, the apparent error covariance Σ_a becomes identical to Σ_{EV} in Eq.(2.3) but with Ψ_{XX} inflated by a factor of 1+α:

$$\Sigma_{(1+\alpha)} = (I - \boldsymbol{WH})(1+\alpha)\Psi_{XX}(I - \boldsymbol{WH})^T + \boldsymbol{W}\boldsymbol{R}_{(\boldsymbol{1}+\boldsymbol{\alpha})}\boldsymbol{W}^{\boldsymbol{T}} \tag{2.42}$$

where $\boldsymbol{R}_{(\boldsymbol{1}+\boldsymbol{\alpha})} = \begin{bmatrix} R & 0 \\ 0 & (1+\alpha)\Psi_{XX} \end{bmatrix}$. The KF solution for (2.42) is identical to the standard KF solution but with $\Sigma_{k|k-1}$ replaced by $(1+\alpha)\Sigma_{k|k-1}$:

$$\widehat{X}_{k|k} = \left[H_k^T R_k^{-1} H_k + \left\{(1+\alpha)\Sigma_{k|k-1}\right\}^{-1}\right]^{-1} \left[H_k^T R_k^{-1} Z_k + \left\{(1+\alpha)\Sigma_{k|k-1}\right\}^{-1} \widehat{X}_{k|k-1}\right] \tag{2.43}$$

With **WH**=I in (2.43) for the VIKF solution, we have $\Sigma_{(1+\alpha)} = \boldsymbol{W}\boldsymbol{R}_{(\boldsymbol{1}+\boldsymbol{\alpha})}\boldsymbol{W}^{\boldsymbol{T}}$for the apparent filtered error variance of $\widehat{X}_{k|k}$ in (42). The error covariance of $\widehat{X}_{k|k}$, $\Sigma_{k|k}$, is given by (2.3) as:

$$\begin{aligned} \Sigma_{k|k} = \boldsymbol{WRW^T} &= \left[\boldsymbol{H^T R^{-1}_{(1+\alpha)} H}\right]^{-1} \boldsymbol{H^T R^{-1}_{(1+\alpha)} R R^{-1}_{(1+\alpha)} H} \left[\boldsymbol{H^T R^{-1}_{(1+\alpha)} H}\right]^{-1} \\ &= \Sigma_{(1+\alpha),k|k} \Sigma^{-1}_{(1+\alpha)^2,k|k} \Sigma_{(1+\alpha),k|k} \end{aligned} \tag{2.44}$$

In (2.44), the inflated filtered error covariance, $\Sigma_{\beta,k|k}$, where β denotes the multiplicative inflation factor, is given by:

$$\begin{aligned} \Sigma_{\beta,k|k} &= \beta\Sigma_{k|k-1} - \beta\Sigma_{k|k-1}H_k^T\left[H_k\beta\Sigma_{k|k-1}H_k^T + R_k\right]^{-1}H_k\beta\Sigma_{k|k-1} \\ &= \left[H_k^T R_k^{-1} H_k + \left(\beta\Sigma_{k|k-1}\right)^{-1}\right]^{-1} \end{aligned} \tag{2.45}$$

Computationally, evaluation of (2.43) and (2.44) requires solving two (m×n) and an (m×m) linear systems. As in the original formulation of CBPKF, iterative reduction of α is necessary to ensure $\Sigma_{k|k} \leq \Sigma_{k|k-1}$.

The above approximation assumes that the CB penalty, Σ_{CB}, is proportional to the error covariance, Σ_{EV}. To help ascertain how KF, CBPKF and the VIKF approximation may differ, we compare in Table 2-1 their analytical solutions for gain κ_k, and filtered error variance $\sigma^2_{k|k}$ for the 1D case of m=n=1. The table shows that the VIKF approximation and CBPKF are identical for the 1D problem except that the CB penalty for CBPKF is twice as large as that for the VIKF approximation. To visualize the differences, Fig. 2-1 shows κ_k and $\sigma^2_{k|k}$ for KF, the VIKF approximation

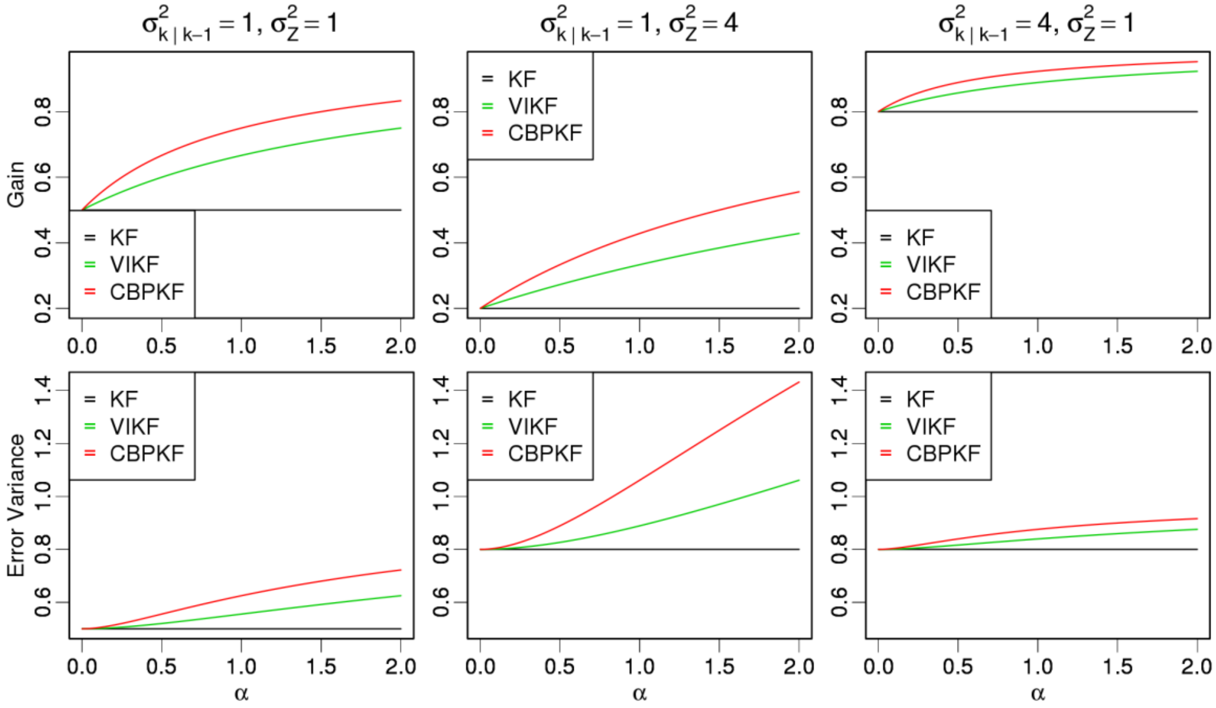


Figure 2-1 Comparison of κ_k and $\sigma^2_{k|k}$ for KF, VIKF and CBPKF for three different cases: $\sigma^2_{k|k-1} = 1$ and $\sigma^2_Z = 1$ (Case 1), $\sigma^2_{k|k-1} = 1$ and $\sigma^2_Z = 4$ (Case 2), $\sigma^2_{k|k-1} = 4$ and $\sigma^2_Z = 1$ (Case 3)

Table 2-1 Comparison of gain and filtered error variance among KF, the VIKF approximation, and CBPKF.

	Gain	**Filtered error variance, $\boldsymbol{\sigma^2_{k\|k}}$**
KF	$\frac{h\sigma^2_{k\|k-1}}{h^2\sigma^2_{k\|k-1}+\sigma^2_Z}$	$\frac{\sigma^2_Z}{h^2\sigma^2_{k\|k-1}+\sigma^2_Z}\sigma^2_{k\|k-1}$
VIKF approx.	$\frac{h(1+\alpha)\sigma^2_{k\|k-1}}{h^2(1+\alpha)\sigma^2_{k\|k-1}+\sigma^2_Z}$	$\frac{\{(1+\alpha)^2h^2\sigma^2_{k\|k-1}+\sigma^2_Z\}\sigma^2_Z}{\{(1+\alpha)h^2\sigma^2_{k\|k-1}+\sigma^2_Z\}^2}\sigma^2_{k\|k-1}$
CBPKF	$\frac{h(1+2\alpha)\sigma^2_{k\|k-1}}{h^2(1+2\alpha)\sigma^2_{k\|k-1}+\sigma^2_Z}$	$\frac{\{(1+2\alpha)^2h^2\sigma^2_{k\|k-1}+\sigma^2_Z\}\sigma^2_Z}{\{(1+2\alpha)h^2\sigma^2_{k\|k-1}+\sigma^2_Z\}^2}\sigma^2_{k\|k-1}$

and CBPKF for the three cases of $\sigma^2_{k|k-1} = 1$ and $\sigma^2_Z = 1$ (left), $\sigma^2_{k|k-1} = 1$and $\sigma^2_Z = 4$(middle),and $\sigma^2_{k|k-1} = 4$ and $\sigma^2_Z = 1$ (right). For all cases, we set h to unity and varied $\alpha$from 0 to 1.The figure indicates that, compared to KF, the VIKF approximation and CBPKF prescribe appreciably larger gains, that the increase in gain is larger for CBPKF solutions are the smallest for $\sigma^2_{k|k-1} > \sigma^2_Z$, a reflection of the diminished impact of CB owing to the comparatively smaller uncertainty in the observations. The above development suggests that one may be able to approximate CBPKF very closely with the VIKF-based formulation by adjusting α in the latter. Below, we evaluate the performance of CBPKF relative to KF and the VIKF-based approximation of CBPKF.

2.5 Evaluation and Results

For comparative evaluation, we carried out the synthetic experiments of (Seo et al., 2018a). We assume the following linear dynamical and observation models with perfectly known statistical parameters:

$$X_k = \Phi_{k-1}X_{k-1} + W_{k-1} \tag{2.46}$$

$$Z_k = H_kX_k + V_k \tag{2.47}$$

where X_k and X_{k-1} denote the state vectors at time steps k and k-1, respectively, Φ_{k-1} denotes the state transition matrix at time step k-1 assumed as $\Phi_{k-1} = \varphi_{k-1}I$, W_{k-1} denotes the white noise vector, $w_{j,k-1} \sim N(0, \sigma^2_{w_{k-1}})$, j=1,…,m, with $Q_{k-1} = E[W_{k-1}W^T_{k-1}]$, and V_k denotes the observation error vector, $v_{i,k} \sim N(0, \sigma^2_{v_k})$, i=1,…,n. The number of observations, n, is assumed to be time-invariant. The observation errors are assumed to be independent among themselves and of the true state. To assess comparative performance under widely varying conditions, we randomly perturbed φ_{k-1}, $\sigma_{w,k-1}$ and $\sigma_{v,k}$ above according to (2.48) through (2.50) below, and used only those deviates that satisfy the bounds:

$$\phi^p_{k-1} = \phi_{k-1} + \gamma_\phi \varepsilon_\phi \qquad 0.5 \le \phi^p_{k-1} \le 0.95 \tag{2.48}$$

$$\sigma^p_{w,k-1} = \sigma_{w,k-1} + \gamma_w \varepsilon_w \qquad \sigma^p_{w,k-1} \ge 0.01 \tag{2.49}$$

$$\sigma^p_{v,k} = \sigma_{v,k} + \gamma_v \varepsilon_v \qquad \sigma^p_{v,k} \ge 0.01 \tag{2.50}$$

In the above, the superscript p signifies that the variable is a perturbation, $\varepsilon_{\phi,}$ ε_w and ε_v denote the normally-distributed white noise for the respective variables, and γ_ϕ, γ_w and γ_v denote the standard deviations of the white noise added to ϕ_{k-1}, $\sigma_{w,k-1}$ and $\sigma_{v,k}$, respectively. The parameter settings (see Table 2-1) are chosen to encompass less predictable (small φ_{k-1}) to more predictable (large φ_{k-1}) processes, certain (small $\sigma_{w,k-1}$) to uncertain (large $\sigma_{w,k-1}$) model dynamics, and more informative (small $\sigma_{v,k}$) to less informative (large $\sigma_{v,k}$) observations. The bounds for ϕ^p_{k-1} in (2.48) is based on the range of lag-1 serial correlation representing moderate to high predictability where CBPKF and KF are likely to differ the most. The bounding of the perturbed values $\sigma^p_{w,k-1}$ and $\sigma^p_{v,k}$ in (2.49) and (2.50), respectively, is necessary to avoid the observational or model prediction uncertainty becoming unrealistically small. Very small $\sigma^p_{w,k-1}$ and $\sigma^p_{v,k}$render the information content of the model prediction, $\Sigma_{k|k-1}$, and the observation, Z_k, respectively, very large, and hence keep the filters operating in unrealistically favorable conditions for extended periods of time. We then apply KF, CBPKF and the VIKF approximation to obtain $\hat{X}_{k|k}$ and $\Sigma_{k|k}$, and verify them against the assumed truth. To evaluate the performance of CBPKF relative to KF, we calculate percent reduction in root mean square error (RMSE) by CBPKF over KF conditional on the true state exceeding some threshold between 0 and the largest truth.

Figure 2-2 show the percent reduction in RMSE by CBPKF over KF for Cases 1 (left), 5 (middle) and 9 (right) representing Groups 1, 2 and 3 in Table I, respectively. The three groups differ most significantly in the variability of the dynamical model error, γ_w, and may be characterized as nearly stationary (Group 1), nonstationary (Group 2), and highly nonstationary (Group 3). The range of α values used is [0.1, 1.2] with an increment of 0.1. The numbers of state variables, observations, and updating cycles used in Fig. 2-2 are 1, 10, and 100,000 for all cases. The dotted line at 10% reduction in the figure serves as a reference for significant improvement. The figure shows that, at the extreme end of the tail, CBPKF with α of 0.7, 0.6 and 0.5 reduces RMSE by about 15, 25 and 30% for Cases 1, 5 and 9, respectively, but at the expense of increasing unconditional RMSE by about 5%.The general pattern of reduction in RMSE for other cases in Table 2-2

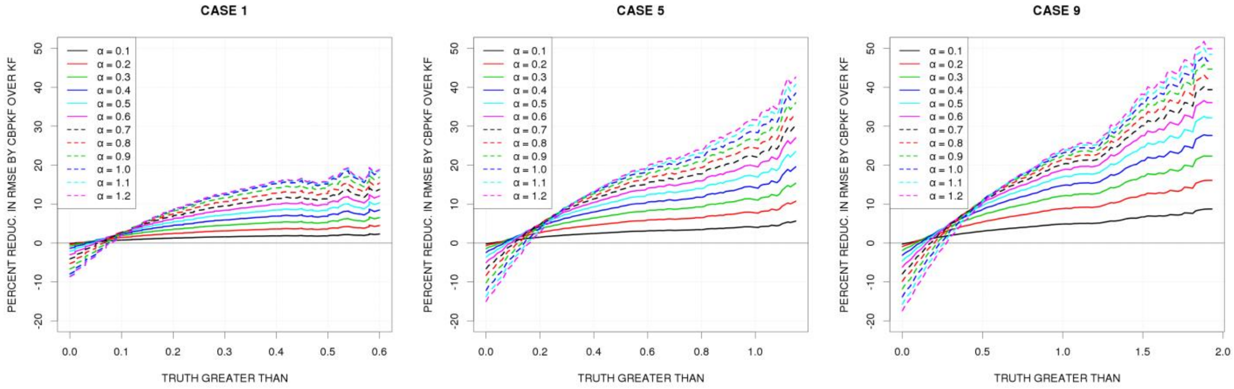


Figure 2-2 Percent reduction in RMSE by CBPKF over KF for a range of values of α for Cases 1 (left), 5 (middle) and 9 (right).

Table 2-2 Parameter settings for the 12 cases considered.

Group	Case	$\sigma_{w,k-1}$	γ_w	$\sigma_{v,k}$	γ_v	φ_{k-1}	γ_φ
1	1	0.1	0.01	1.5	0.4	0.7	0.1
	2	0.1	0.01	1.5	0.4	0.7	0.8
	3	0.1	0.01	1.5	1.2	0.7	0.1
	4	0.1	0.01	1.5	1.2	0.7	0.8
2	5	0.1	0.1	1.5	0.4	0.7	0.1
	6	0.1	0.1	1.5	0.4	0.7	0.8
	7	0.1	0.1	1.5	1.2	0.7	0.1
	8	0.1	0.1	1.5	1.2	0.7	0.8
3	9	0.1	0.2	1.5	0.4	0.7	0.1
	10	0.1	0.2	1.5	0.4	0.7	0.8
	11	0.1	0.2	1.5	1.2	0.7	0.1
	12	0.1	0.2	1.5	1.2	0.7	0.8

is similar within each group and is not shown. We only note here that larger variability in observational uncertainty (i.e., larger γ_v) reduces the relative performance of CBPKF somewhat, and that the magnitude of variability in predictability (i.e., γ_φ) has relatively small impact on the relative performance.

It was seen in Table 2-1 that the VIKF approximation is identical to CBPKF for m=n=1 but for the multiplicative scaler weight for the CB penalty. Numerical experiments indicate that, whereas the above relationship does not hold for other m or n, one may very closely approximate CBPKF with the VIKF-based formulation by adjustingα. For example, the VIKF approximation with α increased by a factor of 1.25 to 1.90 differ from CBPKF only by 1% or less for all 12 cases in Table 2-2 with m=1 and n=10. The above findings indicate that the VIKF approximation may be used as a computationally less expensive alternative for CBPKF. Table 2-3 compares the CPU time among KF, CBPKF and the VIKF approximation for 6 different combinations of m and n based using Intel(R) Xeon(R) Gold 6152 CPU @ 2.10GHz. The computing time is reported in multiples of the KF's. Note that the original formulation of CBPKF quickly becomes extremely expensive as the dimensionality of the problem increases whereas the CPU time of the VIKF approximation stays under 3.5 times that of KF for the size of the problems considered.

Table 2-3 Comparison of computing time among KF, CBPKF and VIKF approximation.

Dimensionality		**Normalized Computing Time**		
m	n	KF	CBPKF	VIKF approx.
1	10	1	5.23	1.51
1	40	1	18.41	2.74
5	10	1	6.44	1.67
5	40	1	24.03	2.88
10	10	1	14.27	2.03
10	40	1	27.96	3.46

If the filtered error variance is unbiased, one would expect the mean of the actual error squared associated with the variance to be approximately the same as the variance itself. To verify this, we show in Fig. 2-3 the filtered error variance vs. the actual error squared for KF (left), the VIKF approximation (middle) and CBPKF (right) for all ranges of filtered error variance. For reference, we plot the one-to-one line representing the unbiased error variance conditional on the magnitude of the filtered error variance and overlay the local regression fit through the actual data points using the R package locfit (Catherine Loader, 2013). The figure shows that all three provide conditionally unbiased estimates of filtered error variance as theoretically expected, and that the VIKF approximation and CBPKF results are extremely similar to each other.

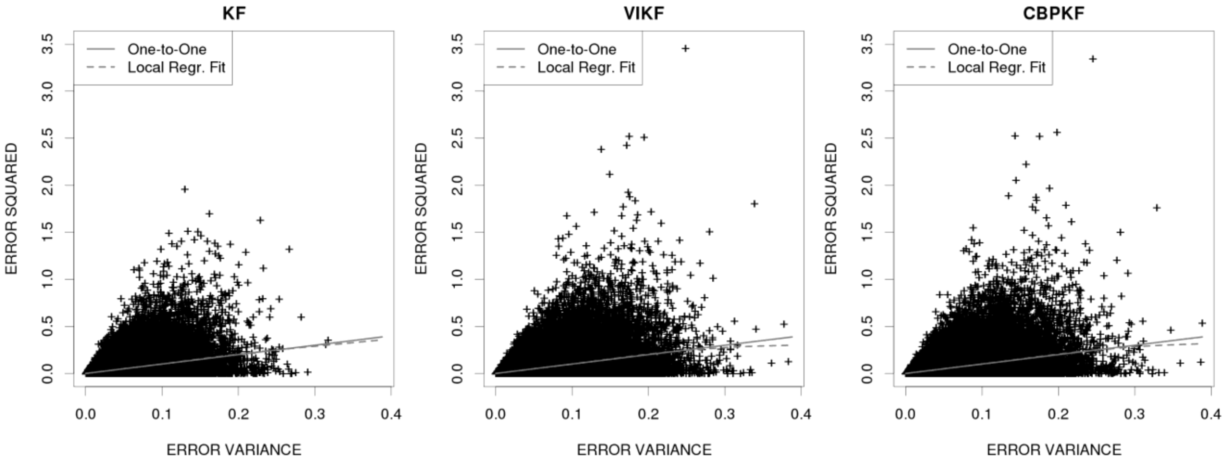


Figure 2-3 Filtered error variance vs. error squared for KF (left), the VIKF approximation (middle) and CBPKF (right). The one-to-one line is shown in black and the local regression fit is shown in green.

2.6 Adaptive CBPKF

Whereas CBPKF or the VIKF approximation significantly improves the accuracy of the estimates over the tails, it deteriorates performance near the median. Fig. 2-2 suggests that, if α can be prescribed adaptively such that a small/large CB penalty is effected when the system is in the normal/extreme state, the unconditional performance of CBPKF would improve. Because the true state of the system is not known, adaptively specifying α is necessarily an uncertain proposition. There are, however, certain applications in which the normal-vs.-extreme state of the system may be ascertained with higher accuracy than others. For example, the soil moisture state of a catchment may be estimated from assimilating precipitation and streamflow data into hydrologic models (Seo et al., 2003; Lee et al., 2011, 2012, 2015; Lee and Seo, 2014; Rafieeinasab et al., 2014b). If α is prescribed adaptively based on the best available estimate of the state of the catchment, one may expect improved performance in hydrologic forecasting. In this section, we apply adaptive CBPKF in the synthetic experiment and assess its performance. An obvious strategy for adaptively filtering is to parameterize α in terms of the KF estimate (i.e., the CBPKF estimate with $\alpha = 0$) as the best guess for the true state. The premise of this strategy is that, though it may be conditionally biased, the KF estimate fuses the information available from both the observations and the dynamical model, and hence best captures the relationship between α and the departure of the state of the system from median. A similar approach has been used in fusing radar rainfall data and rain gauge observations for multisensor precipitation estimation in which ordinary cokriging estimate was used to prescribe α in CB-penalized cokriging (Kim et al., 2018, p. 201).

Necessarily, the effectiveness of the above strategy depends on the skill of the KF estimate; if the skill is very low, one may not expect significant improvement. Fig. 2 suggests that, qualitatively, α should increase as the state becomes more extreme. To that end, we employed the following model for time-varying α:

$$\alpha_k = \gamma\left\|\hat{X}_{k|k}^{KF}\right\| \tag{2.51}$$

where α_k denotes the multiplicative CB penalty factor for CBPKF at time step k, $\left\|\hat{X}_{k|k}^{KF}\right\|$ denotes some norm of the KF estimate at time step k, and γ denotes the proportionality constant.

Figure 2-4a shows the RMSE reduction by adaptive CBPKF over KF with $\alpha_k = \gamma\left|\hat{X}_{k|k}^{KF}\right|$ for the 12 cases in Table 2-2 m=1 and n=10. The γ values used were 3.0, 1.0 and 0.5 for Groups 1, 2 and 3 in Table 2-2, respectively. The figure shows that adaptive CBPKF performs comparably to KF in the unconditional sense while substantially improving performance in the tails. The rate of reduction in RMSE with respect to the increasing conditioning truth, however, is now slower than that seen in Fig. 2-2 due to the occurrences of incorrectly specified α. To assess the uppermost

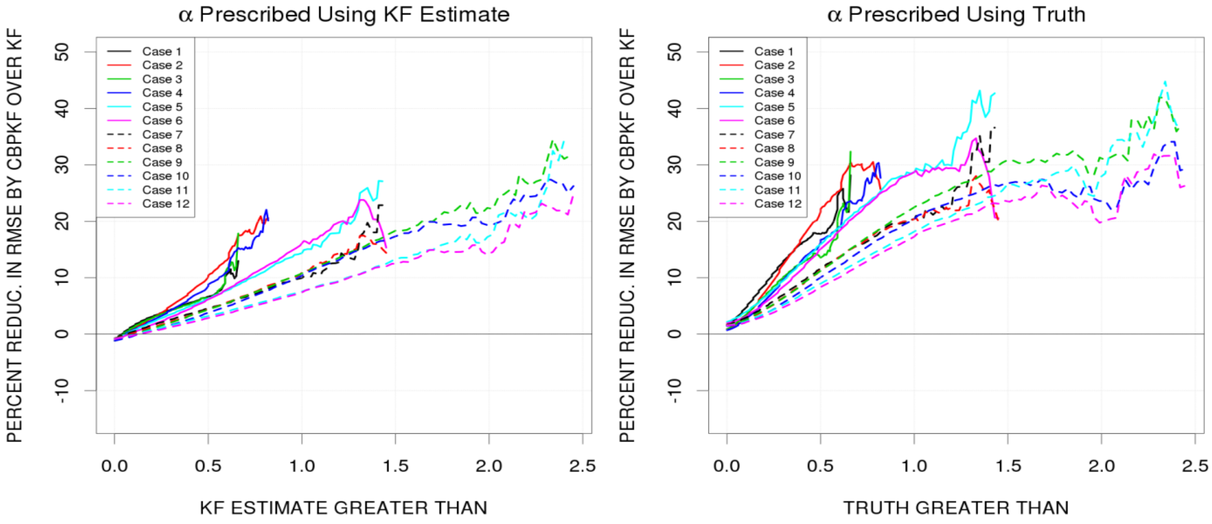


Figure 2-4 Percent reduction in RMSE by adaptive CBPKF over KF in which α is prescribed using the KF estimate (left) and the truth (right).

bound of the feasible performance of adaptive CBPKF, we also specified α with perfect accuracy under (2-51) via $\alpha_k = \gamma|X_k|$ where X_k denotes the true state.The results are shown in Fig. 2-4b for which the γ values used were 3.0, 1.5 and 1.0 for Groups 1, 2 and 3 in Table II, respectively. The figure indicates that adaptive CBPKF with perfectly prescribed α greatly improves performance, even outperforming KF in the unconditional sense. Fig. 2-4 suggests that, if α can be prescribed more accurately with additional sources of information, the performance of adaptive CBPKF may be improved beyond the level seen in Fig. 2-4a. Finally, we show in Fig. 2-5 the example scatter plots of the KF (black) and adaptive CBPKF (red) estimates vs. truth. They are for Cases 1 and 9 in Table 2-2 representing Groups 1 and 3, respectively. It is readily seen that the CBPKF significantly reduces CB in the tails while keeping its estimates close to the KF estimates in normal ranges.

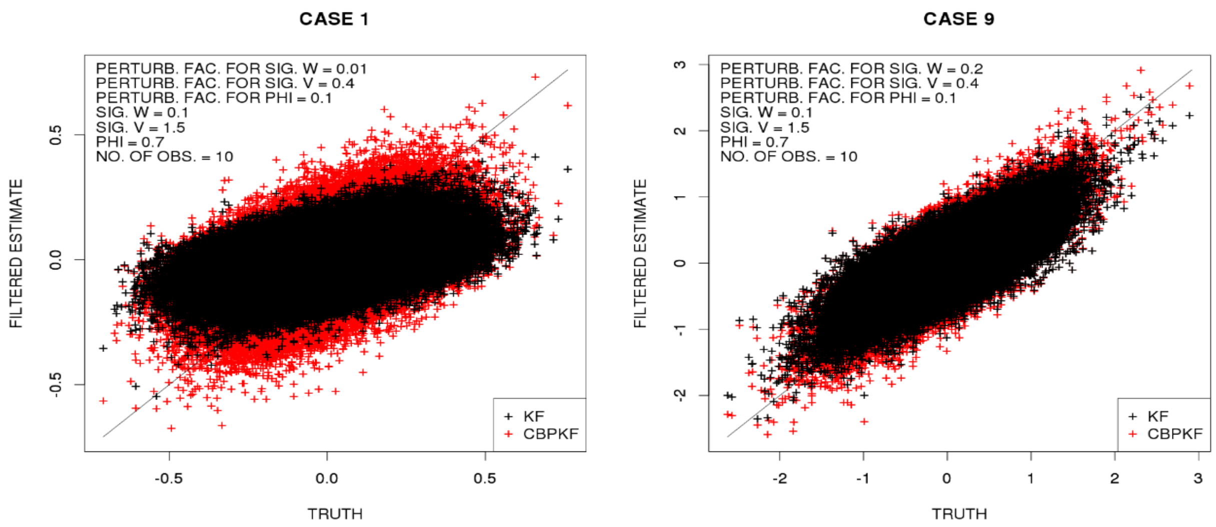


Figure 2-5 Example scatter plots of KF (black) and adaptive CBPKF (red) estimates vs. truth for Cases 1 (left) and 9 (right) in Table 2-2.

2.7 Conclusions

Conditional bias-penalized Kalman filter (CBPKF) has recently been developed to improve estimation and prediction of extremes. The original formulation, however, is computationally very expensive, and deteriorates performance in the normal ranges relative to KF. In this chapter, we present a computationally less expensive alternative based on the variance-inflated KF (VIKF) approximation, and improve unconditional performance by adaptively prescribing the weight for the CB penalty. For evaluation, we carried out synthetic experiments using linear systems with varying degrees of dynamical model uncertainty, observational uncertainty, and predictability. The results indicate that the VIKF-based approximation of CBPKF provides a computationally much less expensive alternative to the original formulation, and that adaptive CBPKF performs comparably to KF in the unconditional sense while improving estimation of extremes by about 20% to 30% over KF. It is also shown that additional improvement may be possible by improving adaptive prescription of the weight to the CB penalty using additional sources of information. The findings indicate that adaptive CBPKF offers a significant addition to the dynamic filtering methods for general application in signal processing and, in particular, when or where estimation of extremes is of importance. The findings in this work are based on idealized synthetic experiments that satisfy linearity and normality. Additional research is needed to assess performance for non-normal problems and for nonlinear problems using the ensemble extension (Lee et al., 2018), and to prescribe the weight for the CB penalty more skillfully.

Chapter 3. CBEnKF and Application to Streamflow Prediction

3.1 Introduction

In this chapter, we describe the ensemble extension of the CBPKF, or conditional bias-penalized ensemble Kalman filter (CBEnKF), toward addressing conditional bias in hydrologic DA. Ensemble forecasting has been fast gaining acceptance in operational water forecasting for uncertainty quantification and risk-based decision making (Cloke and Pappenberger, 2009; Demargne et al., 2014). As in single-valued forecasting, effective real-time DA is critical in ensemble forecasting to keeping the model states in line with the unfolding reality and hence to increasing predictive skill (Noh et al. 2018). The potency of manual adjustment of model states (Seo et al. 2003, 2009) is well recognized in single-valued forecasting, in particular, of extreme events. In ensemble forecasting, however, such manual DA may not be considered feasible due to the higher-order nature of the information content. The purpose of the CBEnKF is to offer an ensemble DA technique which is superior to the EnKF for prediction of extreme events such as floods. In this work, we apply the CBEnKF for streamflow DA using the lumped Sacramento Soil Moisture Accounting (SAC, Burnash et al., 1973) and unit hydrograph (UH, Chow et al., 1988) models. We then comparatively evaluate the CBEnKF with the EnKF in a set of twin experiments wherein streamflow, precipitation and potential evapotranspiration (PE) observations are assimilated into the SAC-UH for high flow events. The DA experiments are carried out for ten headwater catchments in Texas (U.S.) in diverse climatological regions and of contrasting hydrologic response characteristics. The new and significant contributions of this chapter are: the development of the CBEnKF technique, comprehensive comparative evaluation of the CBEnKF with the EnKF for prediction of high flows using widely used and extensively studied hydrologic models, and the attribution and interpretation of the skill added by the CBEnKF. We note here that the CBPKF and CBEnKF are not bias-aware techniques used to address unconditional bias (Dee and da Silva, 1998; Dee, 2005; Drecourt et al., 2006; De Lannoy et al., 2007a; De Lannoy et al., 2007b; Pauwels and De Lannoy, 2015). As with the KF or EnKF, it is readily possible to render the CBPKF or CBEnKF, respectively, bias-aware via state augmentation. In this work, we deal with systematic biases by introducing and estimating time-varying biases in the control variables in both the EnKF and CBEnKF in an identical manner. This chapter is an adaptation of Lee et al. (2019). The remainder of this chapter is organized as follows. Section 3.2 describes the methodology. Section 3.3 describes the study basins, evaluation metrics and observational uncertainty modelling. Section 3.4 presents the results and discussion. Section 3.5 provides the conclusions and future research recommendations.

3.2 Methodology

This section describes the DA problem, the CBEnKF solution, algorithmic considerations for implementation, the hydrologic models used, and the sequential assimilation approach based on fixed-lag smoothing.

3.2.1 Problem description and formulation

The assimilation problem at hand may be described as follows. Given the a priori model soil moisture states and observations of streamflow, precipitation and PE, update the model soil moisture states and multiplicative biases for observed precipitation and PE by assimilating observed streamflow, precipitation and PE. This nonlinear DA problem may be solved via a number of different techniques, including EnKF with state augmentation (Lorentzen and Nævdal 2011). The main motivation for an ensemble approach is to capture the nonlinear model dynamics without having to approximate them. Among the ensemble techniques, EnKF is by far the most popular owing to its simplicity, and hence serves as a benchmark in this work. The purpose of state augmentation is to render the nonlinear observation equation nominally linear for the KF solution. For an ensemble technique with state augmentation, one may write the dynamical and observation equations, or models, as follows. The nonlinear dynamical model is written as:

$$X_k = M(X_{k-1}) + W_{k-1} \tag{3.1}$$

where X_k denotes the $(n_c \times 1)$ model state, or control, vector, where n_c denotes the number of variables in the control vector, M( ) denotes the dynamical model for the state variables, and W_{k-1} denotes the dynamical model error at time step $k-1$. The observational model is written as:

$$Z_k = H_k Y_k + V_k \tag{3.2}$$

where Z_k denotes the (n×1) observation vector where n denotes the total number of observations. The (n×m) observation matrix, H_k, where $m=n_c+n$, consists of the $(n \times n_c)$ zero matrix and $(n \times n)$ identity matrix, i.e., $H_k = [0 \quad I]$. The (m×1) augmented state vector Y_k is given by $Y_k = [X_k \quad G(X_k)]^T$ where X_k denotes the $(n_c \times 1)$ state vector defined in Eq.(3.1) and $G(X_k)$ denotes the generally nonlinear function which maps the state variables to the (n×1) augmented state subvector. The above state augmentation reflects the nonlinear relations between the state variables and the observed variables, such as those between the state variables, X_k, and the observed flow, Z_k^Q, while keeping the observation equation linear in appearance (Evensen, 2003; Clark et al., 2008; Lorentzen and Naevdal, 2011; Houtekamer and Zhang, 2016). Lastly, V_k denotes the (n×1) observation error vector at time step k with its covariance matrix R.

Because the observations of streamflow, precipitation and PE are generally uncertain, one may expect the DA solutions based solely on minimizing error variance such as the KF or EnKF to be conditionally biased as explained in the Introduction Section. The objective of the CBPKF or CBEnKF is to reduce the Type-II CB by minimizing the weighted sum of the error covariance and the expectation of the Type-II CB squared (Seo et al., 2018a, 2018b; Shen et al., 2019) as shown below:

$$\Sigma_k = E_{Y_k,Y_k^*}[(Y_k - Y_k^*)(Y_k - Y_k^*)^T] + \alpha E_{Y_k}\left[\left(Y_k - E_{Y_k^*}[Y_k^*|Y_k]\right)\left(Y_k - E_{Y_k^*}[Y_k^*|Y_k]\right)^T\right] \tag{3.3}$$

where Y^* denotes the vector of the estimated states, the first and second terms represent the error covariance and the CB penalty, respectively, and α denotes the weight given to the latter. If $\alpha=0$, the CBPKF or CBEnKF is reduced to the KF or EnKF, respectively.

3.3 Conditional bias-penalized Ensemble Kalman Filter (CBEnKF)

Because the hydrologic model dynamics in Eq.(3.1) are generally nonlinear, one may not apply the CBPKF directly. Instead, an ensemble filter is necessary. The formulation of the ensemble version of the CBPKF, or CBEnKF, is identical to that of the EnKF except for the Kalman gain and the filtered error covariance. For a general description of EnKF, the reader is referred to Lorentzen and Nævdal (2011). For a description in the context of the DA problem at hand, the reader is referred to Rafieeinasab et al. (2014b). Below, we describe the CBEnKF solution which reflects the important corrections to Seo et al. (2018a) as reported in Seo et al. (2018b).

As in the EnKF, the updated state vector of the i^{th} ensemble member, $Y_{k|k}^i$, is given by:

$$Y_{k|k}^i = Y_{k|k-1}^i + K_k\left[Z_k - H_k Y_{k|k-1}^i\right] \tag{3.4}$$

where $Y_{k|k-1}^i$ denotes the (m×1) forecast state vector of the i^{th} ensemble member, K_k denotes the (m×n) CB-penalized Kalman gain matrix, Z_k denotes the (n×1) observation vector and H_k denote the (n×m) observation matrix. The gain, K_k, and the error covariance, $\Sigma_{k|k}$, are given by:

$$K_k = \left[\varpi_{1,k} H_k + \varpi_{2,k}\right]^{-1} \varpi_{1,k} \tag{3.5}$$

$$\Sigma_{k|k} = \left[\varpi_{1,k}H_k + \varpi_{2,k}\right]^{-1}\left(\varpi_{1,k}R_k\varpi_{1,k}^T + \varpi_{2,k}\Sigma_{k|k-1}\varpi_{2,k}^T\right)\left[\varpi_{1,k}H_k + \varpi_{2,k}\right]^{-1} \tag{3.6}$$

In Eqs.(3.5) and (3.6), the (m×n) and (m×m) weight matrices for the observations and model prediction, $\varpi_{1,k}$ and $\varpi_{2,k}$, respectively, are given by:

$$\varpi_{1,k} = \widehat{H}_k^T\Gamma_{11,k} + \Gamma_{21,k} \tag{3.7a}$$

$$\varpi_{2,k} = \widehat{H}_k^T\Gamma_{12,k} + \Gamma_{22,k} \tag{3.7b}$$

In Eq.(3.7), the (m×n) modified observation matrix, $\widehat{H}_k^T$, and the (n×n), (n×m) and (m×m) matrices, $\Gamma_{11,k}$, $\Gamma_{12,k}$ $(= \Gamma_{21,k}^T)$ and $\Gamma_{22,k}$, respectively, are given by:

$$\widehat{H}_k^T = H_k^T + \alpha C_k^T \tag{3.8}$$

$$\Gamma_{22,k}^{-1} = \Lambda_{22,k} - \Lambda_{21,k}\Lambda_{11,k}^{-1}\Lambda_{12,k} \tag{3.9a}$$

$$\Gamma_{11,k} = \Lambda_{11,k}^{-1} + \Lambda_{11,k}^{-1}\Lambda_{12,k}\Gamma_{22,k}\Lambda_{21,k}\Lambda_{11,k}^{-1} \tag{3.9b}$$

$$\Gamma_{12,k} = -\Lambda_{11,k}^{-1}\Lambda_{12,k}\Gamma_{22,k} \tag{3.9c}$$

In Eqs.(3.8) and (3.9), the (n×m) gain matrix for the observation vector, C_k, and the (n×n), (n×m) and (m×m) modified error covariance matrices, $\Lambda_{11,k}$, $\Lambda_{12,k}$ $(= \Lambda_{21,k}^T)$ and $\Lambda_{22,k}$, respectively, are given by:

$$C_k = \left[H_k\Sigma_{k|k-1}G_{2,k}^{-1} + R_kH_k\right]\left[G_{2,k}^{-1}\Sigma_{k|k-1}G_{2,k}^{-1} + 2(H_k^TR_kH_k + \Sigma_{k|k-1})\right]^{-1}G_{2,k}^{-1} \tag{3.10}$$

$$\Lambda_{11,k} = R_k + \alpha(1-\alpha)C_k\Sigma_{k|k-1}C_k^T + \Lambda_{12,k}H_k^T + H_k\Lambda_{21,k} \tag{3.11a}$$

$$\Lambda_{12,k} = -\alpha C_k\Sigma_{k|k-1} \tag{3.11b}$$

$$\Lambda_{22,k} = \Sigma_{k|k-1} \tag{3.11c}$$

In Eqs.(3.10) and (3.11), R_k and $\Sigma_{k|k-1}$ denote the (n×n) observation error covariance matrix, $Cov[V_k, V_k]$, and the (m×m) forecast error covariance matrix, respectively, and $G_{2,k}^{-1} = H_k^T H_k + I$. The forecast error covariance matrix is estimated by:

$$\Sigma_{k|k-1} = \overline{\left(Y_{k|k-1} - \overline{Y_{k|k-1}}\right)\left(Y_{k|k-1} - \overline{Y_{k|k-1}}\right)^T} \tag{3.12}$$

where the overbar represent an average over the ensemble members and $\overline{Y_{k|k-1}}$ denotes the ensemble mean vector of the state vector $Y_{k|k-1}$ (Evensen, 2003). With $\alpha = 0$, we have $\Lambda_{12,k} = 0, \Lambda_{11,k} = R_k, \Lambda_{22,k} = \Sigma_{k|k-1}, \widehat{H}_k^T = H_k^T, \Gamma_{22,k}^{-1} = \Lambda_{22,k}, \Gamma_{11,k} = R_k^{-1}, \Gamma_{12,k} = 0, \varpi_{1,k} = H_k^T R_k^{-1}, \varpi_{2,k} = \Sigma_{k|k-1}^{-1}, \Sigma_{k|k} = \left[H_k^T R_k^{-1} H_k + \Sigma_{k|k-1}^{-1}\right]^{-1}$, and $K_k = \Sigma_{k|k} H_k^T R_k^{-1}$, which reduces the CBPKF solution to the KF. For further details, the reader is referred to Seo et al. (2018a,b) and Shen et al. (2019).

3.3.1 Algorithmic considerations

Because α in Eq.(3.3) can take on any nonnegative value, the CBPKF does not guarantee $\Sigma_{k|k} \leq \Sigma_{k|k-1}$ a priori. It is hence necessary to check the above inequality (or a less stringent necessary condition or conditions) and, if not satisfied, reduce α and repeat the calculations. For the CBEnKF, however, sampling errors in both $\Sigma_{k|k}$ and $\Sigma_{k|k-1}$ render the above check impractical, particularly when the ensemble size is small. For this reason, instead we check if the sample $\Sigma_{k|k}$ lies within the theoretically expected bounds using the apparent CBPKF error covariance, $\Sigma_{k|k}^a$, which represents Eq.(3.3) at minimum:

$$\Sigma_{k|k}^{a} = B\left[\varpi_{1,k} H_k + \varpi_{2,k}\right]^{-1} \tag{3.13}$$

where the (m×m) scaling matrix, B, is given by $B = \alpha\Sigma_{k|k-1}[\varpi_{1,k}\widehat{H}_{1,k} + \varpi_{2,k}] + I$. Because it represents the sum of error covariance and expectation of CB squared, Eq.(3.13) represents an upper bound for $\Sigma_{k|k}$. A lower bound for $\Sigma_{k|k}$ may also be obtained from Eq.(3.13) by setting $\alpha = 0$ or, equivalently, $B = I$. With the above, we have:

$$\left[\varpi_{1,k} H_k + \varpi_{2,k}\right]^{-1} \leq \Sigma_{k|k} \leq B\left[\varpi_{1,k} H_k + \varpi_{2,k}\right]^{-1} \tag{3.14}$$

For solving the linear systems in the KF or the EnKF, the least computationally-expensive method is the Cholesky decomposition (Press, 2007; Krishnamoorthy and Menon, 2013). In the CBPKF or CBEnKF, the linear systems may or may not be positive semidefinite. If the Cholesky decomposition is not successful in inverting $\Lambda_{11,k}$ or $\Gamma_{22,k}^{-1}$, or solving for K_k or C_k, it is necessary to use the LU decomposition (Press, 2007). Finally, if Eq.(3.14) is not satisfied, it is necessary to reduce α and repeat the steps. The initial value of α, or α_0, should be chosen carefully so that it is large enough to effect the CB penalty but not too large to incur too many iterations for successive reduction or over-correct the CB. An effective strategy is to start with a sufficiently large α_0 and apply the filter in a hindcast mode over a sufficiently long analysis period and examine the distribution of the final α. A good choice for α_0 should result in iterative reduction of α only for a small fraction of the times for computational economy. We are currently exploring adaptive approaches for prescribing α for improved performance and reduced computing time, and the interested reader is referred to Shen et al. (2019).

3.3.2 Hydrologic models

We used the SAC (Burnash et al., 1973) for rainfall-runoff modeling and the UH (Chow et al., 1988) for runoff routing. The SAC is a conceptual soil moisture accounting model and uses two layers of conceptual soil storages, i.e., the upper zone and lower zone, to model soil water balance in the vertical. The lower zone is usually much thicker than the upper zone and provides moisture to meet the evapotranspiration demands (Koren et al., 2004). Each soil layer includes tension and free water contents, specifically, the Upper Zone Tension Water Content (UZTWC), Upper Zone Free Water Content (UZFWC), Lower Zone Tension Water Content (LZTWC), Lower Zone Free Primary Content (LZFPC) and Lower Zone Free Supplemental Content (LZFSC). The UZTWC is directly fed by rainfall. Rainfall exceeding the tension water capacity becomes the UZFWC which generates interflow or percolates into the lower zone. Surface runoff is generated from free water contents depending on the capacities of upper zone free water and lower zone tension and free water. Surface runoff includes both fast and slow runoffs. Fast runoff consists of the rainfall intensity-dependent surface runoff from the unsaturated area, direct runoff from the saturated area, and impervious runoff from the impervious area. Slow runoff processes produce interflow, supplemental baseflow, and primary baseflow. The UH model is used to route the runoff components to produce streamflow at the outlet of the basin.

The SAC is a long-standing operational model used by the National Weather Service (NWS) for hydrologic forecasting at most River Forecast Centers. The input forcings of the model include mean areal precipitation (MAP) and mean areal potential evapotranspiration (MAPE). In this work, we use a 1-hr time step for all basins given the flashiness and short response time of many of the basins. To refine the SAC parameters and to estimate UHs that are valid at a 1-hr time step, we used the Adjoint-Based OPTimizer (AB_OPT; (Seo et al., 2009) which corrects long-term biases in MAP and MAPE, derives empirical unit hydrographs, and locally optimizes the SAC parameters. In this way, systematic biases in inputs, systematic timing errors, and model parametric uncertainties are reduced as much as possible so that DA may address the initial condition uncertainties as intended.

3.3.3 Assimilation approach

The CBEnKF and the EnKF are applied as sequential assimilators but using the fixed-lag smoother formulation (Figure 3-1) (Schweppe, 1973). The size of the fixed lag, or the assimilation window, is chosen to be the same as the length of the UH (see Table 3-1). In this way, the time lag between runoff generation and the resulting flow at the basin outlet is captured in the DA process (Seo et al., 2009; Rafieeinasab et al., 2014b). The motivation for the fixed-lag smoother approximation using a filter formulation, as opposed to directly using a smoother formulation (Evensen and van Leeuwen, 2000), is to handle both cold and warm restarts with a clearly identifiable assimilation window so that human forecasters may easily interpret the results and make decisions (Seo et al., 2009). In this approach, the control vector in Eq.(3.1) consists of $X_k = (X_k^S, X_k^P, X_k^E, X_k^R)^T$, i.e., the six SAC states, X_k^S, the multiplicative adjustment factors for MAP and MAPE, X_k^P and X_k^E, respectively, and the additive error to the Total Channel Inflow (TCI), X_k^R (see

Table 3-1 Hydrologic attributes of the study basins. T_P and $Q_{95|S}$ denote time-to-peak and 95th percentile observed flow within the significant events, respectively. T_P is estimated from the empirical unit hydrograph; dryness index is calculated as the ratio of mean annual potential evaporation to mean annual precipitation; $Q_{95|S}$ is calculated from the events whose peak flow exceeds 100 m^3/s (see text for additional details).

Basin	USGS ID	Area (km2)	Dryness index	UH duration (hr)	TP (hr)	Q95|S (m3/s)
Justin	08053500	1037	1.47	80	6	166.0
Houston	08076000	137	0.66	32	5	87.2
Greenville	08017200	212	0.92	35	16	106.4
Brays Bayou	08075000	246	0.39	24	3	80.8
Lyons	08110100	508	1.25	75	18	130.0
McKinney	08058900	427	0.91	41	14	135.5
Madisonville	08065800	870	1.05	50	21	164.5
Midfield	08162600	435	0.89	80	17	172.7
Quinlan	08017300	197	0.82	25	12	113.7
Schroeder	08176900	932	1.23	53	14	232.8

Figure 3-1 and Table 3-2). In the above, X_k^S, is valid at the beginning of the assimilation window; X_k^P, X_k^E, and X_k^R are valid over the entire assimilation window; and X_k^E and X_k^P represent time-varying systematic biases in PE and precipitation data (Gourley and Vieux, 2005; Seo et al., 2010; Stisen and Sandholt, 2010; Habib et al., 2013). The time evolution of X_k^S is given by the SAC. The time evolution of X_k^P, $X_k^E$and X_k^R is modeled as autoregressive-1 with a lag-1 (hr) serial correlation coefficient of 0.96 for all three (Smith and Krajewski, 1991; Anagnostou and Krajewski, 1999; Seo et al., 1999). The observation vector in Eq.(3.2) consists of $Z_k = \left(Z_k^P, Z_k^E, Z_k^R, Z_k^Q\right)^T$, i.e., the observed MAP, Z_k^P, the climatological MAPE, Z_k^E, the observed additive error in TCI, Z_k^R, and the

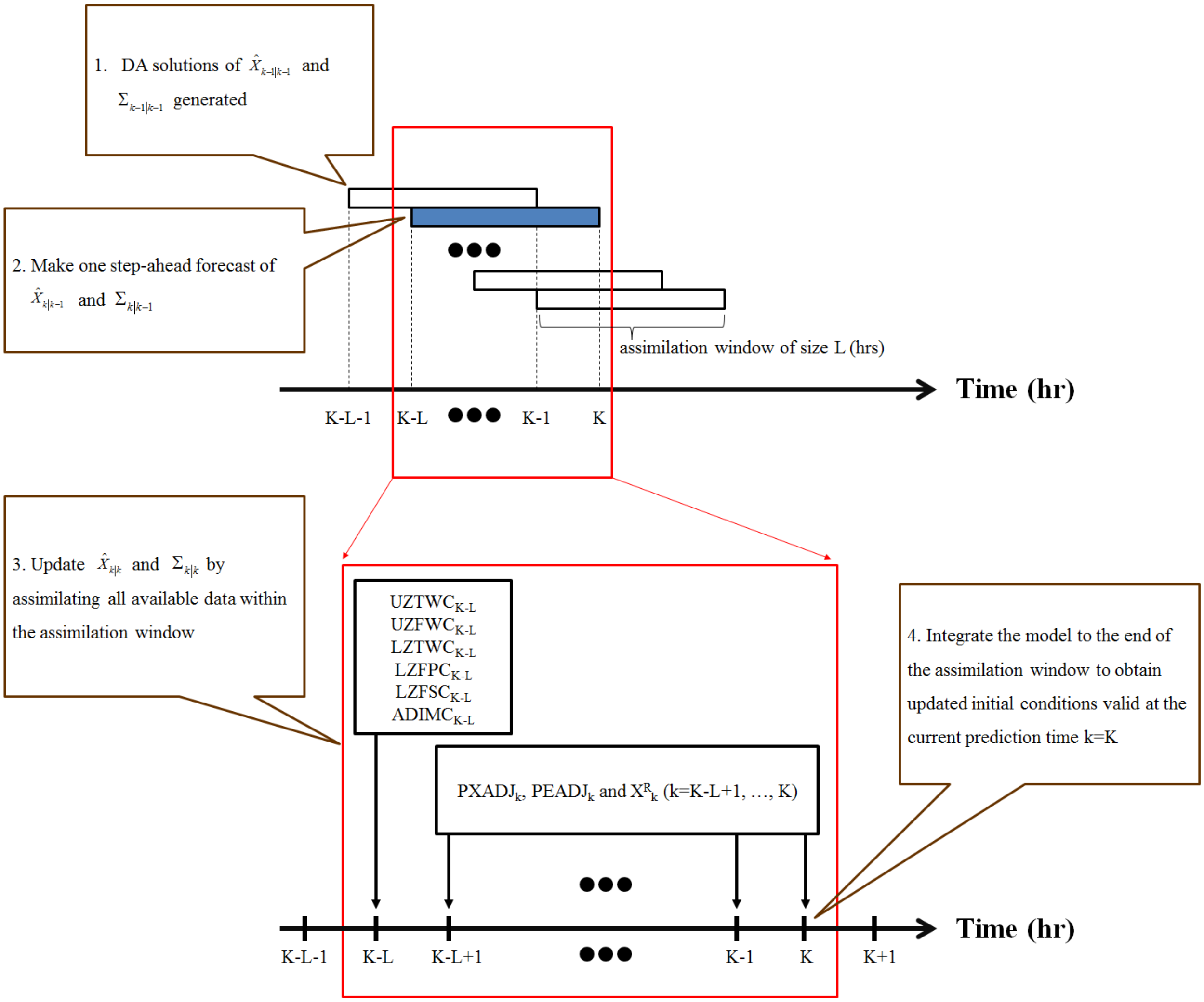


Figure 3-1 A schematic of the DA cycles based on the fixed-lag smoother formulation. See Table 3-2 for description of the control variables.

Table 3-2 Control variables solved for by the CBEnKF and EnKF for streamflow prediction using the SAC-UH

Control variable	Description
UZTWC	Upper Zone Tension Water Content
UZFWC	Upper Zone Free Water Content
LZTWC	Lower Zone Tension Water Content
LZFPC	Lower Zone Primary Free Water Content
LZFSC	Lower Zone Supplemental Free Water Content
ADIMC	Additional impervious area water content
PXADJ	Multiplicative adjustment factor for MAP
PEADJ	Multiplicative adjustment factor for MAPE
$\mathbf{X^R}$	Additive error to the TCI

streamflow observations, Z_k^Q. The observation error vector, $V_k = \left(V_k^P, V_k^E, V_k^R, V_k^Q\right)^T$, consists of the observation errors associated with $Z_k = \left(Z_k^P, Z_k^E, Z_k^R, Z_k^Q\right)^T$. It is assumed that the elements of V_k are uncorrelated among themselves and with X_k, and have mean of zero and covariance of R_k. The nonlinear observation equation, $G(X_k)$, consists of the model-simulated observations of MAP, or $Z_k^P X_k^P$, MAPE, or $Z_k^E X_k^E$, additive error in TCI, or X_k^R, and streamflow, or X_k^Q. In reality, Z_k^R is not observed. As such, we set it to zero under the assumption that the model-simulated TCI is unbiased at all scales.

The observations to be assimilated for each cycle include the MAP, MAPE, and streamflow valid over the assimilation window. The motivation for updating the multiplicative adjustment factors for MAP and MAPE is to account for dynamically-varying biases in the two forcing variables. In SAC, static adjustment factors already exist for MAP and MAPE, referred to as PXADJ and PEADJ, respectively (Seo et al., 2009). In this work, we include the two factors in the control vector and allow them to vary dynamically in the DA process. The MAP observations used in this work are from the Multisensor Precipitation Estimator (Seo et al., 2010). They are subject to various sources of error which manifest themselves as time-varying biases as well as random errors

(Smith et al., 1996; Fulton et al., 1998). Because MAP is a spatially averaged quantity, purely random errors tend to cancel out whereas systematic biases do not. As such, it is important in streamflow DA to correct dynamically-varying biases in MAP over the assimilation window so that the SAC states valid at the prediction time and the TCIs valid over the assimilation window are as accurate as possible (Seo et al., 2003, 2009). The motivation for the additive error to TCI is to allow for the time-varying errors in SAC-UH that cannot be addressed by adjusting the model states, PXADJ, and PEADJ alone. The above formulation has proven very effective in several previous DA studies (Seo et al., 2003, 2009; Lee et al., 2011, 2012, 2016; Lee and Seo, 2014; Rafieeinasab et al., 2014b). In this work, the adjustment factors and the additive error in TCI are assumed to be temporally uniform within the assimilation window following Lee et al. (2012) who found that the use of temporally varying errors does not significantly improve performance while substantially increasing the computational amount. An assimilation cycle results in updated SAC states valid at the end of the assimilation window, or the prediction time, and updated TCIs valid over the assimilation window (see Figure 3.1). In the forecast generation process, the DA solution is used only for the initial conditions of the SAC and the TCIs valid over the assimilation window whereas all other control variables assume their default values. Accordingly, PXADJ and PEADJ are set to unity and the additive error to TCI is to zero for simulation over the forecast horizon.

For assimilation and validation, we used the US Geological Survey hourly streamflow data at the outlets of all study basins. Our focus in this work is on flood forecasting. As such, we identified only the significant events from the observed streamflow data and used the resulting partial duration time series, where a significant event is defined as a hydrograph whose peak exceeds 100 m^3/s. Once a significant event is identified, the beginning and ending hours of that event are determined by subtracting and adding 72 hours from and to the first and last hours of the flow duration over 100 m^3/s, respectively. In this way, the comparative evaluation is focused on the performance of DA for relatively large events only, for which streamflow DA is likely to be affected by significant CB in the state variables. The MAP data used were operationally produced by the West Gulf River Forecast Center using the Multisensor Precipitation Estimator (Seo et al., 2010). For MAPE, the monthly climatology of daily estimates was used. Because the DA cycles consist only of the periods of significant flow, PE does not play a significant role, and hence the lack of dynamically-varying MAPE is unlikely to impact the outcome in any significant way.

Because both the CBEnKF and the EnKF solve unconstrained minimization, there is no guarantee a priori that the updated states of all members lie within the feasible region of the SAC states. Various approaches have been reported in the literature to solve constrained KF and EnKF (Gupta and Hauser, 2007; Wang et al., 2009). In this work, we use the trial-and-error approach of Wang et al. (2009); if not all updated ensemble members are in the feasible region, we additionally sample random perturbations for the infeasible members and update the states until all members are feasible. The above strategy was used for both the CBEnKF and the EnKF in an identical manner. To initiate ensemble DA, each SAC state is perturbed by sampling from a lognormal distribution with mean and standard deviation given by the current SAC state and approximately 17% of the square root of the maximum soil water content, respectively. The perturbations for PXADJ and PEADJ are sampled from a lognormal distribution with mean and standard deviation given by the current state and 0.17, respectively. The perturbations for the additive error of the TCI are sampled from a normal distribution with the same mean and standard deviation (in mm/hr) as above. The model states are randomly perturbed independently of one another for simplicity. Similarly, the zero-mean dynamical model errors, W_{k-1} in Eq. (3-1), are assumed to be independent of one another and have the standard deviation of approximately 4% of the current state for the SAC states and 0.03 for all other variables. The ensemble size is 30 throughout this work. The above setup was identical for both the EnKF and the CBEnKF. For additional details on perturbation and ensemble generation, the reader is referred to Rafieeinasab et al. (2014).

3.4 Study basins, evaluation metrics and observational uncertainty modelling

In this section, we describe the study basins, evaluation metrics and observational uncertainty modelling.

3.4.1 Study basins

Ten hydroclimatologically diverse headwater basins in Texas (U.S.) are used (see Figure 3-2). These catchments are a subset of the basins identified as particularly challenging in operational forecasting by the NWS West Gulf River Forecast Center. The basins have been used by the authors and their collaborators for hydrologic DA and related research in recent years (Seo et al., 2009; Lee et al., 2011, 2012, 2016; Lee and Seo, 2014; Rafieeinasab et al., 2014b). The basins are wetter in the east and drier in the west as indicated by the dryness index calculated for the period of 1996 to 2005 (see Table 3-1) where the dryness index is defined as the ratio of mean annual PE to mean annual precipitation (Budyko, 1974). Table 3-1 also presents the time-to-peak (T_P) and the 95th percentile observed flow within the selected events ($Q_{95|S}$). Compared to the basins in the humid region, those in the semi-arid region pose larger challenges in streamflow forecasting due to the generally quicker catchment response, larger uncertainties in the initial conditions, possible structural uncertainties associated with unmodeled processes such as infiltration into dry channel beds, and larger parametric uncertainties due to the above factors and the limited number of large events available for calibration. In addition, the large spatio-temporal variability of precipitation

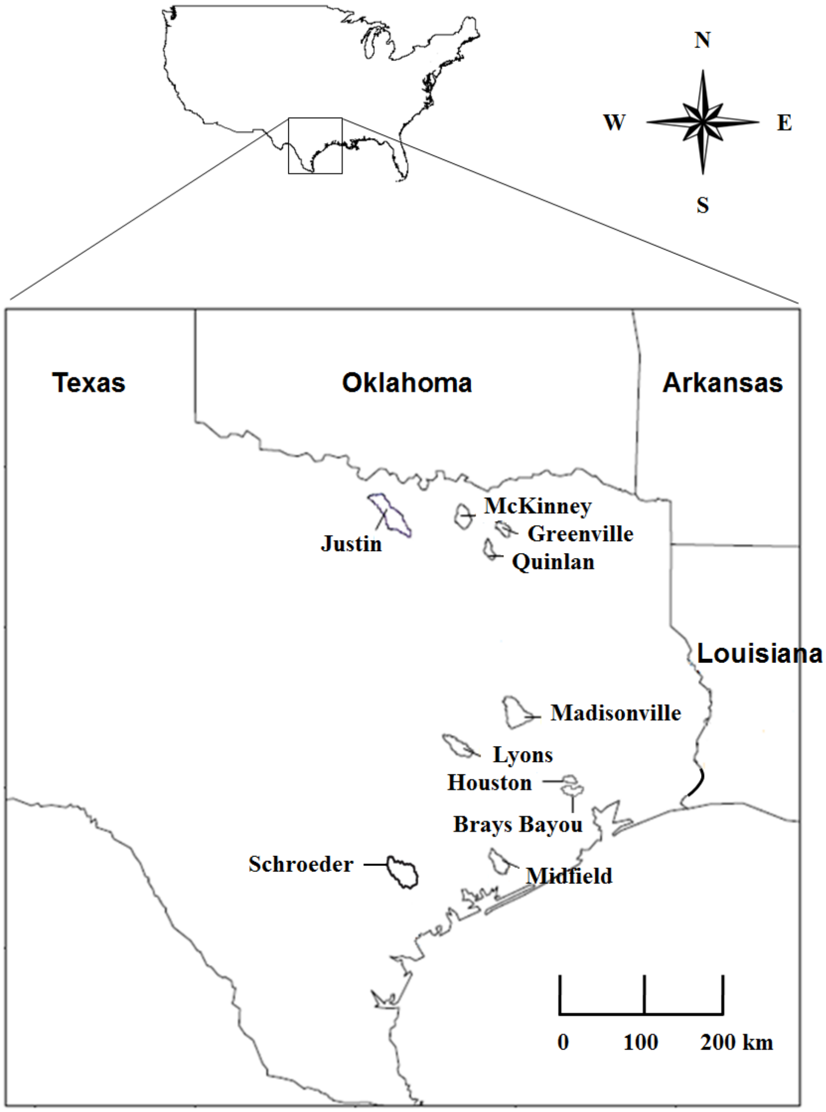


Figure 3-2 Study basins in Texas, USA

from highly convective storms adds large uncertainties in lumped modelling particularly for large basins. As such, the study basins offer an extremely challenging test for streamflow DA.

3.4.2 Evaluation metrics

To evaluate the comparative performance of the CBEnKF with the EnKF, we verified both ensemble mean and ensemble streamflow forecasts from hindcasting experiments. For ensemble mean forecasts, we used the root mean square error, maximum errors of under- and over-prediction, correlation coefficient and mean square error skill score (MSE-SS)(Lee et al., 2011, 2012;

Lee and Seo, 2014). The results indicate that the MSE-SS represents the overall comparative performance very well and we hence focus below on the MSE-SS results. The MSE-SS for the CBEnKF ensemble mean in reference to the EnKF ensemble mean is given by:

$$\text{MSE-SS} = 1 - \frac{\frac{1}{N}\sum_{k=1}^{N}\left(\overline{Q_k}^{CBEnKF} - Z_k^Q\right)^2}{\frac{1}{N}\sum_{k=1}^{N}\left(\overline{Q_k}^{EnKF} - Z_k^Q\right)^2} \tag{3.15}$$

where $\overline{Q_k}^{CBEnKF}$, $\overline{Q}_k^{EnKF}$, Z_k^Q, and N denote the CBEnKF ensemble mean flow, EnKF ensemble mean flow, verifying observed flow, and the number of paired sets of Z_k^Q and $\overline{Q_k}^{CBEnKF}$, respectively. A perfect ensemble mean forecast would yield an MSE-SS of unity. No improvement in the ensemble mean forecast by the CBEnKF over the EnKF would yield an MSE-SS of zero. For ensemble forecasts, we used a large number of verification measures available in the Ensemble Verification System (EVS)(Brown et al., 2010), including the Brier score, Brier skill score, continuous ranked probability score (CRPS), CRPS decomposition (Hersbach, 2000), continuous ranked probability skill score (CRPSS)(Brown et al., 2014), reliability diagram, relative operating characteristic and others. The results indicate that the mean CRPS, its decomposition, and mean CRPSS represent the comparative performance very well and hence we focus on their results in this paper. The CRPS measures the integral squared difference between the cumulative distribution function (CDF) of ensemble forecast and that of the verifying observation. A more skilful ensemble forecast will produce a smaller CRPS than the reference ensemble forecast. The mean CRPSS of the CBEnKF-based ensemble forecast in reference to the EnKF-based is given by:

$$\overline{CRPSS} = 1 - \frac{\overline{CRPS}_{CBEnKF}}{\overline{CRPS}_{EnKF}} = 1 - \frac{\frac{1}{N}\sum_{i=1}^{N}\int_0^{\infty}\left(F_i^{CBEnKF}(Q) - F_i^{obs}(Q)\right)^2 dQ}{\frac{1}{N}\sum_{i=1}^{N}\int_0^{\infty}\left(F_i^{EnKF}(Q) - F_i^{obs}(Q)\right)^2 dQ} \tag{3.16}$$

where the $\overline{CRPS}_{CBEnKF}$ and $\overline{CRPS}_{EnKF}$ denote the mean CRPS of the CBEnKF and the EnKF ensemble forecast, respectively, $F_i^{CBEnKF}(Q)$ and $F_i^{EnKF}(Q)$ denote the CDFs of the i^{th} CBEnKF and EnKF ensemble forecast, respectively, and $F_i^{obs}(Q)$ denotes the CDF of the verifying observation (i.e., the Heaviside step function). A perfect ensemble forecast would yield a CRPSS of unity. No improvement by the CBEnKF over the EnKF ensemble forecast would yield a CRPSS of zero. Following Hersbach (2000), the CRPS can be decomposed into reliability (REL), resolution (RES), and uncertainty (UNC), or into REL and potential CRPS ($CRPS_{POT}$):

$$CRPS = REL - RES + UNC = REL + CRPS_{POT} \tag{3.17}$$

Smaller REL indicates more reliable ensembles (desirable) and larger RES means better resolution (desirable). The RES component (=UNC – $CRPS_{POT}$) is positive if the ensemble forecast is better than the climatological ensemble forecast (Hersbach, 2000). The UNC component reflects climatologic uncertainties in the observations and does not relate to forecast attributes. The $CRPS_{POT}$ represents the CRPS achievable by calibrating forecast probabilities for perfect reliability (Hersbach, 2000). Similarly to the CRPS, smaller $CRPS_{POT}$ indicates improved performance by the subject ensemble forecast.

3.4.3 Observational uncertainty modelling

Observational errors are generally heteroscedastic and correlated in time (Lee et al., 2011). Due to lack of observed truth, however, their modeling is a large challenge. In this work, we assume that all observational errors are temporally uncorrelated and homoscedastic, and use a wide range of settings for their variances. In this way, one may assess the comparative performance of the CBEnKF with the EnKF under widely varying, but identical conditions of the quality of the observations being assimilated and that of the model used. Generally speaking, one would expect the observation error variance for MAPE, σ_E^2, to be smaller than that for MAP, σ_P^2. Given that we are using climatological MAPE rather than dynamically estimated values, however, we assumed $\sigma_E^2 = \sigma_P^2$ for simplicity. We then used the settings of 10, 100 and 1000 (mm/hr)^2 for $\sigma_E^2(=\sigma_P^2)$ to represent small, medium and large uncertainties, respectively, in input forcings. The same three settings were used for the error variance to represent small, medium and large model uncertainties in the SAC-UH. Table 3-3 lists the resulting 9 possible combinations of σ_P^2 and σ_R^2, represented by C1 through C9.

Harmel et al. (2006) report streamflow measurement errors for small watersheds of 42%, 19%, 10%, 6% and 3% for the worst case, typical maximum, typical average, typical minimum, and the best case scenarios, respectively. Baldassarre and Montanari (2009) report that the overall error affecting river discharge observations ranges from 6.2% to 42.8%, at the 95% confidence level, with an average value of 25.6%. Given the dependence of streamflow observation errors on the flow, one may consider prescribing larger or smaller σ_Q^2 depending on the magnitude of the observed flow. The error associated with Z_k^Q in Z_k in Eq.(3.2), however, reflects not only the observational error of streamflow but also the modeling errors associated with the simulated flow in

Table 3-3 Observational error variances used for different basins

Case	$\boldsymbol{\sigma_P^2}$ **(mm/hr)^2**	$\boldsymbol{\sigma_R^2}$ **(mm/hr)^2**	$\boldsymbol{\sigma_Q^2}$ **(m^3/s)^2**	
C1	10	10	100 for Justin, Houston, Greenville, Brays Bayou, Midfield	1000 for Lyons, McKinney, Madisonville, Quinlan, Schroeder
C2	10	100		
C3	10	1000		
C4	100	10		
C5	100	100		
C6	100	1000		
C7	1000	10		
C8	1000	100		
C9	1000	1000		

G(X_k). Note that, whereas the DA formulation addresses uncertain SAC states, biases in MAP and MAPE, and errors in the TCI, it does not address errors arising from model deficiencies such as timing errors (Liu et al., 2011). As such, a larger σ_Q^2 would be more appropriate if the model has large simulation errors which the DA procedure cannot address even though the basin produces smaller peak flows. For this reason, a priori determination of σ_Q^2 is not readily possible and some form of sensitivity analysis is usually necessary. In this work, we used σ_Q^2 of 100 and 1,000 $(m^3/s)^2$, carried out assimilation experiments using both settings of σ_Q^2, and compare the relative performance for each basin. To illustrate the above process, we show in Figure 3-3 the MSE-SS results for Midfield for C1 through C9 in Table 3-3 for σ_Q^2 of 100 and 1,000 $(m^3/s)^2$. The figure shows that

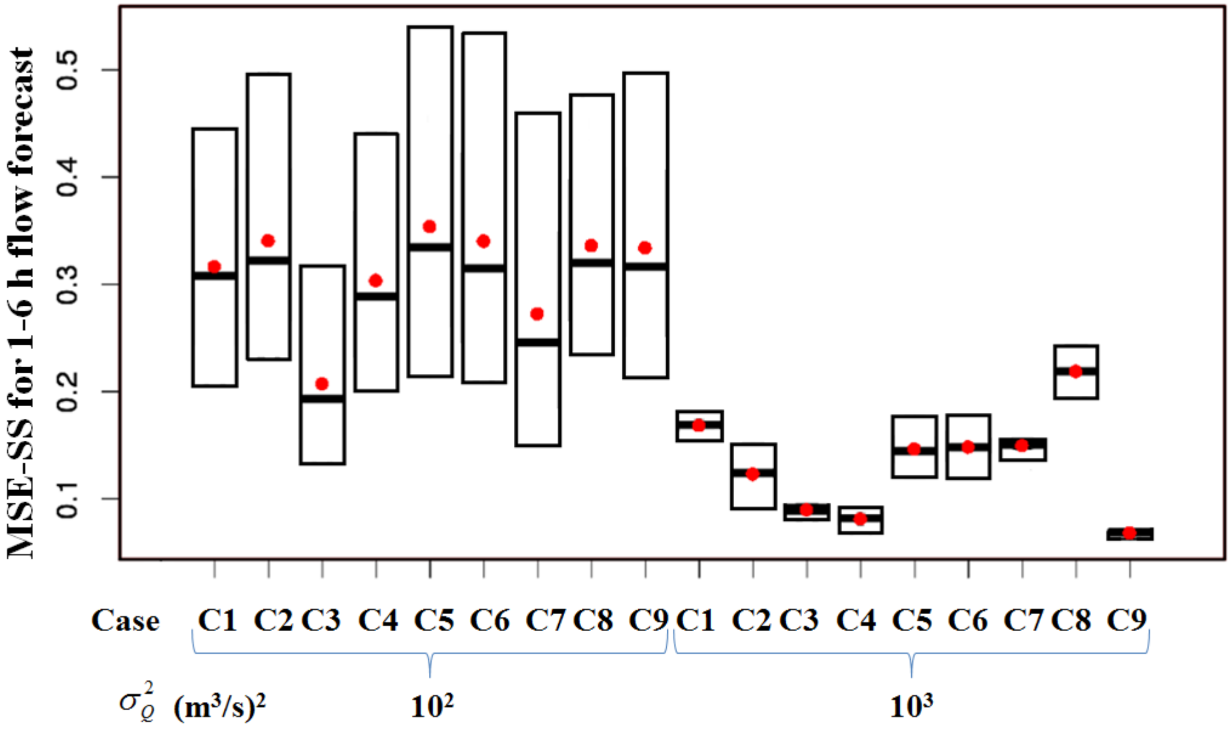


Figure 3-3 Box plots of the MSE-SS of the CBEnKF 1-6 h ensemble mean flow forecast over the EnKF for Midfield for the 9 different combinations of σ_P^2 and σ_R^2, and two different values of σ_Q^2 in Table 3-3. A total of 6 data points are used at each box where each data point is from each of the first 6 lead hour ensemble mean flow. The top and bottom of the boxes are associated with the 1 and 6 hr-ahead predictions, respectively. Red dots and the bars in the middle of boxes represent the mean and the median, respectively.

the CBEnKF improves over the EnKF in the ensemble mean sense for all cases, but that the margin of improvement is larger with $\sigma_Q^2 = 100$ (m^3/s)^2. Similar results for all other basins suggest σ_Q^2 of 100 (m^3/s)^2 for Justin, Houston, Greenville and Brays Bayou and σ_Q^2 of 1,000 (m^3/s)^2 for Lyons, McKinney, Madisonville, Quinlan and Schroeder. In terms of the observed 95th percentile flow, Schroeder, Madisonville, Midfield and Justin produce the largest peak flows whereas Greenville, Houston and Brays Bayou produce the smallest. Comparison of the basin groupings based on the choice of σ_Q^2 and those based on the 95th-percentile flow suggests that the smaller σ_Q^2 of 100 (m^3/s)^2 might not be appropriate for Justin and Midfield which produce the largest peak flows. Examination of the hydrographs indicated, however, that the above two basins have significantly smaller timing errors (Liu et al., 2011) which makes the smaller σ_Q^2 more realistic. In the opposite case, Quinlan has significantly larger timing errors though its 95th-percentile flow is not very large which makes the larger σ_Q^2 of 1,000 (m^3/s)^2 more appropriate. With the above choices, we arrive at the nine cases for each of the two groups of $\sigma_Q^2 = 100$ (m^3/s)^2 and $\sigma_Q^2 = 1{,}000$ (m^3/s)^2 as shown in Table 3-3. It is worth noting that the resulting upper bounds for the observation error variances are at least an order of magnitude larger than those used in Rafieeinasab et al. (2014). The rationale for the larger bounds is to use uncertainty estimates that are realistic for large events and to assess the relative performance between the CBEnKF and the EnKF under the limiting conditions of very large and very small observational and modeling uncertainties.

3.5 Results and discussion

In this section, we present the comparative evaluation results. For each basin, the CBEnKF and EnKF were applied for prediction of all events whose peak flow exceeds 100 m^3/s within the 10-yr period of record. We first present illustrative examples of the CBEnKF and EnKF results in

Subsection 3.5.1. We then describe comparative verification of ensemble mean and ensemble predictions in Subsections 3.5.2 and 3.5.3, respectively.

3.5.1 Illustrative examples

To help intuitive understanding of how the CBEnKF and EnKF predictions may compare, we first present in Figure 3-4 illustrative examples of predicted hydrographs for selected events only for a subset of the basins due to space limitations. In the figure, it is readily seen that the CBEnKF ensembles are of higher quality than the EnKF in terms of both ensemble mean and spread. In general, the CBEnKF ensembles are able to encompass the observed peak flow whereas the EnKF ensembles are not. Examination of the hydrographs for all basins may be summarized as follows. For Greenville and Quinlan, the overall shape of the simulated hydrograph indicated the presence of timing errors. For McKinney and Midfield, it was observed that the EnKF ensembles lack spread in the recession phase. The above suggests that one or more SAC states for the soil water contents may have been completely full, thereby suppressing the ensemble spread. The above observations point out the need for parametric uncertainty modeling to mitigate the negative impact of timing errors and to account for the uncertainties in the soil water holding capacities of the soil moisture accounting model.

Figure 3-5 shows an example set of ensemble mean results for a single event at Lyons. The figure shows the open-loop, and DA-updated streamflow, UZTWC, UZFWC, LZTWC, LZFSC, LZFPC, precipitation, PE and the TCI. They illustrate how the CBEnKF- and EnKF-updated states may compare during the course of an event. For streamflow, the verifying observations are also shown. In the figure, each time step represents the ending hour of the assimilation window, and the associated results represent the updated states and the model-simulated flow valid at that hour (i.e., the prediction time). The figure hence represents the time series of the analysis results from

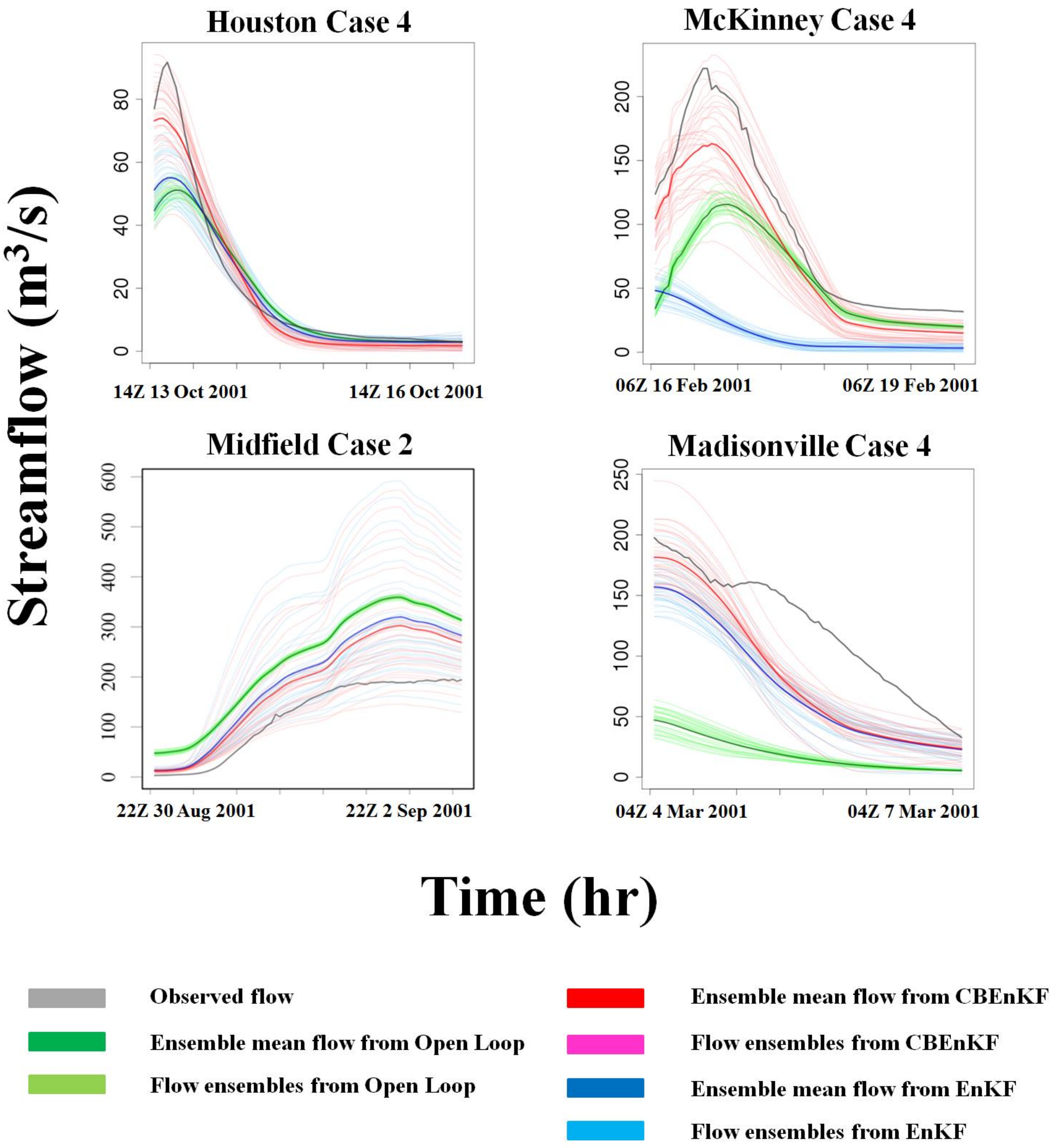


Figure 3-4 Example ensemble streamflow forecasts for Houston, McKinney, Midfield and Madisonville.

successive DA cycles during the course of the event. The observation uncertainty parameters associated with Figure 3-5 are $\sigma_Q^2 = 1{,}000$ (m^3/s)^2, and $\sigma_P^2 = \sigma_E^2 = \sigma_R^2 = 1{,}000$ (mm/h)^2. It is seen that both the CBEnKF and EnKF streamflow analyses are much closer to the observed than the open-loop analysis, and that the CBEnKF streamflow analysis tracks the observed more closely than the

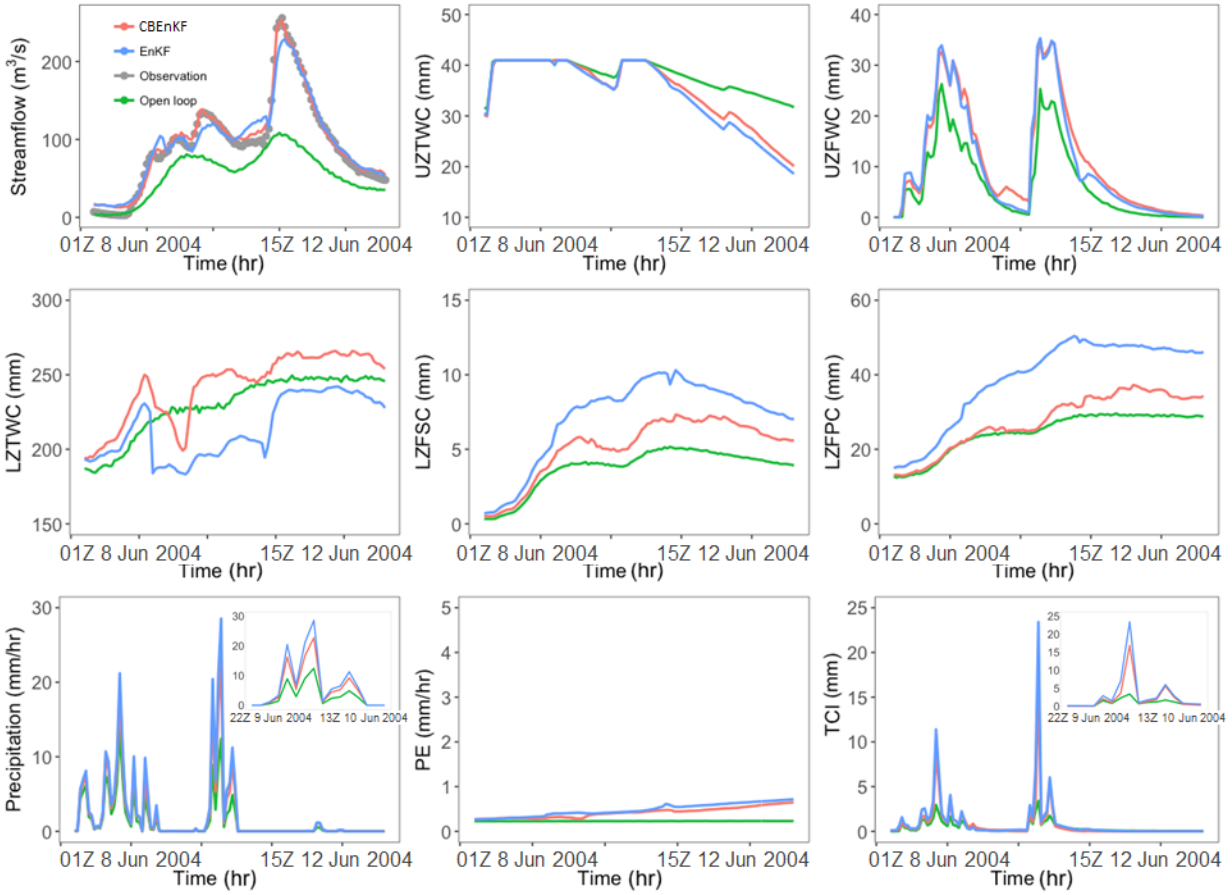


Figure 3-5 Ensemble mean analyses valid at each hour of a single event in Case 9 (C9) for Lyons.

EnKF (upper left). The corresponding CBEnKF and EnKF analyses of the state variables may at first appear similar, but significant differences exist in the lower zone soil moisture states throughout the event as well as in the UZFWC and the additive error in the TCI at certain DA cycles. Because there are no truths available, it is not possible to assess the accuracy of the state variables. One may nevertheless observe that, compared to the EnKF, the CBEnKF solutions are closer to those of the open loop simulation especially in the three lower zone soil moisture states of LZTWC, LZFSC and LZFPC, precipitation, and the TCI while producing more accurate streamflow analysis. Visual examination of many results similar to Figure 3-5 suggests that the CBEnKF- and EnKF-updated states differ by varying degrees among different combinations of the observational uncertainty parameters and different state variables, but that, in general, the CBEnKF-updated states tend to maintain the joint relationships among the different states of the open-loop simulation more closely than the EnKF-updated.

To illustrate the above point, we show in Figure 3-6 the scatter plots between the UZFWC and LZFPC, the fastest and the slowest varying SAC states, respectively, from the open-loop, EnKF-updated and CBEnKF-updated ensemble results. In the top left plot of Figure 6, each "stripe" of data points represents an open-loop simulation associated with the peak flow portion of the hydrograph associated with a significant event (see the bottom plot in Figure 3-6). The herringbone-like patterns are due to the ensemble members from the same assimilation cycle possessing similar updated states. Figure 6 indicates that the CBEnKF-updated states resemble the joint relationships

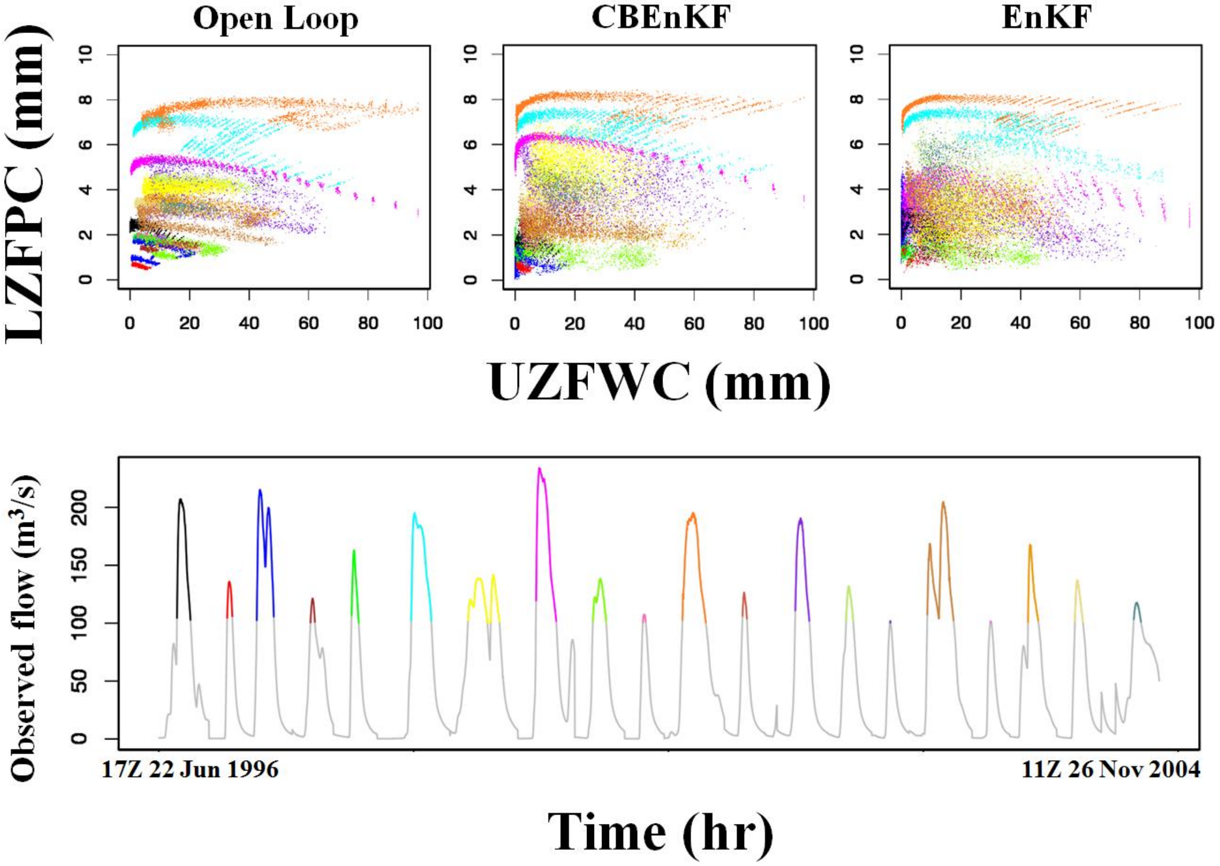


Figure 3-6 (Upper) Scatter plots of UZFWC vs. LZFPC analysis ensembles valid at prediction time for Case 1 (C1) of Midfield and (Lower) the corresponding observed flow time series. The same colors between the scatter and the time series plots represent the same events. Only those portions of $Z^Q \geq 100$ (m^3/s) are shown for clarity where Z^Q denotes observed streamflow.

between the UZFWC and LZFPC in the open-loop states more closely than the EnKF-updated even though the locations of the stripes are quite different among the three results in the bivariate phase space. Observations from Figs 3-5 and 3-6 and many similar results suggest that the CBEnKF retains the open-loop model dynamics among the different soil moisture states more closely than the EnKF, and that, when both the CBEnKF and EnKF produce streamflow analyses of similar quality, the CBEnKF-updated states stay closer to the open-loop states than the EnKF states. This supports the argument that the improved skill by the CBEnKF is due in part to the updated states being more in line with the open-loop model dynamics. Similar results were also obtained by Noh et al. (2018) in which a multi-scale version of the fixed-lag smoother (see Figure 3-1) was compared for streamflow DA with the single-scale version as implemented in this work but using variational assimilation rather than the EnKF or CBEnKF. They showed that minimizing adjustment to the model states while producing high-quality streamflow analyses significantly improves predictive skill in DA solutions albeit at a substantially higher computational cost.

3.5.2 Ensemble mean streamflow evaluation

Here we present the scatter plots and MSE-SS results to assess the performance of ensemble mean predictions. Figure 3-7 shows the scatter and quantile-quantile plots of 3 hr-ahead ensemble mean streamflow forecast vs. verifying observation. Also shown in the figure are the sample CBs computed over the 10 sub-ranges that equally divide the entire range of observed flow. It is seen that the CBEnKF further reduces CBs in the ensemble mean sense beyond what the EnKF is able to for the entire range of flow. Table 4 summarizes the mean absolute CB and RMSE of 3 hr-ahead ensemble mean streamflow forecasts shown in Figure 3-7. The table shows that the CBEnKF generally reduces both the mean absolute CB and RMSE further than the EnKF. For Justin and Mid-

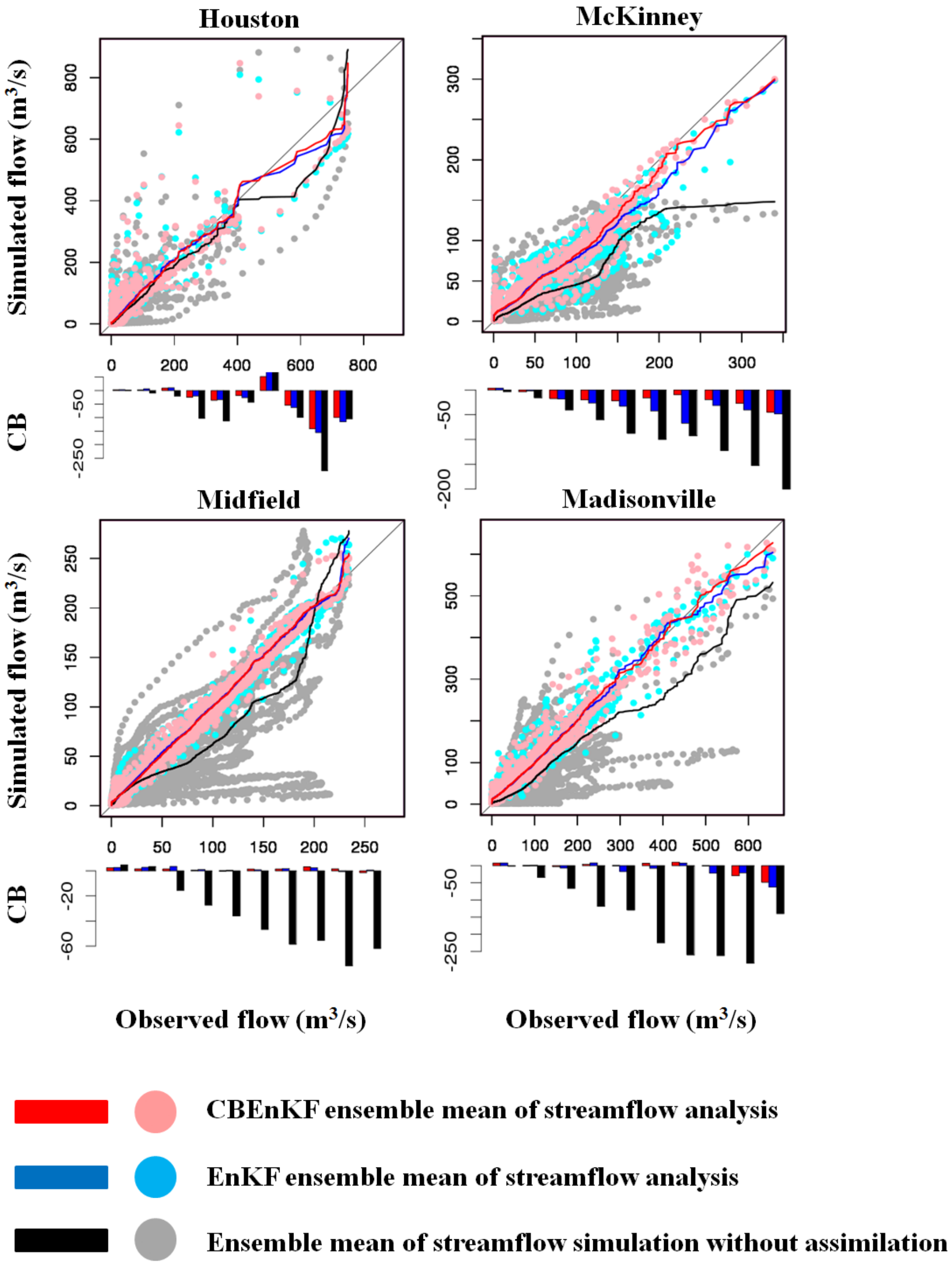
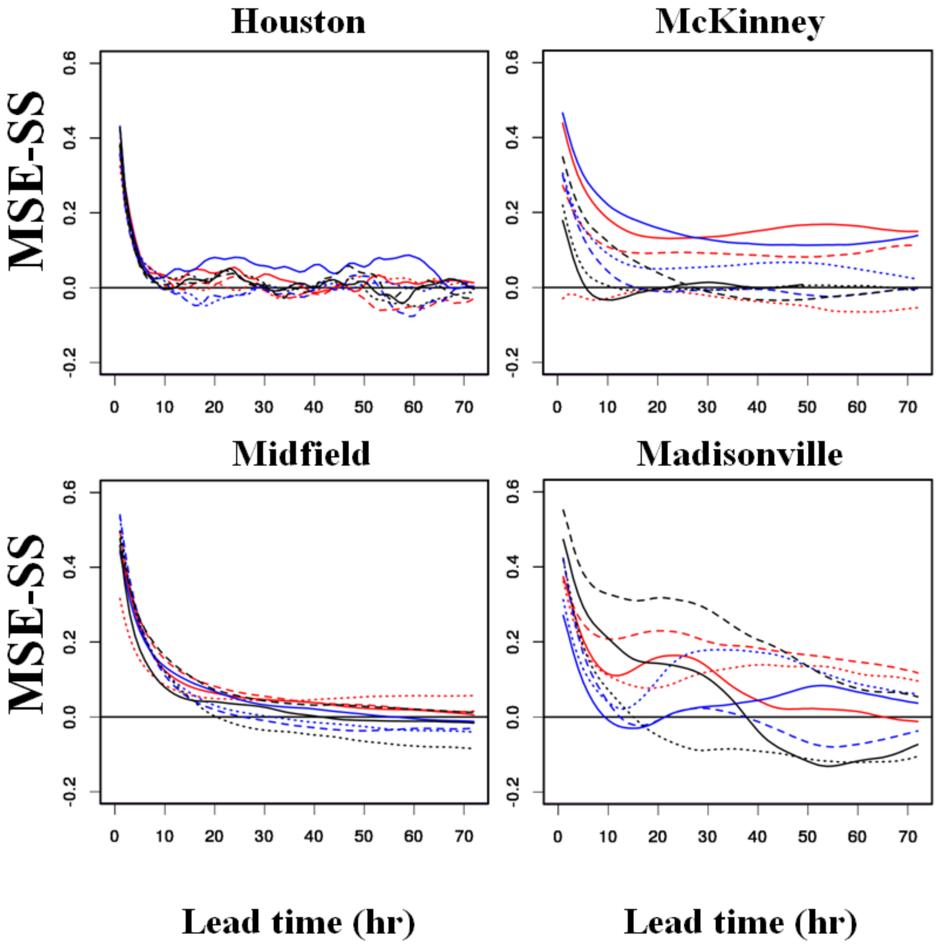


Figure 3-7 Scatter and quantile-quantile plots of ensemble mean simulated flow vs. observed at lead time of 3 hrs. The simulations are from the open loop (blue line), EnKF (green line) and CBEnKF (red line) runs for significant events. The bar plots show the CB calculated as the mean difference between simulated and observed flows for each of the ten equally divided flow ranges. For each basin, a single case from the nine cases of C1 to C9 in Table 3 is shown, chosen based on the MSE-SS results in Figure 3-8. The selected cases are C4, C4, C2 and C8 for Houston, McKinney, Midfield, Madisonville, respectively.

field, the reduction in both the mean absolute CB and RMSE by the EnKF over open-loop simulation is already very large, and that for the CBEnKF is similarly large. Figure 3-8 shows the MSE-



	CASE	σ_P^2 (mm/hr)²	σ_R^2 (mm/hr)²
———	C1	10	10
— — —	C2	10	100
- - - - -	C3	10	1000
———	C4	100	10
— — —	C5	100	100
- - - - -	C6	100	1000
———	C7	1000	10
— — —	C8	1000	100
- - - - -	C9	1000	1000

Figure 3-8 MSE-SS of the CBEnKF ensemble mean streamflow forecast over the EnKF as a function of lead time

SS of the CBEnKF ensemble mean forecast for a subset of the basins for all verifying observed flow within the periods of significant events as a function of lead time. The reference forecast is the EnKF ensemble mean forecast. The positive MSE-SS means that the CBEnKF outperforms the EnKF in the ensemble mean sense. In each plot, there are 9 MSE-SS curves representing the 9 different cases in Table 3-3. The multi-basin mean for all 10 basins of the MSE-SS of the CBEnKF over the EnKF is 0.37 for lead time of 1 hr, and 0.15 over lead times of up to the time-to-peak of the basin. For Houston and Brays Bayou, the two most quickly-responding small urban basins in the humid region, the positive impact of the CBEnKF over the EnKF diminishes rather quickly due to the short basin memory and relatively small CB in the base simulation (for Houston, see also Figure 3-7). For larger basins, the positive impact extends to longer lead times. For Lyons, McKinney and Madisonville, the increased skill in the CBEnKF ensemble mean forecast remains significantly high well beyond the first 24 hrs of lead time for a number of observational uncertainty parameter combinations. For the largest basin, Justin, on the other hand, the skill dissipates rather quickly due to the short time-to-peak associated with rain falling in the lower part of this very elongated basin which dominates the generation of large flows at the outlet (see Figure B1 of Seo et al., 2009). Figure 3-8 also shows that the CBEnKF improves over the EnKF in the ensemble mean sense under widely varying combinations of σ_P^2 ($=\sigma_E^2$) and σ_R^2. The above observations indicate that the CBEnKF is robust and may be expected to perform well even if the uncertainty parameters are not prescribed with accuracy, an important consideration for real-world applications.

3.5.3 Ensemble streamflow evaluation

Here we present the ensemble verification results to assess the performance of ensemble predictions. Figure 3-9 shows the mean CRPSS of the CBEnKF ensemble forecast for a subset of the basins. The reference is the EnKF ensemble forecast. All results in Figure 3-9 are conditional on the verifying observation exceeding the 95th-percentile flow of the associated basin within the significant events (see Table 3-1). As such, the figure assesses comparative skill specifically for prediction of large flows for which there exists only limited hydrologic memory for DA to rely on.

For each basin, there are 9 mean CRPSS curves corresponding to the 9 combinations of σ_P^2 (=σ_E^2

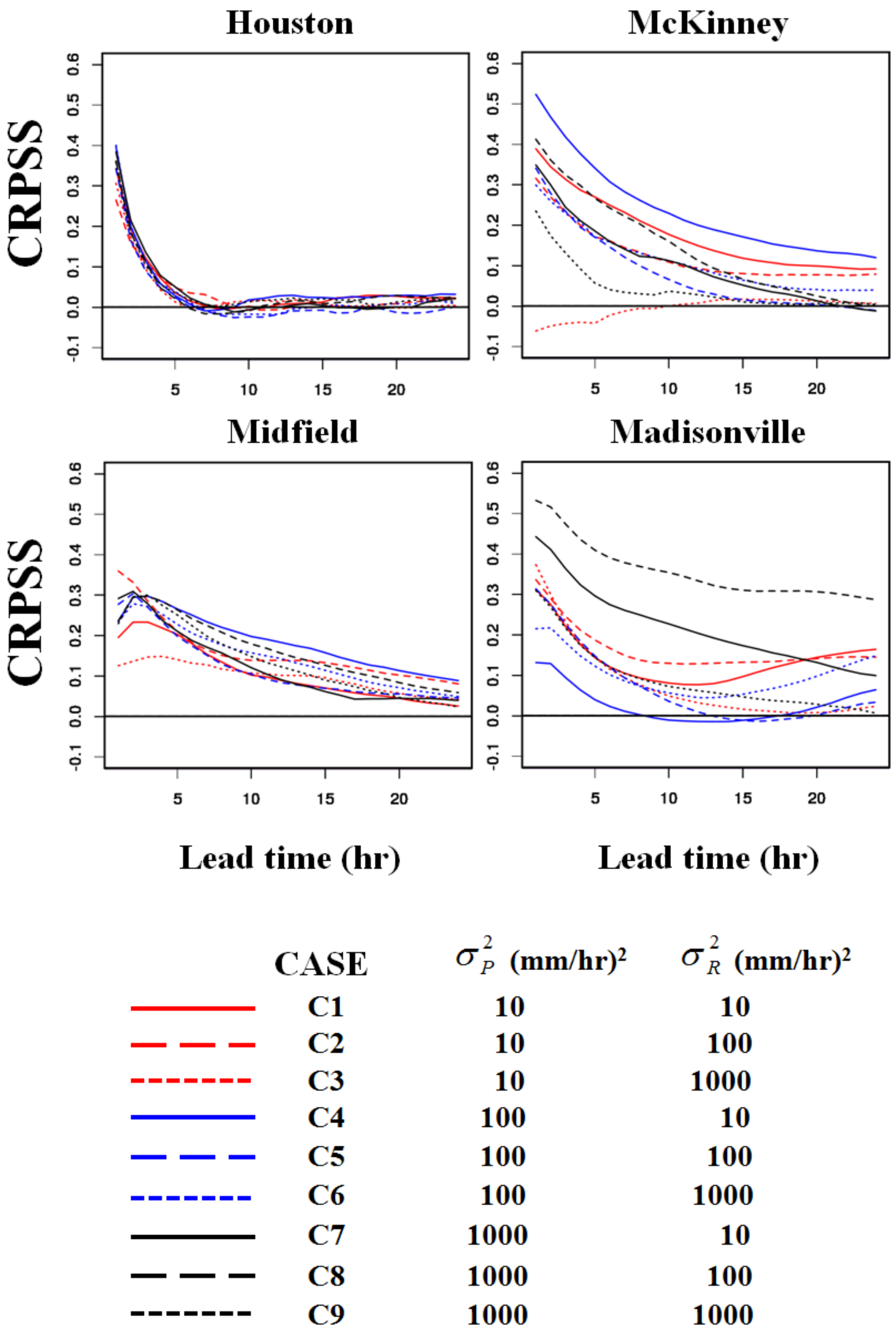


Figure 3-9 Mean CRPSS vs. lead time of the CBEnKF ensemble streamflow forecast conditional on the observed flow exceeding the 95th percentile within significant events. The reference is the EnKF ensemble forecast.

) and σ_R^2 given the basin-specific σ_Q^2 (see Table 3-3). Figure 3-9 shows that the CBEnKF considerably improves over the EnKF for all basins shown for short lead times. As with the MSE-SS, the increased skill quickly dissipates for Houston and Brays Bayou. For the larger basins, Lyons, McKinney and Madisonville, the increase skill persists into significantly longer lead times. For the rest of the basins, Justin, Greenville, Midfield, Quinlan and Schroeder, the increased skill stays significant for the first several hours or more but dissipates within the first 24 hours or so. It is observed in Figure 3-9 that, for some mean CRPSS curves for Midfield, the skill is lower at lead time of zero (i.e., analysis) than at lead time of 1 hr. Examination of the comprehensive verification results from the EVS indicates that, for large basins such as Midfield, the analysis ensembles are not always as reliable as the 1-hr forecast ensembles. The above observations point out the need for improving uncertainty modeling particularly for large basins, a topic left as a future endeavor.

For attribution of the increased skill in CBEnKF ensembles, we carried out mean CRPS decomposition for 1-6 hr-ahead streamflow ensemble forecasts. Figure 3-10 shows the mean CRPSS, CRPS, REL, RES and CRPS_{POT} for all basins. For each basin, there are 9 data points representing the 9 combinations of σ_P^2 ($=\sigma_E^2$) and σ_R^2 given the basin-specific σ_Q^2. In the figure, the results are connected across all basins only to aid visualization of the differences between the CBEnKF and EnKF results. The figure shows that the CBEnKF generally improves both reliability and resolution of ensemble forecasts over the EnKF for all basins. The mean CRPSS varies from 0.06 to 0.19 with a mean of 0.12. The mean improvement for all basins in the REL, RES, CRPS_{POT} and CRPS by the CBEnKF over the EnKF are 17, 7, 10 and 11%, respectively. Because the magnitude of the

RES is generally much larger than that of the REL at short lead time, the above reduction in the

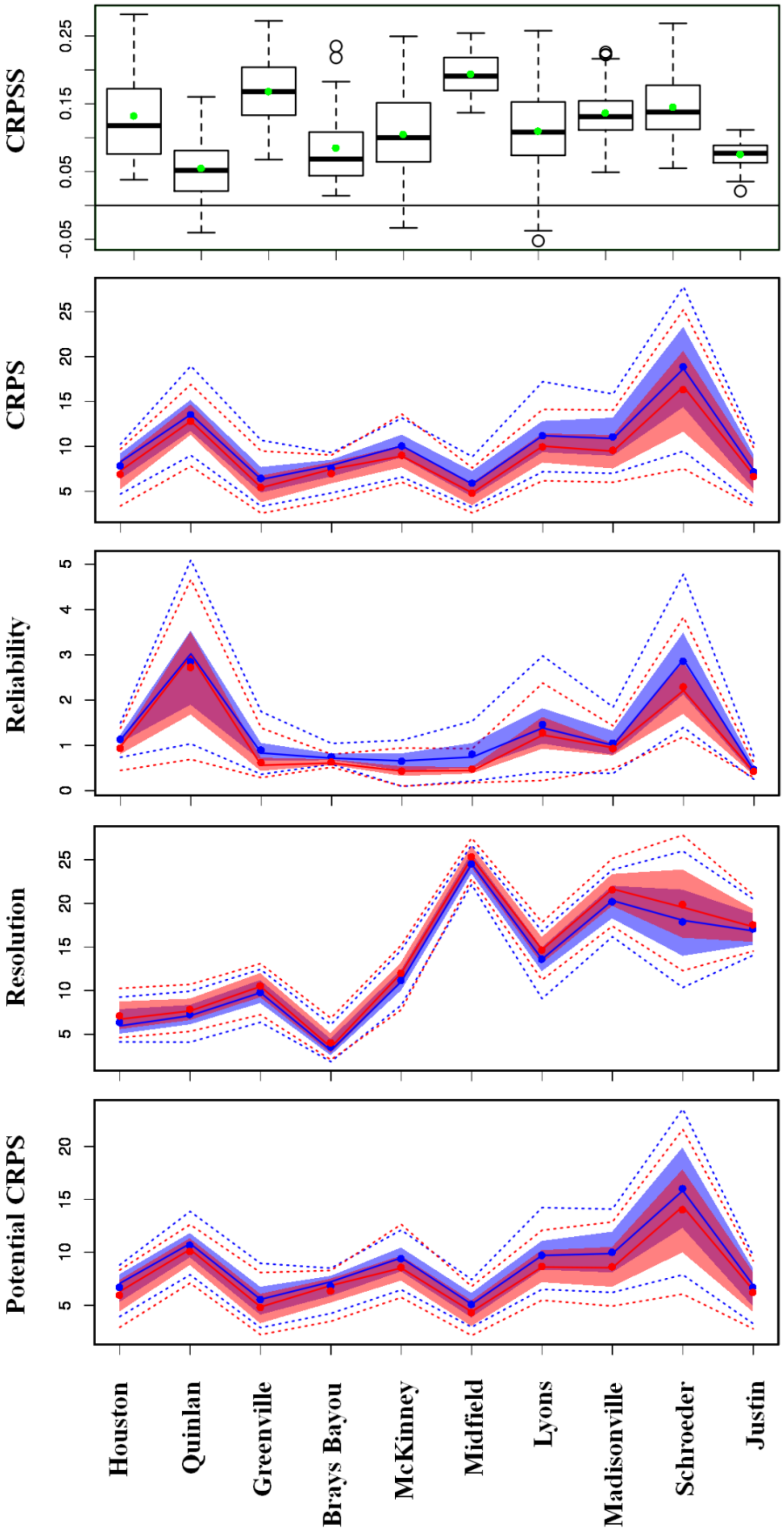


Figure 3-10 (From top to bottom) Mean CRPSS of the CBEnKF 1-6h flow forecast over the EnKF and corresponding mean CRPS, REL, RES and CRPS$_{POT}$. Red and blue colors are used for the CBEnKF and EnKF, respectively. Shaded areas represent 5-95th percentile ranges. Lower and upper dotted lines denote minimum and maximum, respectively. Solid lines and dots denote the median and mean, respectively.

REL and increase in the RES are comparable in magnitude in terms of reducing the mean CRPS of the CBEnKF ensembles. Figure 3-11 shows the mean CRPSS of 1 hr-ahead CBEnKF ensemble forecast in reference to the 1 hr-ahead EnKF forecast for all basins as a function of the conditioning threshold. For each threshold on the x-axis, a total of 90 data points are represented in the vertical which include 9 cases for each of the 10 basins. The figure shows that, the larger the verifying flow is, the larger the margin of improvement by the CBEnKF is. The mean CRPSS of the CBEnKF ensembles in reference to the EnKF is about 0.2 for all ranges of verifying flow and increases to about 0.3 or larger for observed flows at or above the 95th percentile flow within the significant events (see Table 3-1). That the gain in skill increases as the threshold increases is an extremely appealing attribute of the CBEnKF and is consistent with the CBPKF results (Seo et al., 2018a,b; Shen et al. 2019). The above indicates that SAC-UH simulations of floods generally contain significant CB for the study area, and that the CBEnKF is able to reduce its ill effects and significantly improve prediction.

The results presented above pertain only to prediction of significant events. When the assimilation cycles include both the storm and inter-storm periods, the skill of the CBEnKF predictions

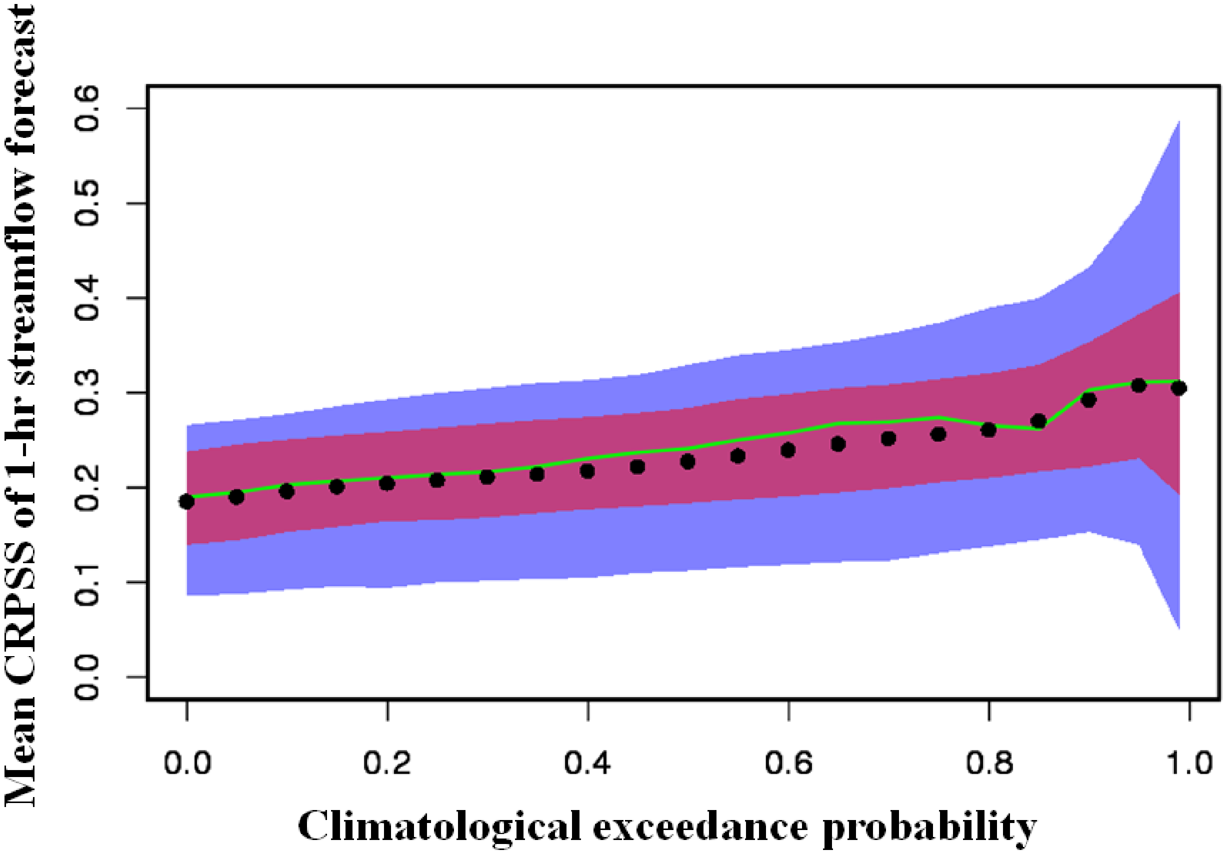


Figure 3-11 Mean CRPSS of the CBEnKF 1-hr streamflow forecast over the EnKF for all basins as a function of the conditioning verifying observed flow exceeding the climatological probability shown on the x-axis. The probability levels used range from 0 to 0.99. At each probability level, there are 90 values (=10 basins× 9 cases) from which the 5, 25, 50, 75, 95th percentiles and the mean value were calculated. Blue and red shaded areas represent the 5-95th and 25-75th percentile ranges, respectively. Green line and black dot denote the median and mean, respectively.

in reference to the EnKF is expectedly smaller. For example, for C5 and C9 at Madisonville, the mean CRPSS for flood periods only and for all periods are about 0.35 and 0.20, respectively. When the assimilation cycles include only the inter-storm periods, the CBEnKF shows little improvement over the EnKF. When only the periods of extremely low flows are included, however, the CBEnKF improves over the EnKF. The above results are fully expected in that the EnKF already provides the optimal solution when the true state of the system is at or near median, which corresponds generally to low flow conditions, and that the CBEnKF provides improvement over the tails of the distribution (Seo et al., 2018a, 2018b; Shen et al., 2019). The above observations suggest that, in real-time operation, one may normally run the EnKF and supersede it with the CBEnKF only in non-normal conditions.

3.6 Conclusions and future research recommendations

The conditional bias-penalized ensemble Kalman filter (CBEnKF) is developed and evaluated for real-time assimilation of streamflow data to improve flood forecasting. The ensemble mean results show that, for the ten basins in TX, the CBEnKF improves the mean square error skill score (see Eq. (3.15)) over the EnKF on average by about 0.15 over lead times of up to the time-to-peak of the basin. The ensemble results show that on average the CBEnKF improves the mean continuous ranked probability skill score (see Eq.(3.16)) over the EnKF by about 0.2 for all observed flows in significant events, and by about 0.3 or larger for observed flows exceeding the 95th percentile in those events. That the margin of improvement is larger for larger flows is a particularly appealing aspect of the CBEnKF. Both the ensemble mean and ensemble results show that the considerable improvement by the CBEnKF over the EnKF is realized over a wide range of combinations of observational uncertainties. As such, one may expect the CBEnKF to improve over the EnKF even if the uncertainty modeling may not be very accurate, an important consideration in operational forecasting. Visual comparison of forecast hydrographs with verifying observations indicates that the CBEnKF ensemble forecasts are generally of higher quality in the mean and spread than the EnKF ensembles, but that, for a few basins, the quality of both the CBEnKF and EnKF ensembles is compromised by phase errors in the UH and lack of parametric uncertainty modeling in the maximum soil water storages of the SAC. The CRPS decomposition indicates that the improved skill in CBEnKF ensembles is due to improvement in both reliability and resolution, and that their contributions are comparable in magnitude at short lead times. The examination of the updated and open-loop states indicates that the CBEnKF is able to improve the quality of streamflow analysis over the EnKF by adjusting the model states beyond what the EnKF is able to, and that, when the CBEnKF and EnKF produce streamflow analyses of comparable quality, the CBEnKF generally makes smaller adjustments to the open-loop state variables than the EnKF while still outperforming the EnKF in streamflow prediction. The last observation is supported by the examination of the joint relationships of the SAC states with and without DA in that the CBEnKF-updated states tend to follow the patterns of the open-loop states more closely than the EnKF-updated. It is hence surmised that the improved skill is due in part to the CBEnKF-updated states being more in line with the open-loop model dynamics than the EnKF-updated.

The CBEnKF in its current form is computationally expensive (Shen et al. 2019). Compared to the EnKF, the CBEnKF additionally requires inverting an (n×n) and two (m×m) matrices in each assimilation cycle where m and n denote the number of state variables and observations, respectively. Toward remedying the situation, at least two approaches seem possible. The first is to build a hybrid filter which runs the EnKF in “normal” conditions but switches to the CBEnKF in non-normal conditions. The second is to reduce the dimensionality of the linear systems via the singular value decomposition or its variant. The latter will also allow assessment of dynamically-varying information content in the observation and forecast error covariance matrices. The state augmentation approach used in this work renders the dimensionality of the state vector very large. Alternative approaches are possible which greatly reduces dimensionality but at some expense of reduced accuracy. The weight for the CB penalty, α, is specified in this work by prescribing the initial value, α_0, and reducing successively until a valid solution is obtained. If α could be specified adaptively without iterations based on the state of the dynamical system, the computational burden associated with the CBEnKF would be significantly reduced (Shen et al. 2019). We are currently evaluating lower-dimensional reformulation of the DA problem, the use of the randomized decomposition approaches for rank reduction (Halko et al., 2011), and machine learning-based state classification approaches for adaptive filtering, and the results will be reported in the near future.

Chapter 4. Adaptive CBEnKF for CB-Aware DA

4.1 Introduction

Data assimilation (DA) is widely used to reduce uncertainty in the initial conditions (IC) of prediction models by fusing the actual and model-predicted observations (Evensen, 1994; Evensen and van Leeuwen, 2000; Moradkhani et al., 2005; Liu and Gupta, 2007; Reichle, 2008; Liu et al., 2012). For cost-effective DA operation and observation, it is necessary to consider not only the DA method but also the flow-dependent information content. For example, an additional streamflow observation at the base of the rising limb of an unfolding hydrograph has a larger marginal information content relative to the model prediction than one at the lower recession limb. It is hence desirable to make more frequent observations of streamflow in the early rise phase of the streamflow response and schedule the DA cycles accordingly. In the context of DA, the marginal information content of an observation may be expressed in terms of the reduction in uncertainty in the ICs due to DA relative to the uncertainty without DA. If the assimilation completely eliminates the uncertainty in the ICs, the marginal information content of the observation would be 100%. If it does not reduce the uncertainty at all, the observation has no marginal information content. For optimal information fusion, observation and DA operation, it is hence important that the DA method be able to model the predictive uncertainty as accurately as possible.

In the CB-penalized techniques, the largest additional source of uncertainty and potential subjectivity is the scalar weight for the CB penalty, α (see Eq. (4.1) in Section 0). In the context of multiple linear regression, α may be considered analogous to the tuning parameter, λ, in Tikhonov regularization (Tikhonov and Arsenin, 1977; Tikhonov et al., 1995) or ridge regression (Hoerl, 1962; Hoerl and Kennard, 1970), or its L_1 version, the Least Absolute Selection and Shrinkage Operator LASSO, (Tibshirani, 1996, 1997), although the nature of the penalty is quite different (Jozaghi et al., 2020). The choice for the value of α depends on whether the state of the system that manifests the predictor-predictand relationship being observed is in a normal or extreme state. If the system is at or near its median state, α should be set to near zero so that CBEnKF reduces to EnKF and produces the least squares solution. If the system is in an extreme state, α should be set to a large value so that CB is reduced as much as possible. In this work, we describe an adaptive extension of CBEnKF in which α is prescribed in real time flow-dependently. The premise of the approach is that, by significantly reducing the parametric uncertainty in α, one may quantify the marginal information content of the observations more accurately, which in turn improves the DA-aided analysis and prediction, particularly when the hydrologic system is in non-normal states. Toward the above end, we describe adaptive CBEnKF, comparatively evaluate with EnKF, and assess the flow-dependent information content in relation to streamflow response to precipitation at the catchment scale. For quantification of marginal information content, we use the degrees of freedom for signal (DFS) (Rodgers, 2000; Zupanski et al., 2007) which may be considered as a skill score for information fusion. The new and significant contributions of this chapter are: development of CB-aware DA in the form of adaptive CBEnKF for assimilation of streamflow observations, comparative evaluation of adaptive CBEnKF with EnKF, and advances in understanding of flow-dependent information content in hydrologic prediction and its temporal variations in relation to streamflow response at the catchment scale. This chapter is an adaptation of Shen et al. (2020) and is organized as follows. Section 4.2 describes the hydrologic models, streamflow and precipitation data, and basins used for the study. Section 0 describes CBEnKF and formulation of the control vector, adaptive extension, flow-dependent information content and uncertainty modeling. Section 4.4 presents the results. Section 4.5 provides conclusions and future research recommendations.

4.2 Models, Data and Basins Used

In this section, we describe the hydrologic models, the forcings and streamflow data and the basins used in this study. They were used in a number of previous studies of hydrologic modeling and DA (Seo et al., 2009; Lee et al., 2012, 2016, 2019; Lee and Seo, 2014; Rafieeinasab et al., 2014b). As such, we limit the descriptions only to those that are essential to the development of this paper.

4.2.1 Hydrologic Models Used

We used the Sacramento soil moisture accounting model (SAC) (Burnash et al., 1973) for rainfall-runoff modeling and the unit hydrograph (UH) (Chow et al., 1988) for runoff routing. They represent long-standing hydrologic models used at the River Forecast Centers (RFC) of the National Weather Service (NWS) for operational hydrologic forecasting for headwater basins. The SAC inputs mean areal precipitation (MAP) and mean areal potential evapotranspiration (MAPE) and outputs the total channel inflow (TCI) which represents the sum of surface and subsurface runoffs into the channels. In this work, we assume clairvoyant forcing forecasts and use (observed) MAP and MAPE for future input as well. The UH inputs the TCI and outputs discharge at the catchment outlet. The SAC is a conceptual soil moisture accounting model and uses two layers of soil storages, i.e., the upper and lower zones, to model soil water balance in the vertical. Surface runoff includes both fast and slow runoffs. Fast runoff consists of the rainfall intensity-dependent surface runoff from the unsaturated area, direct runoff from the saturated area, and impervious runoff from the impervious area. Slow runoff processes produce interflow, supplemental baseflow, and primary baseflow. The UH model is used to route the runoff components to produce streamflow at the outlet of the basin.

The objective of DA in this work is reduce the IC uncertainty. As such, it is important to reduce other sources of error, such as systematic biases in observed forcings, systematic timing errors, and model parametric uncertainties, as much as possible so that one may ascertain the effectiveness of DA for the intended purpose of reducing the IC uncertainty. In this work, we used hourly MAP given the flashiness and short response time of a number of the study basins. For MAPE, we used climatological mean values adjusted via calibration. In SAC, MAP and MAPE may be adjusted for long-term biases via multiplicative correction factors. To estimate the adjustment factors, derive the empirical UH and refine the SAC parameters, we used the Adjoint-Based OPTimizer (AB_OPT) (Seo et al., 2009). For details on the calibration technique used in AB_OPT, the reader is referred to (Kuzmin et al., 2008).

4.2.2 Data Used

The observations assimilated are MAP, MAPE, and streamflow. The MAP observations are from the Multisensor Precipitation Estimator (Fulton et al., 1998; Seo, 1998; Seo et al., 1999, 2010; Young et al., 2000) produced operationally by the West Gulf River Forecast Center (WGRFC). For MAPE, the monthly climatology of daily estimates were disaggregated to hourly estimates based on the diurnal cycle of solar radiation. Because the DA cycles consist only of the periods of significant flow, potential evapotranspiration does not play an important role and the lack of dynamically-varying MAPE does not impact the outcome in any significant way. For assimilation and validation, we used the US Geological Survey's hourly streamflow data at the catchment outlet. Because our focus is on flood forecasting, we identified only the significant events from the observed streamflow data and used the resulting partial-duration time series. A significant event is defined as a hydrograph whose peak flow exceeds 100 m^3/s. Once an event is identified, the beginning and ending hours of the event were determined by subtracting and adding 72 hrs from and to the first and last hours of the partial-duration exceeding 100 m^3/s, respectively. The choice of 72 hrs is made to include the most significant portions of the primary baseflow response in the assimilation process.

4.2.3 Basin Used

A total of 6 basins was used in the WGRFC's service area: DCJT2, LYNT2, MDST2, MTPT2, REFT2 and SBMT2 (see Figure 4-1 and Table 4-1 for location and hydroclimatological attributes). In Table 4-1, the dryness index (DI) is defined as the ratio of mean MAPE to mean MAP (Budyko, 1974). A value of DI greater and less than unity indicates the catchment is water- and energy-limited, respectively. They are part of the WGRFC DA testbed which includes a group of 23 basins identified as particularly challenging for operational streamflow forecasting. The

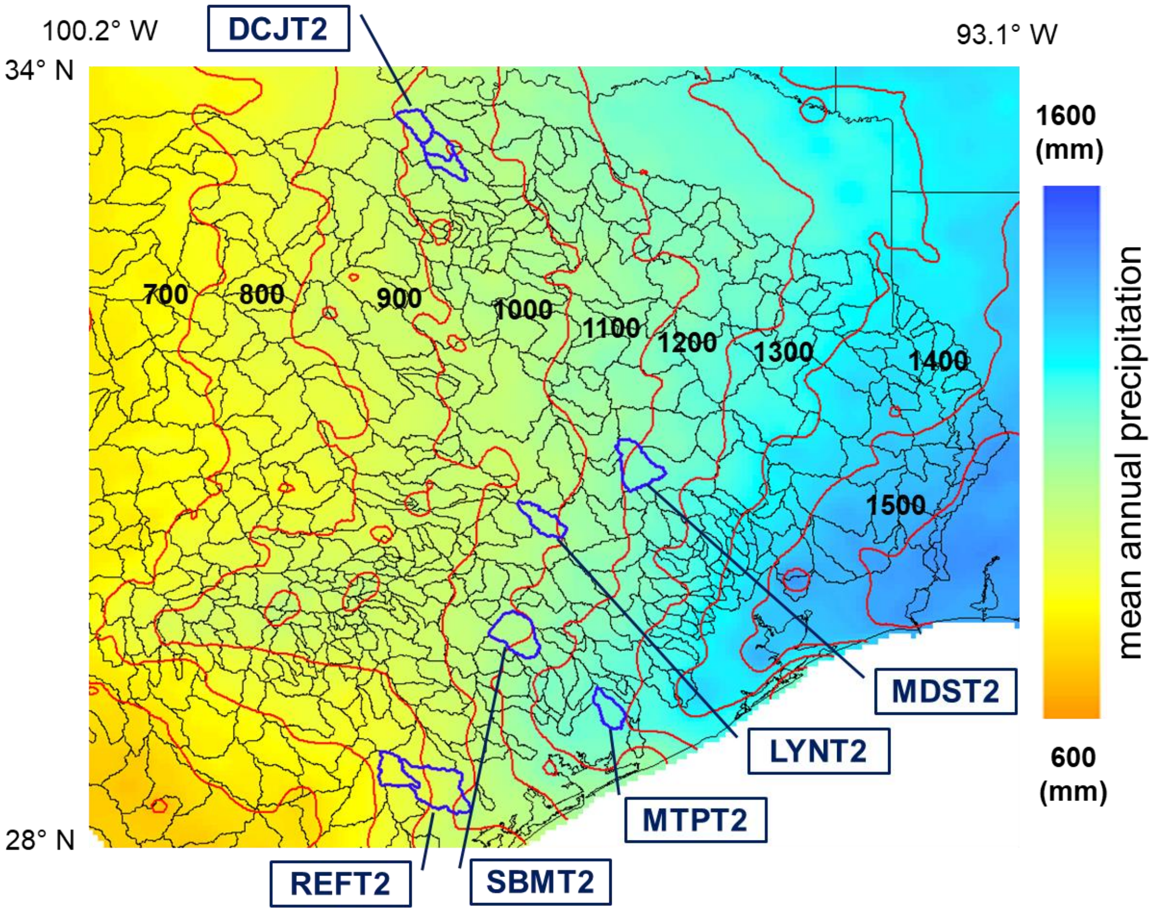


Figure 4-1 Locations of the 6 basins in the WGRFC's service area used in the study, DCJT2, LYNT2, MDST2, MTPT2, REFT2 and SBMT2, with mean annual precipitation and its contour lines

testbed has been previously used by the authors and their collaborators for hydrologic DA and related research (Seo et al., 2009; Liu et al., 2011; Lee et al., 2012, 2016, 2019; Rafieeinasab et al., 2014b; Kim et al., 2018). The 6 basins used in this work are located along the longitudinal line of mean annual precipitation of approximately 1,000 mm. Due to the strong east-to-west gradient in precipitation climatology and the influence of the Gulf of Mexico, significant hydroclimatological variations exist among the basins ranging from semi-arid to humid, providing a degree of diversity (see Figure 4-1, Table 4-1).

Table 4-1 Study basins and their attributes (see Figure 4-1 for locations)

	Basin ID	Basin name	USGS ID	Area (km^2)	Time to peak (hr)	Average annual discharge (m^3/s)	Average annual precipitation (mm)	Dryness index
1	**DCJT2**	Justin – Denton Creek	08053500	1039	6	2.38	684	1.47
2	**LYNT2**	Lyons – Davidson Creek	08110100	508	18	2.45	858	1.25
3	**MDST2**	Madisonville – Bedias Creek	08065800	870	21	6.40	933	1.05
4	**MTPT2**	Midfield – Tres Palacios	08162600	435	17	4.89	1126	0.89
5	**REFT2**	Refugio – Mission River	08189500	1787	39	4.48	748	1.06
6	**SBMT2**	Sublime – Navidad River	08164300	896	26	5.01	934	1.34

Streamflow forecasting is generally more challenging for more arid basins due to the quicker catchment response, larger uncertainties in the ICs, generally larger structural uncertainties due to unmodeled processes such as infiltration into dry channel beds, and larger parametric uncertainties due to the above factors and the limited number of large events available for calibration (Alizadeh et al., 2020). In addition, the large spatio-temporal variability of precipitation from convective events adds uncertainty in lumped modelling particularly for large basins (Smith et al., 2004). Hence, the study basins collectively offer a rather challenging test for streamflow DA. In addition to the hydroclimatological variations, we also considered the size of the basins as explained below. Given the hydrologic models, observations and DA method used, the potency of DA depends primarily on the memory of the hydrologic system (Alizadeh et al., 2020). In general, the larger the storage of surface and subsurface water is, the larger the streamflow predictability is. Streams in semi-arid basins may be ephemeral due to intermittent soil water in some of the storages, resulting in significantly smaller predictability. Accordingly, one may expect the potential gains from DA to be relatively small for small urban basins or semi-arid basins even though they may be relatively large in size. For this reason, we chose larger basins in the 23-basin testbed even though they are less likely to satisfy the UH assumptions (Chow et al., 1988; Liu et al., 2011).

4.3 Methods Used

In this section, we describe adaptive CBEnKF and how flow-dependent marginal information content is assessed using DFS. The technical details of CBEnKF are not central to the development of this work. For a complete mathematical and algorithmic description of CBEnKF, the reader is referred to Lee et al., (2019). For a complete derivation of CBPKF, the reader is referred to Seo et al. (2018a, 2018b) and Shen et al. (2019). For a simpler and computationally less-expensive approximation of CBPKF via variance-inflated KF, the reader is referred to Shen et al. (2019).

4.3.1 CBEnKF and Formulation of Control Vector

CBEnKF (Lee et al., 2019) is an ensemble extension of CBPKF (Seo et al., 2018a, 2018b) and is completely analogous to what EnKF (Evensen, 1994; Lorentzen and Naevdal, 2011) is to KF (Kalman, 1960). Whereas KF minimizes error variance only, CBPKF minimizes a linearly weighted sum of error variance and expectation of the Type-II error squared:

$$\Sigma_k = E_{X_k,X_k^*}[(X_k - X_k^*)(X_k - X_k^*)^T] + \alpha \cdot E_{X_k}\left[\left(X_k - E_{X_k^*}[X_k^*|X_k]\right)\left(X_k - E_{X_k^*}[X_k^*|X_k]\right)^T\right] \qquad 4.1$$

where X_k^* denotes the vector of the estimated states, the first and second terms represent the error covariance and the CB penalty, respectively, and α denotes the scalar weight given to the latter. If α=0, the CBPKF and CBEnKF reduce to the KF and EnKF, respectively. The CBEnKF and EnKF are applied as sequential assimilators in this work but using the fixed-lag smoother formulation (Lee et al., 2019) The size of the fixed lag, or the assimilation window (see Figure 4-2), is chosen to be the same as the length of the UH. In this way, the time lag between the runoff generation at the most upstream areas of the catchment and the resulting flow appearing at the basin outlet is captured in the DA process (Seo et al., 2009; Rafieeinasab et al., 2014b). The motivation for the fixed-lag smoother approximation, as opposed to directly using a smoother formulation (Evensen and van Leeuwen, 2000), is to handle both cold and warm restarts with a clearly identifiable assimilation window for human forecaster-supervised operation (Seo et al., 2009). The duration of the UH is 40, 40, 53, 81, 72 and 43 hrs for DJCT2, LYNT2, MDST2, MTPT2, REFT2 and SBMT2, respectively (see Figure 4-2).

As with any models, hydrologic prediction using SAC-UH is subject to various sources of uncertainty. Because MAP is a spatially averaged quantity, purely random errors tend to cancel out whereas systematic errors do not. To address time-varying biases in the forcings, we include in the control vector, X_k, in Eq. (4.1) the multiplicative adjustment factors to MAP and MAPE, X_k^P and

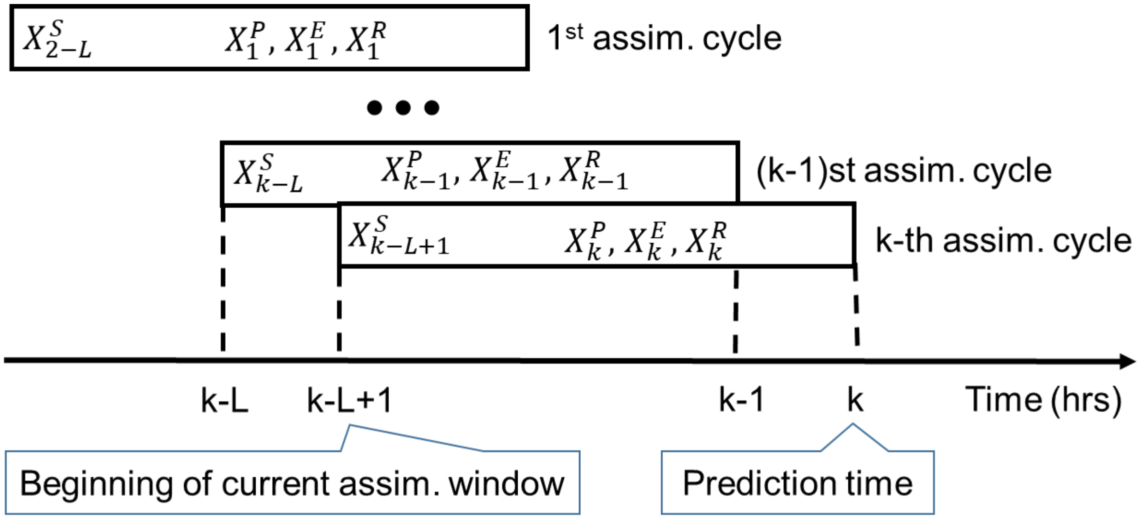


Figure 4-2 Schematic of the assimilation cycles based on the fixed-lag smoother formulation

X_k^E, respectively. We assume that $\ln(X_k^P)$ and $\ln(X_k^E)$ are normally distributed and autoregressive (AR)-1 (Smith and Krajewski, 1991; Anagnostou and Krajewski, 1999; Seo et al., 1999):

$$LX_k^P = \Phi_P LX_{k-1}^P + W_P \tag{4.2}$$

$$LX_k^E = \Phi_E LX_{k-1}^E + W_E \tag{4.3}$$

where LX_k^P and LX_k^E denote $\ln(X_k^P)$ and $\ln(X_k^E)$, respectively, Φ_P and Φ_E denote lag-1 autocorrelation for $\ln(X_k^P)$ and $\ln(X_k^E)$, respectively, and W_P and W_E denote the white noise for the respective AR(1) models. The SAC soil moisture states are conceptual quantities which may also be expressed as volumetric soil water contents (Lee et al., 2011; Koren et al., 2014). In reality, soil water is seldom observed at catchment scale. As such, the SAC states are subject to potentially large structural and parametric uncertainties of the soil moisture accounting model in addition to input uncertainty in observed forcings (Carpenter et al., 2001; Yilmaz et al., 2008; Lee and Seo, 2014; Lee et al., 2016). In this work, we include all 6 SAC state variables, X_k^S, valid at the beginning of the assimilation window in the control vector for updating. The temporal evolution of the SAC states is modelled as:

$$X_{k-L+1}^S = M(X_{k-L}^S) + W_{k-L}^S \tag{4.4}$$

where X_{k-L+1}^S denotes the SAC states valid at the beginning of the assimilation window at the k-th assimilation cycle, $M(\quad)$ denotes the SAC, and W_{k-L}^S denotes the model error for state transition from the (k-1)-st to the k-th assimilation cycle (see Figure 4-2). The zero-mean errors, W_{k-L}^S, are assumed to be independent of one another and have standard deviation of approximately 4% of the time-varying SAC states.

As noted above, the study area is hydroclimatologically challenging for streamflow prediction (Alizadeh et al., 2020). It is very possible that the ensemble subspace for streamflow spanned by X_k^P, X_k^E and X_{k-L+1}^S may fail to encompass the observed flow due to a combination of large hydrologic and input uncertainties (Krzysztofowicz, 1999; Seo et al., 2006). For this reason, we include the additive error to TCI, X_k^R, to the control vector. The control variable X_k^R may be best understood in the context of the UH operation expressed as a continuous convolution shown below:

$$Z_Q(t) = \int_0^t \{I(\tau) + w(\tau)\} u(t-\tau) d\tau \quad I(\tau) + w(\tau) \geq 0 \tag{4.5}$$

where $\mathrm{Z}_Q(\mathrm{t})$ denotes the observed streamflow valid at time t, $I(\tau)$ denotes the TCI, $w(\tau)$ denotes the additive error to TCI, and $u(\quad)$ denotes the UH. The control variable, X_k^R, is a discrete implementation of w(   ) in Eq. (4.5) and represents the time-varying errors in SAC-UH that cannot be addressed by adjusting the SAC states, MAP and MAPE alone. We assume that X_k^R is normally distributed and AR(1):

$$X_k^R = \Phi_R X_{k-1}^R + W_R \tag{4.6}$$

where Φ_R and W_R denote the lag-1 correlation coefficient and the white noise for the AR(1) model, respectively. With the addition of X_k^R, the control vector in Eq.(1) for the k-th assimilation cycle is given by $X_k = (X_{k-L+1}^S, X_k^P, X_k^E, X_k^R)^T$. The above formulation has proven very effective in the study area in numerous previous studies (Seo et al., 2003; Lee et al., 2011, 2012, 2016; Lee and Seo, 2014; Rafieeinasab et al., 2014; Noh et al., 2018).

The control variables, X_k^P, X_k^E and X_k^R, are applied to MAP, MAPE and TCI valid over the assimilation window whereas X_{k-L+1}^S is applied at the beginning of the window only (see Figure 4-2). The highest-dimensional implementation of the above control vector would be to allow X_k^P, X_k^E

and X_k^R to vary at hourly scale within the assimilation window. The lowest-dimensional implementation is to assume temporal uniformity of X_k^P, X_k^E and X_k^R over the entire assimilation window. Lee et al. (2012) have shown using VAR that modeling the above control variables at an hourly scale improve performance only rather modestly while substantially increasing the computational amount and possibly contributing to under-determinedness of the inverse problem. As such, we chose the lowest-dimensional implementation in this work. It is possible, however, this choice may compromise the performance of DA for large basins with lengthy response time for fast runoff. Note that, for such basins, even a very small $w(\tau)$ may greatly alter the adjusted runoff hyetograph due to the nonlinearity in $I(\tau) + w(\tau), 0 \leq \tau \leq t$ (see Eq. (5)). For assimilation of streamflow observation using linear filtering, state augmentation is necessary to render the observation equation linear in appearance (Lorentzen and Naevdal, 2011; Rafieeinasab et al., 2014; Lee et al., 2019). In this work, we assimilate observed streamflow valid at the prediction time only, which adds another control variable to the control vector. There are hence a total of 10 control variables, i.e., the 6 SAC states for X_{k-L+1}^S, 1 each for X_k^P, X_k^E and X_k^R, and the model-simulated streamflow analysis, representing the lowest-dimensional formulation possible for streamflow DA using SAC-UH. For the CBEnKF solution and the algorithm for the DA problem formulated above, the reader is referred to Lee et al. (2019).

Collectively, the control variables employed in this work encompass a very large feasible region for input and hydrologic uncertainties. For some basins, however, the variables may not well represent the actual errors, in which case they only add noise. To assess the relative performance of adaptive CBEnKF under varying levels of hydrologic uncertainty (Lee et al., 2012, 2016), we performed DA under three different levels of hydrologic uncertainty modeling. In the weakly-constrained implementation, or WC, we invoke the adjustment factors to forcings, uncertainty in soil moisture transition dynamics, W_{k-L}^{S}, and uncertainty in runoff depth, w(τ), which results in the control vector of $X_k = (X_{k-L+1}^{S}, X_k^{P}, X_k^{E}, X_k^{R})^T$. In the moderately-constrained implementation, or MC, we drop w(τ) from WC, which results in the control vector of $X_k = (X_{k-L+1}^{S}, X_k^{P}, X_k^{E})^T$. The strongly-constrained implementation, or SC, is the same as MC but the variance of W_{k-L}^{S} is reduced by an order of magnitude to emulate the control vector of $X_k = (X_k^{P}, X_k^{E})^T$. In this way, we assume near-perfect hydrologic models in SC, uncertain soil moisture ICs in MC valid at the beginning of the assimilation window, and uncertain soil moisture ICs and runoff depth in WC. It is important to note that, with temporally uniform X_k^{R} over the assimilation window, we are implicitly assuming perfect UH, which may not be very realistic particularly for large basins under highly convective precipitation (Dooge, 1959; Chow et al., 1988, p. 1; Saghafian, 2006; Liu et al., 2011).

4.3.2 Adaptive Extension

Although CB-penalized estimation significantly improves performance over the tails of the distribution of the predictand, they do not yield minimum unconditional mean squared error (MSE) due to the additional penalty (see Eq. (4.1)). This deterioration in unconditional performance may be reduced by prescribing the weight to the CB penalty, α, adaptively based on the best available estimate of the unknown true state. The basic strategy is to prescribe a large weight if the true state is in the tail and a small weight if it is near median, respectively (Kim et al., 2018; Shen et al., 2019). Because the true state is not known in reality, α can only be prescribed sub-optimally. Using synthetic experiments with known truth, Shen et al. (2019) assessed the upper bound of the possible improvement by CBPKF over KF by prescribing α perfectly. The results show that, with uncertainty in α removed, adaptive CBPKF is not only far superior to KF for large truth but also improves over KF in the unconditional sense as well. Whereas the upper bound of the performance seen in Shen et al. (2019) is not attainable in reality, it is possible to use real-time streamflow observations to infer the state of the hydrologic system and prescribe α adaptively. To that end, we optimize α by minimizing the Continuous Ranked Probability Score (CRPS; Hersbach, 2000) of the ensemble streamflow analysis valid at the prediction time (see Figure 4-2).

Because streamflow data have observational uncertainties, the above optimization of α should ideally be carried out in an ensemble mode using the uncertainty-added streamflow observations (Lee et al., 2019). Due to excessive computation, however, it is assumed in this work that streamflow is observed perfectly for the purpose of optimizing α (but not for CBEnKF itself). A potential negative consequence of this shortcut is that α may be overfit and may deteriorate performance at longer lead times particularly when phase errors are present (Liu et al., 2011). The optimization is carried out by selecting multiple values for starting α and, for each assimilation cycle, running CBEnKF for each of the starting values and identifying the CRPS-minimizing CBEnKF solution. An optimum α of zero is an indication that CB is not important, and that EnKF suffices for the DA cycle. A large value of optimal α is an indication that significant CB exists and one may expect CBEnKF to improve over EnKF.

4.3.3 Flow-Dependent Information Content

To quantify the flow-dependent marginal information content in the observations, we use the DFS (Rodgers, 2000; Zupanski et al., 2007). The DFS is widely used in atmospheric DA for prognostic assessment of the impact of new observations (Fisher, 2003; Zupanski et al., 2007; Zupanski, 2009). The DFS is based on the degrees of freedom (DOF) of the cost function under the normality assumption for the true state, X_k, and describes the number of useful independent quantities in the observation (Rodgers, 2000):

$$\chi_k^2 = E\left[\left(\hat{X}_k - X_{k|k-1}\right)^T \Sigma_{k|k-1}^{-1}\left(\hat{X}_k - X_{k|k-1}\right)\right] + E\left[\left(Z_k - H_k\hat{X}_k\right)^T R_k^{-1}\left(Z_k - H_k\hat{X}_k\right)\right] \quad (4.4)$$

where $\hat{X}_k$ denotes the optimal solution. In Eq. (4.2), the first term is referred to as the DFS, d_s, and is attributable to the state vector. The second term is referred to as the DOF for noise, d_n, and is attributable to the residuals, or noise. The sum of the two equals the number of measurements, i.e., $d_s + d_n = m$ where m denotes the number of measurements (Rodgers, 2000). Because the DOF exceeding the number of observations, m, or the number of states, n, is meaningless, the largest possible DOF is given by the smaller of m and n, where n denotes the dimensionality of the state vector. In the context of KF (or a similar sequential filter), one may express the DFS, d_s, as (Rodgers, 2000):

$$d_s = Tr[(\Sigma_{k|k-1} - \Sigma_{k|k})\Sigma_{k|k-1}^{-1}] = Tr\left[I - \Sigma_{k|k}\Sigma_{k|k-1}^{-1}\right] = Tr[K_k H_k] \quad (4.5)$$

In the above, the Kalman gain, K_k, is given by:

$$K_k = \Sigma_{k|k} H_k R_k^{-1} \quad (4.6)$$

In the above, the updated covariance matrix for the state is given by:

$$\Sigma_{k|k} = \left(H_k^T R_k^{-1} H_k + \Sigma_{k|k-1}^{-1}\right)^{-1} \quad (4.7)$$

where $T_r[\ ]$ denotes the trace of the matrix bracketed, I denotes the identity matrix, and $\Sigma_{k|k}$ and $\Sigma_{k|k-1}$ denote the updated and forecast error covariance matrices of the state variables, respectively. The DFS may also be evaluated by singular-value-decomposing $\widetilde{H}_k$ below:

$$\widetilde{H}_k = R_k^{\frac{1}{2}} H_k \Sigma_{k|k-1}^{\frac{1}{2}} = U\Lambda V^T \quad (4.8)$$

where U and V are unitary matrices and Λ is the diagonal matrix of singular values. The DFS is then given by (Rodgers, 2000; Zupanski, 2009):

$$d_s = Tr[\Lambda^2(\Lambda^2 + I)^{-1}] = \Sigma_{i=1}^{m} \frac{\lambda_i^2}{1 + \lambda_i^2} \tag{4.9}$$

where λ_i denote the i-th largest singular value.

As seen in Eq. (4.8), the DFS quantifies the reduction in uncertainty due to the observation relative to the uncertainty without. If assimilating the observation completely reduces or does not reduce uncertainty at all, we have $d_s = \min\{m, n\}$ or $d_s = 0$, respectively. In sequential DA, one may hence consider d_s as a skill score (when normalized by dividing by $\min\{m, n\}$) for uncertainty reduction due to the assimilation of new observations. Because d_s depends on the relative magnitude of $\Sigma_{k|k}$ and $\Sigma_{k|k-1}$, its utility as a skill score for information fusion depends on the goodness of the uncertainty modeling. It is possible that one may calibrate the uncertainty models to increase d_s but without improving forecast skill. For this reason, it is necessary to accompany the information content analysis with verification of the ensemble analysis and forecast. Lee et al.(2019) have shown that ensemble predictions for flood events from CBEnKF are more reliable and have larger resolution than those from EnKF. One may hence expect adaptive CBEnKF to provide more realistic characterization and quantification of the DFS and further improve predictive skill.

4.3.4 Uncertainty Modelling

The accuracy of the DFS depends on the goodness of the uncertainties modelling. As such, it is particularly important to model the dynamical and observational uncertainties as accurately as possible. In this work, observational errors are modeled as heteroscedastic and temporally uncorrelated. For heteroscedastic modeling of observation errors for streamflow and MAP, we used the following formulations (Sorooshian and Dracup, 1980; Carpenter and Georgakakos, 2004; Weerts and El Serafy, 2006; Rakovec et al., 2012):

$$\mathrm{\sigma_P} \ = \mathrm{C_P Z}_P + \mathrm{C_{Po}}(mm/hr) \tag{3.10}$$

$$\mathrm{\sigma_Q} \ = \mathrm{C_Q} Z_Q + \mathrm{C_{Qo}}\ (mm/hr) \tag{3.11}$$

where $\mathrm{Z_Q}$ and $\mathrm{Z_P}$ denote the streamflow and MAP observations, respectively, σ_Q and σ_P denote the observation error standard deviations for streamflow and MAP, respectively, $\mathrm{C_Q}$ and $\mathrm{C_P}$ denote the multiplicative coefficients for the streamflow and MAP observation error standard deviation, respectively, and $\mathrm{C_{Qo}}$ and $\mathrm{C_{Po}}$ denote the additive coefficients for the streamflow and MAP observation error standard deviation, respectively. The additive coefficients are necessary to keep the variance positive when the observation is zero. For hourly MAPE, we used a fixed observational uncertainty of 1 mm^2. The coefficients in Eqs.(18) and (19) should be carefully chosen as inaccurate heteroscedastic modeling may be counter-productive compared to homoscedastic modeling (Rafieeinasab et al., 2014b). Ideally, the parameters should be estimated based on statistical analysis using the truth. Due to lack of such data, however, the above coefficients are estimated in this work from a combination of the literature (Carpenter and Georgakakos, 2004; Weerts and El Serafy, 2006; Clark et al., 2008; Rakovec et al., 2012) and limited sensitivity analysis. Based on the above, we chose 0.15 and 0.25 for C_Q and C_{Qo}, respectively, and 0.4 and 0.25 for C_P and C_{Po}, respectively.

Because runoff is not directly observed, variance of X_k^R cannot be estimated empirically. In this work, it is inferred from the observed and simulated streamflow via the following steps.

Assuming stationarity of $w(\tau)$ in Eq. (4.5) within the UH duration, we have for error variance of simulated streamflow:

$$\sigma_{\mathrm{eq}}^2 = \sigma_{\mathrm{Rs}}^2 \int_0^t \int_0^t u(t-\tau)\, u(t-s) ds d\tau \tag{3.12}$$

where σ_{eq}^2 denotes the error variance of simulated streamflow and σ_{Rs}^2 denotes the runoff error variance. With Eq. (12), one may develop a regression which relates σ_{eq}^2 with the observed flow, and express σ_{Rs}^2 as a function of σ_{eq}^2 and hence of the observed flow (Rafieeinasab et al., 2014b). For the relationship between σ_{eq}^2 and observed flow, we used simple slope-only linear regression for simplicity. It is likely, however, that nonlinear modeling would improve performance. If significant timing errors exist, they would be incorrectly attributed to σ_{Rs}^2. In such cases, σ_{Rs}^2 estimated as described above is likely to be an overestimate. For this reason, we adjusted the regression slope within 10% based on sensitivity analysis.

4.4 Results

In this section, we present the comparative evaluation results and offer discussion. For each basin, adaptive CBEnKF and EnKF were run for all events whose peak flow exceeded 100 m^3/s within the 12-yr period of record of 1996 through 2007. We then verified ensemble mean streamflow analysis and predictions from the twin hindcasting experiments using CBEnKF and EnKF. Lee et al. (2019) showed that the CBEnKF ensembles are generally more reliable and has better resolution than the EnKF ensembles, and that the CRPS and continuous ranked probability skill score (CRPSS) results are qualitatively very similar to root mean square error (RMSE) and mean square error skill score (MSE-SS) of ensemble mean results, respectively. As such, we focus in this work on the ensemble mean results which are more easily interpretable and relatable. For comparative verification of CBEnKF ensemble predictions, the reader is referred to Lee et al. (2019). For performance measures, we used the RMSE, maximum errors of under- (MAXEU) and over-prediction (MAXEO), correlation coefficient (CORR) and MSE-SS (Lee et al., 2011, 2012; Lee and Seo, 2014). A large ensemble size of 200 was used for all experiments to reduce sampling uncertainty.

4.4.1 Ensemble Mean Streamflow Analysis and Flow-Dependent Information Content

To provide a visual sense of how the ensemble mean analysis from adaptive CBEnKF may compare with that from EnKF, we show in Figure 4-4 to Figure 4-7 examples of the observed hydrograph (black) with the adaptive CBEnKF analysis (red), EnKF analysis (blue) and DA-less ensemble mean simulation (green) overlaid. The DA results in the figure are based on the weakly-

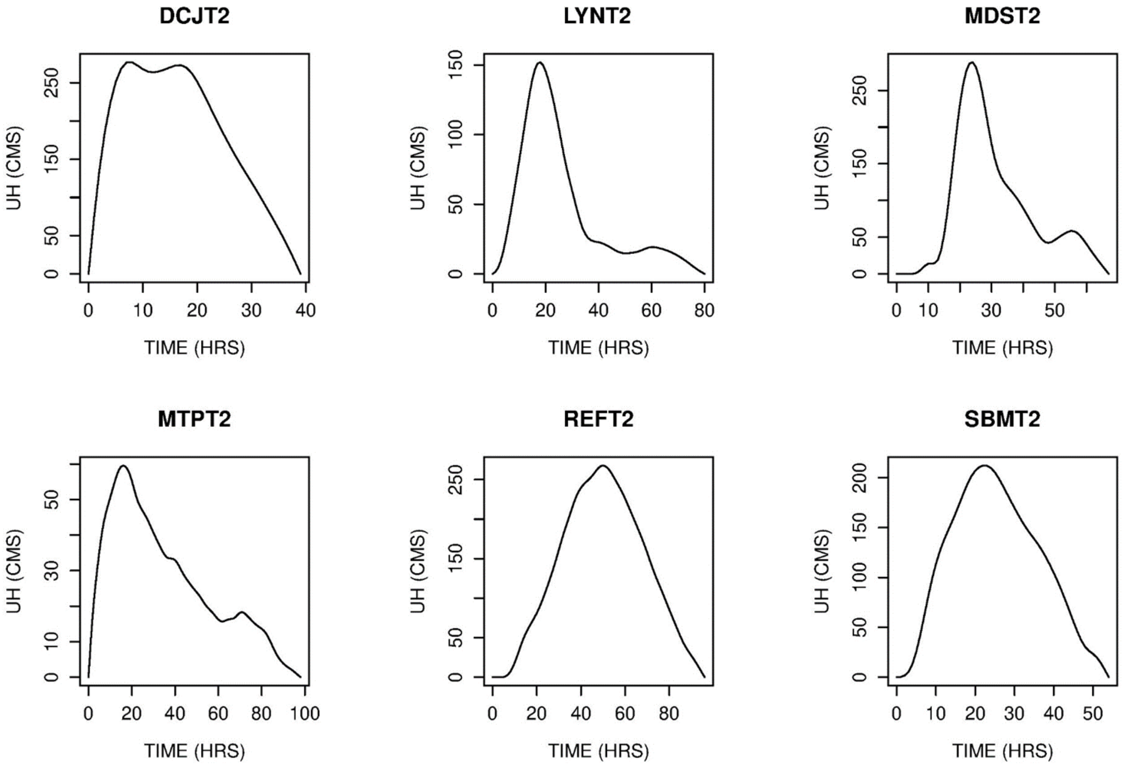


Figure 4-3 Empirical UHs for the 6 basins

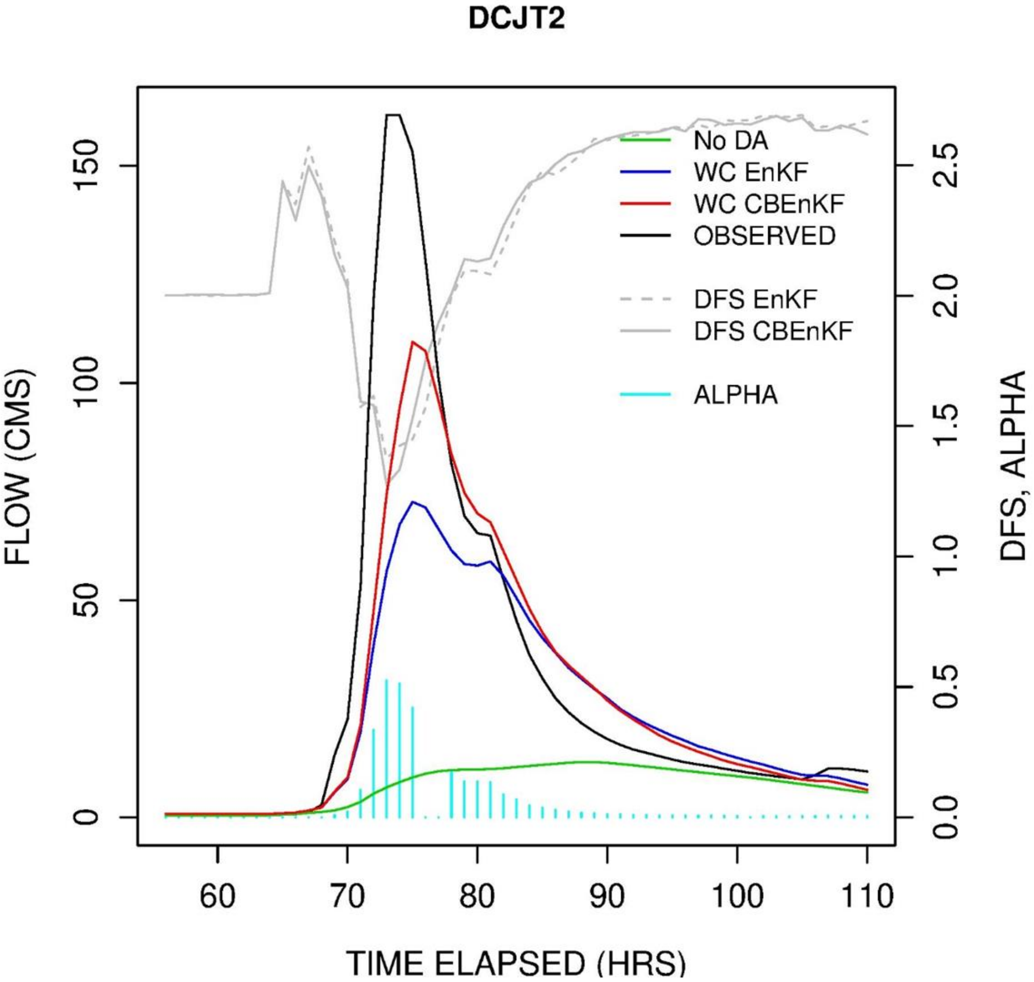


Figure 4-4 Ensemble mean analysis from adaptive CBEnKF (red) compared with that from EnKF (blue) for DCJT2. Also shown are the DA-less ensemble mean simulation (green), observed hydrograph (black), DFS (grey) and final α (cyan)

constrained implementation, or WC (see Section 4.3.1). The vertical cyan bars show the final values of α from adaptive CBEnKF following minimization of the CRPS of ensemble analysis. The larger the value of α is, the larger the relative weight given to the CB penalty is (see Eq. (4.1)). For those assimilation cycles with α of zero, CBEnKF is reduced to EnKF. The solid and dashed gray lines in the figure denote the DFS evaluated by adaptive CBEnKF and EnKF, respectively. At the beginning of each event, the EnKF and adaptive CBEnKF are initialized identically. Adaptive CBEnKF and EnKF are then performed independently of each other every hour under the identical conditions for the entire event. Skillful DA performance in times of very poor ICs is an

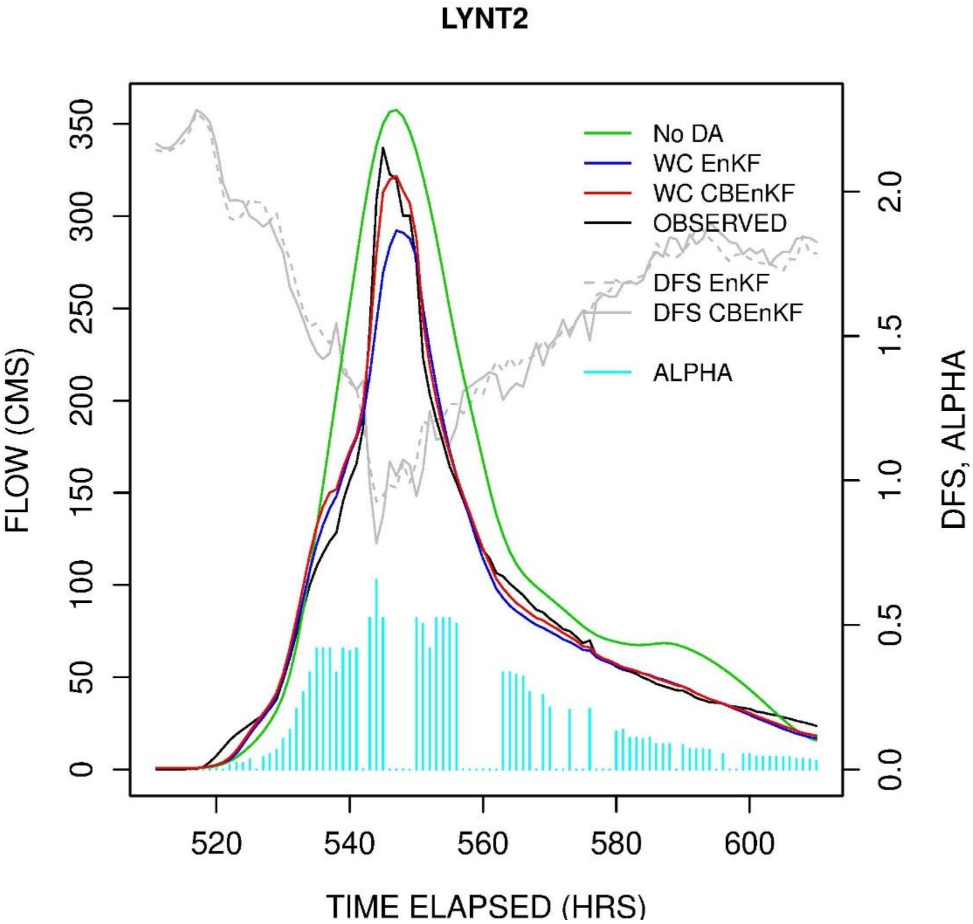


Figure 4-5 Ensemble mean analysis from adaptive CBEnKF (red) compared with that from EnKF (blue) for LYNT2. Also shown are the DA-less ensemble mean simulation (green), observed hydrograph (black), DFS (grey) and final α (cyan)

extremely important consideration for event-based flood forecasting (Kim et al., 2020). To reflect performance under wide-ranging qualities of the ICs, we also included in the evaluation the warmup periods for SAC. Accordingly, in the early part of the assimilation horizon, the base simulation has very large errors in the ICs, resulting in severe under-simulation (see Figs 4-4, 4-6, 4-7) which offers a very tough test for DA.

It is readily seen in Figure 4-4 to Figure 4-7 that the CBEnKF analysis tracks the observed hydrographs more closely than the EnKF analysis. The reduction in RMSE of streamflow analysis

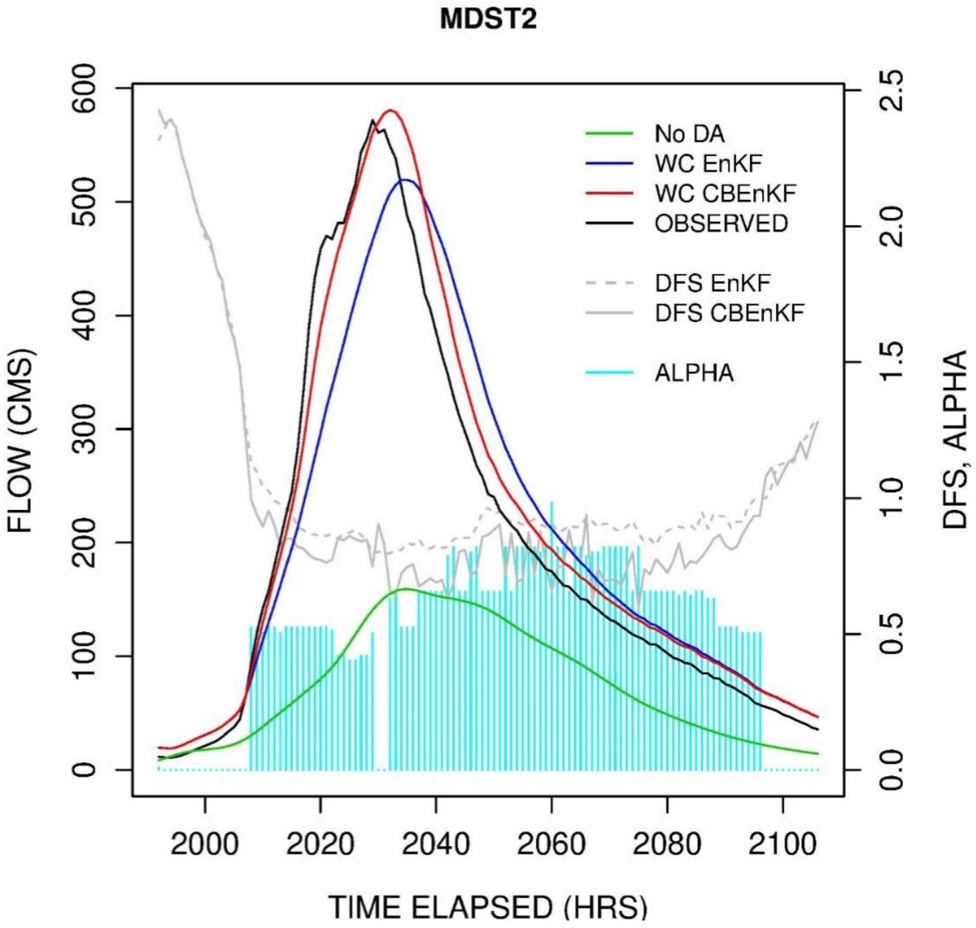


Figure 4-6 Ensemble mean analysis from adaptive CBEnKF (red) compared with that from EnKF (blue) for MDST2. Also shown are the DA-less ensemble mean simulation (green), observed hydrograph (black), DFS (grey) and final α (cyan)

by adaptive CBEnKF over EnKF is 20.8, 29.9, 24.5, 44.1, 10.2 and 17.6% for DCJT2, LYNT2, MDST2, MTPT2, REFT2 and SBMT2, respectively. The differences in the analysis results are due to the fact that accounting for CB tends to increase the model-predicted uncertainty, $\Sigma_{k|k-1}$, particularly in high flows, and hence CBEnKF weighs the observed streamflow more heavily. As a specific example, note in Figure 6 that the EnKF analysis is not able to keep up with the fast-rising observed flow in the rising limb of the hydrograph whereas the CBEnKF analysis is, owing to the significant CB penalty being invoked at almost every assimilation cycle during that period. Figure 4-4 to Figure 4-7 indicates that CB occurs rather frequently. The number of assimilation

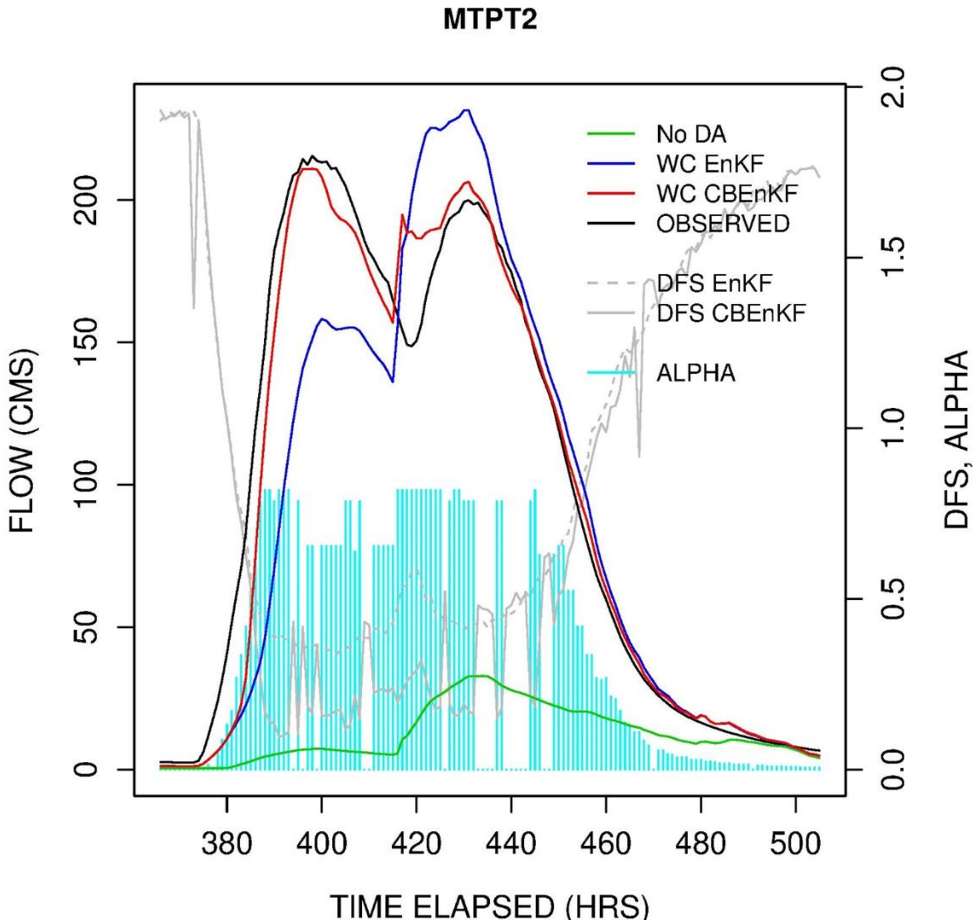


Figure 4-7 Ensemble mean analysis from adaptive CBEnKF (red) compared with that from EnKF (blue) for MTPT2. Also shown are the DA-less ensemble mean simulation (green), observed hydrograph (black), DFS (grey) and final α (cyan)

cycles with $\alpha > 0.1$ is approximately 13, 24, 29, 34, 17 and 22% of the entire partial-duration assimilation horizon for DCJT2, LYNT2, MDST2, MTPT2, REFT2 and SBMT2, respectively. The CB penalty is invoked particularly frequently in the rising limb and in the receding limb surrounding the inflection point where runoff generation transitions from the lower-zone supplemental free water content to lower-zone primary free water content of SAC.

Figure 4-4 to Figure 4-7 shows that DFS varies approximately inversely with the observed streamflow. In general, the DFS values are similar between adaptive CBEnKF and EnKF but differ significantly in high flow periods where α is larger. In such periods, the DFS of adaptive CBEnKF is significantly smaller than that of EnKF due to the larger forecast error covariance resulting from the accounting of CB. The DFS plots show that the observations contribute about 1.5 to 2.5 independent quantities of information in low flow periods but the contribution drops by about 1 to 2 as the hydrograph rises and reaches peak. The above indicates that, in high flows, the marginal information content of the observations is reduced by about a half of that in low flows. Figure 4-7 shows a pronounced example of the above observation where the DFS of adaptive CBEnKF drops to only about 0.1 from about 1.85 as the streamflow rises. Note in the figure that α spikes up when the observed hydrograph begins to rise quickly, reducing the DFS of adaptive CBEnKF well below that of EnKF. It is this CB awareness that allows the adaptive CBEnKF analysis to continue to track the observed flow whereas the EnKF analysis is not able to. Note also in Figure 4-7 that, over the crest of the first peak, the α values fluctuate greatly. This is because, once the CB penalty is invoked in the preceding assimilation cycle, the updated model states are able to produce a realistic streamflow simulation in the current assimilation cycle without the aid of the CB penalty. This positive impact, however, is often short-lived and CB-aware DA again invokes the CB penalty, and the process repeats. Figure 4-4 to Figure 4-7 indicates that, over the crest, the CB penalty is invoked less frequently or not invoked. This is because the model states tend to maintain a temporary steady state over peak flow. Once corrected for CB, no subsequent corrections for CB are necessary until the model exits the temporary steady state.

4.4.2 Ensemble Mean Streamflow Prediction

Figure 4-8 shows the RMSE vs. lead time for all 6 basins. These results include all streamflow observations within the partial-duration time series of all events whose peaks exceed 100 cms. There are three pairs of curves for each basin, each pair consisting of the RMSE vs. lead time curves for adaptive CBEnKF (red) and EnKF (blue). The solid, dotted and dashed lines denote the

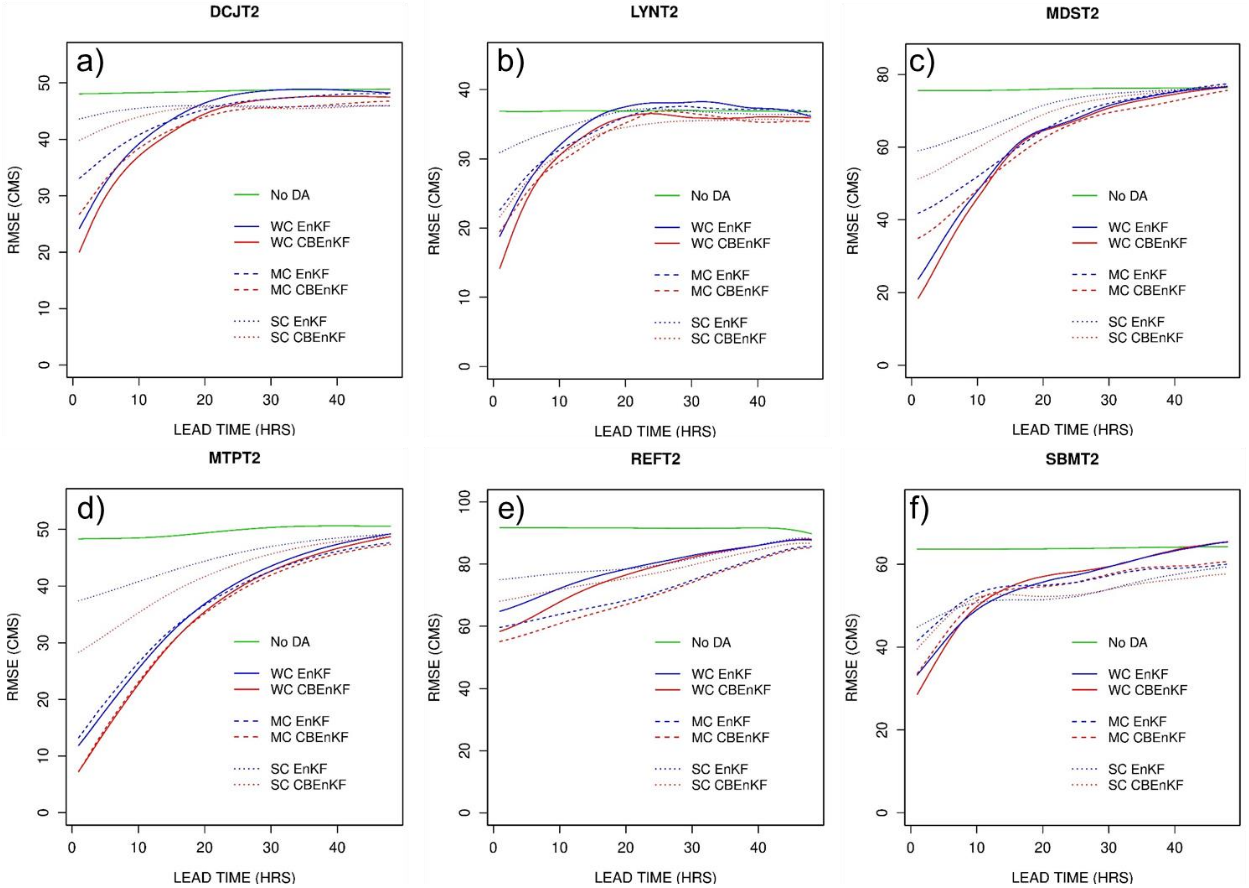


Figure 4-8 RMSE of CBEnKF (red), EnKF (blue) and No DA (green) predictions vs. lead time for the 6 basins. The solid, dotted and dashed lines denote the WC, MC and SC results

WC, MC and SC results (see Section 3.1). The solid green line shows the DA-less ensemble prediction result. For WC, the reduction in RMSE by adaptive CBEnKF over EnKF at lead time of 1 hr (the first data point) is about 20.8, 29.9, 24.5, 44.1, 15.4 and 25.3% with sample sizes of 2236, 1560, 4389, 2964, 1918 and 2840 for DCJT2, LYNT2, MDST2, MTPT2, REFT2 and SBMT2, respectively. The above reductions in RMSE translates into gains of a few to several hours in lead time. Figure 4-8 indicates that, in general, the potency of DA increases as the level of hydrologic uncertainty modeling increases, i.e., WC performs better than MC which performs better than SC. For DCJT2, LYNT2 and MDST2, the potency of DA seen in Figure 4-8 follows the above general pattern. For MTPT2, which is the most humid basin (see Figure 4-1, Table 4-1), little difference is seen in performance between WC and MC, i.e., the inclusion of the additive runoff error, X_k^R (or

$w(\tau)$, in Eq. (5)), has little positive impact. It suggests that, for MTPT2, the SAC-simulated TCI has little bias, and that the positive impact of DA comes from reducing biases in MAP and MAPE and providing the SAC with higher-quality ICs at the beginning of the assimilation window. That MTPT2 shows the largest positive impact of DA among the 6 basins is a reflection of the large storage of water in this humid basin and the resulting stronger hydrologic memory.

For REFT2, which is the largest of the 6 basins, MC outperforms WC, i.e., the additive error in TCI negatively impacts DA performance. For SBMT2, WC performs the best at short lead times but is inferior to MC and SC at longer lead times where SC is better than MC. The above observations for SBMT2 indicate that DA is able to address the input uncertainty but not the hydrologic uncertainty. As explained in Section 4.3.1, X_k^R is modeled as temporally uniform over the entire duration of the UH (see Figure 4-3). It is likely that the above simplification introduces unintended nonlinear distortions in the temporal distribution of SAC- simulated runoff within the assimilation window. Further investigation is needed, however, to ascertain the above conjecture. The study basins are relatively large and heavy convective precipitation events are common. As such, timing errors in streamflow simulation at the outlet tend to vary significantly from event to event. Liu et al. (2011) reported that REFT2 and SBMT2 have significant timing errors for which the model simulation tends to trail the observed. In such cases, DA is likely to over-adjust the control variables when the catchment starts to respond to precipitation. Such over-adjustment makes reducing hydrologic uncertainty particularly challenging and tends to result in phase errors in DA-aided predictions. In Figure 4-8, the local concavity in the RMSE vs. lead time curves for SBMT2 around the lead time of 10 hours is a symptom of such over-adjustment. Though not very pronounced, the slight concavity in the WC and MC results of LYNT2 in Figure 4-8b also suggests timing errors. Figure 4-8 shows that, regardless of the level of hydrologic uncertainty modeled or the apparent severity of timing errors, adaptive CEnKF improves over EnKF consistently and significantly. We also examined the comparative performance of adaptive CBEnKF with EnKF for ME, CORR, MAXEU and MAXEO. Qualitatively, the performance characteristics of adaptive CBEnKF and EnKF are similar but CBEnKF generally outperforms EnKF under each of the above criteria. The above indicates that adaptive CBEnKF improves the overall quality of ensemble mean prediction, in agreement with the ensemble verification results (Lee et al., 2019).

We now turn out attention to comparative evaluation of adaptive CBEnKF conditioned on α. If the adaptive extension prescribes α skillfully, one may expect CBEnKF to outperform EnKF by larger margins than seen in Figure 4-8 when only the predictions associated with larger α are considered. This conditional performance is particularly important from the perspective of flood forecasting in that, the larger the verifying observed flow is, the larger the CB and hence α tend to be (see Figure 4-4 to Figure 4-7). Figs 4-9a,b,c show the RMSE vs. lead time plots for DCJT2, MDST2 and MTPT2 conditioned on $\alpha \geq 0.65$, respectively. These basins are chosen because they are least impacted by timing errors and hence are better-suited for the conditional evaluation. The threshold value of 0.65 represents the largest α that yielded sample size of at least 30 (31, 502 and 205 for DCJT2, MDST2 and MTPT2, respectively). Predictions associated with large α tend occur more frequently over the receding limb than over the rising limb because of the longer duration of the former. As such, the verifying observed flow for these predictions tends to decrease as the lead time increases, which explains why the RMSE associated with the No DA results in Figure 6abc decreases as the lead time increases. Comparison of Figure 4-9 with Figure 4-8 shows that the margin of improvement by adaptive CBEnKF over EnKF is significantly larger when a large CB penalty is invoked. The reduction in RMSE by adaptive CBEnKF over EnKF for 1 hr-ahead prediction is approximately 24.0, 36.4 and 57.3% for DCJT2, MDST2 and MTPT2, respectively. Expectedly, the margin of improvement decreases as the conditioning α decreases. For $\alpha = 0$, the adaptive CBEnKF results are very close to the EnKF results but still slightly superior due to the cumulative positive impact of adaptive CBEnKF in the preceding assimilation cycles. Figs 4-9d,e,f show the RMSE vs. lead time plots for DCJT2, MDST2 and MTPT2 for verifying observed flow exceeding 200 cms regardless of α. The relative reduction in RMSE by adaptive CBEnKF over EnKF is 29.1, 46.9 and 57.0% for ensemble mean analysis associated with Figs 6def for DCJT2, MDST2 and MTPT2, respectively. Figs 4-9d,e,f are to be compared with Figs 4-8a,c,d, respectively, which show the same results but for all ranges of verifying observed flow. Note that the y-

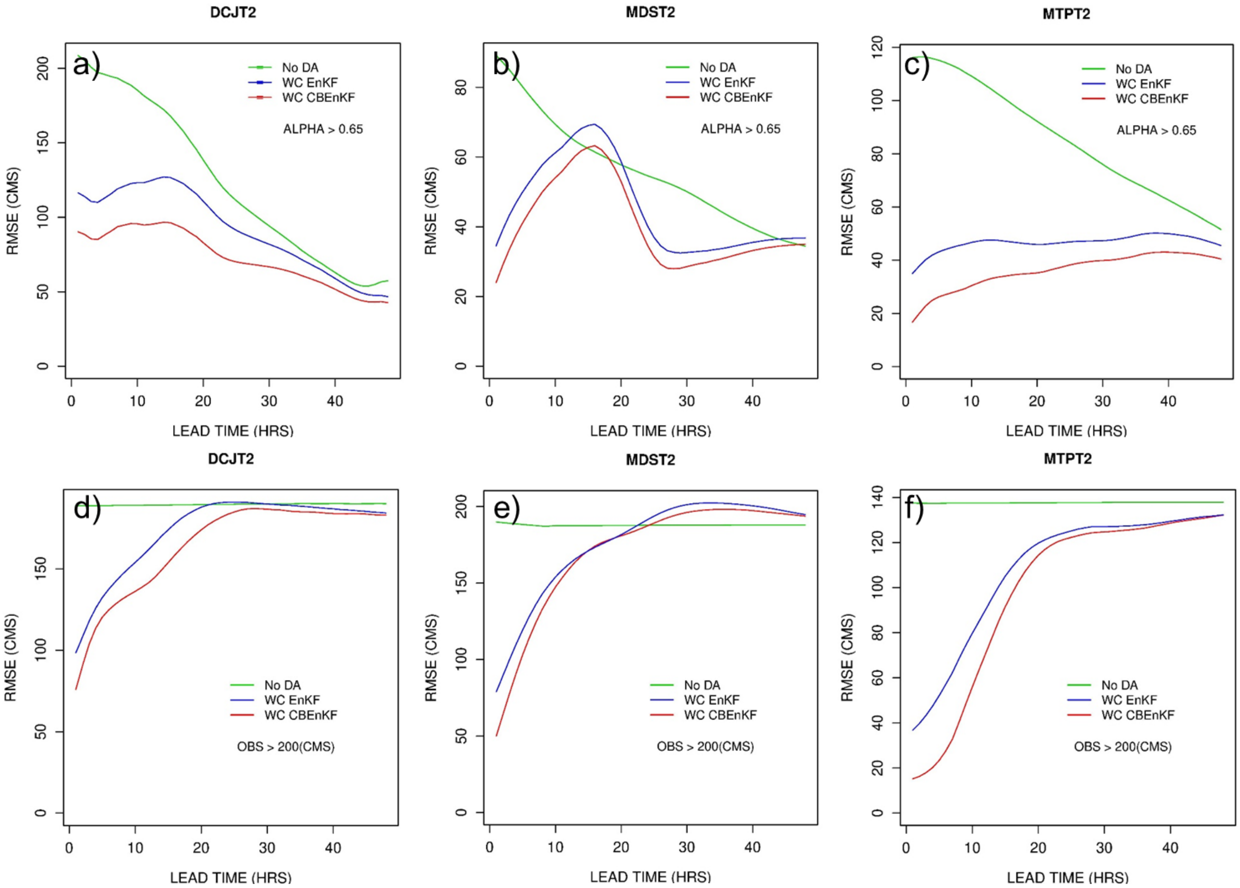


Figure 4-9 RMSE of CBEnKF (red) and EnKF (blue) and No DA (green) predictions for a) DCJT2, b) MDST2 and c) MTPT2 conditioned on $\alpha > 0.65$, and d) DCJT2, e) MDST2 and f) MTPT2 conditioned on the verifying observed flow exceeding 200 cms

axis range varies from panel to panel in these plots. It is seen that the marginal improvement by adaptive CBEnKF over EnKF is larger when only high flows are considered, in agreement with Lee et al. (2019).

Finally, we examine the comparative performance of adaptive CBEnKF for rising vs. falling limbs. Given that the rate of change in streamflow is generally larger in the rising limb than in the falling limb, one may hypothesize that CB, and hence the positive impact of CB-aware DA, may be larger in the rising limb. To test this, we conditioned the predictions on increasing vs. decreasing observed flow as a proxy for rising vs. falling limbs. Admittedly, this stratification is only approximate in that streamflow may increase and decrease within falling and rising limbs,

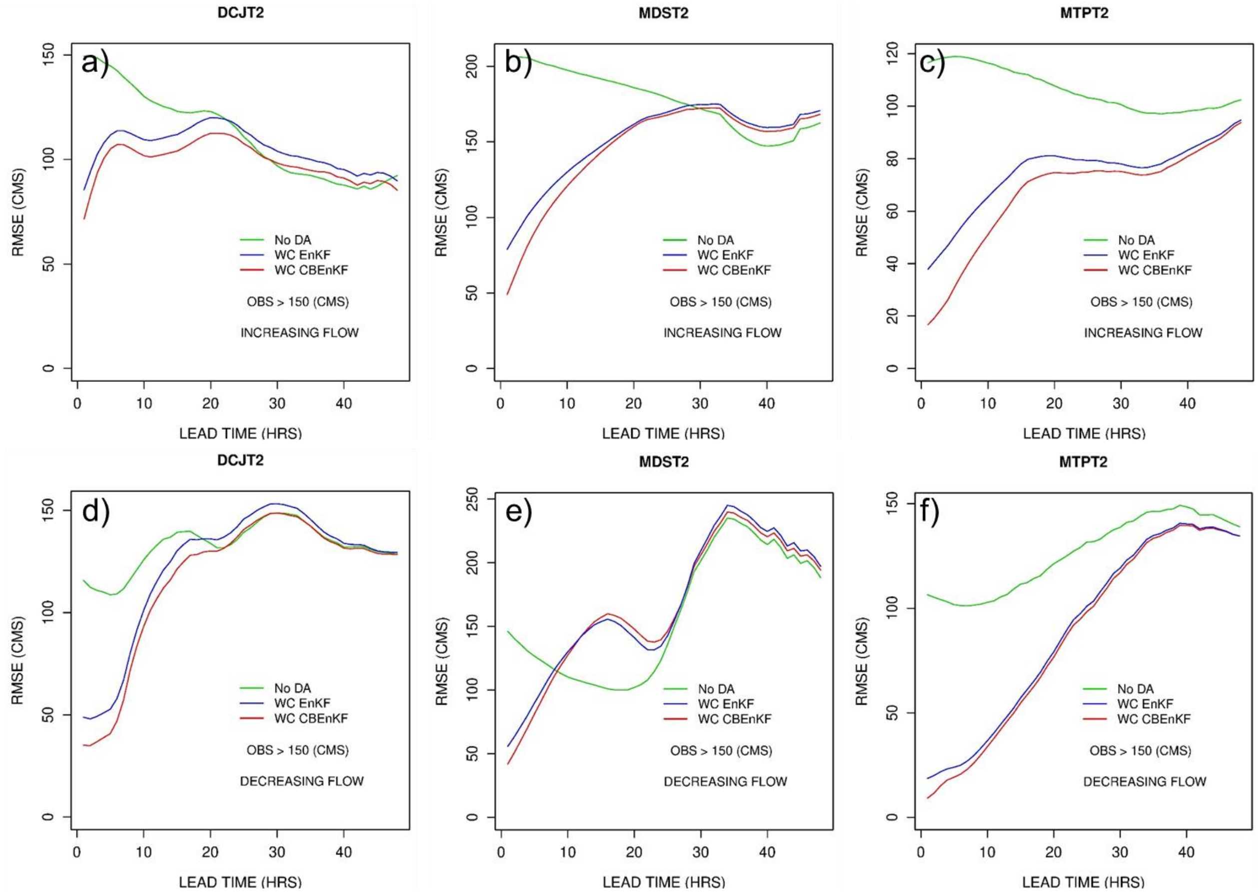


Figure 4-10 RMSE of CBEnKF (red) and EnKF (blue) and No DA (green) predictions conditioned on the verifying observed flow exceeding 150 cms for a) DCJT2, b) MDST2 and c) MTPT2 for increasing observed flow and d) DCJT2, e) MDST2 and f) MTPT2 for decreasing observed flow.

respectively. Figure 4-10 shows the RMSE vs. lead time plots for DCJT2, MDST2 and MTPT2 conditional on the verifying observation exceeding 150 cms. For each pair, the upper and lower panels are further conditioned on increasing vs. decreasing flows, respectively. Because the magnitude of flow is generally smaller for decreasing flow and the sample size is not the same, head-to-head comparisons are not readily possible. Nevertheless, it may be seen in Figure 7 that the marginal improvement by adaptive CBEnKF over EnKF tends to be larger over the rising limb than over the falling limb. Not surprisingly, the rising-vs.-falling limb results for the timing error-impacted basins, LYNT2, REFT2 and SBMT2, are unclear as they are masked by varying degrees of phase errors in the DA results. Additional research is needed for more rigorous evaluation based on a large number of basins in different hydroclimatological regions.

4.5 Conclusions and future research recommendation

Floods are associated with extreme states of the hydrologic system which can only be observed with significant uncertainty. As such, data assimilation (DA) for flood forecasting is subject to conditional bias (CB) which impacts optimal information fusion. This work describes CB-aware DA based on an adaptive extension of the conditional bias-penalized ensemble Kalman filter (CBEnKF) and comparatively evaluates with the ensemble Kalman filter (EnKF) for 6 headwater basins in Texas using the NWS's operational lumped hydrologic models. CB-aware DA and the degrees of freedom for signal are then used to quantify the marginal information content of the observations. The main findings, conclusions and future research recommendations follow below.

CB arises very frequently in varying magnitude when assimilating streamflow observations during the catchment's response to precipitation and drainage. In general, the larger the discharge is, the larger the CB is. Adaptive CEnKF improves over EnKF consistently and significantly regardless of the level of hydrologic uncertainty modeled. The improvement is particularly large during sharp rises of the outlet hydrograph with large peak flow. CB-aware DA is hence likely to have the largest positive impact for forecasting of rising flood flows in fast-responding basins. The flow-dependent marginal information content in the observations varies with the streamflow response of the catchment and the magnitude of CB, and tends to decrease and increase in the rising and falling phases of the hydrograph, respectively. The overall findings indicate that CB-aware DA with information content analysis offers an objective framework for dynamically balancing the flow-dependent predictive skill of hydrologic models, quality and frequency of hydrologic observations, and scheduling of the DA cycle toward cost-effective improvement of operational flood forecasting. Timing errors remain a very significant challenge in streamflow DA. Novel, phase error-agnostic approaches such as multi-scale bias correction (Noh et al., 2018) warrant further attention. Additional research is also needed toward computationally less expensive and algorithmically simpler alternatives and approximations to CBPKF and CBEnKF such as variance-inflated KF (VIKF) and its ensemble extension (Shen et al., 2019). In this work, α was optimized in less than an efficient manner. Additional research is needed to utilize the closed-form expression for $\partial X_{k|k}/\partial\alpha$ available in VIKF for efficient gradient-based minimization.

Chapter 5. DA for WRF-Hydro for Event-Based Urban Flash Flood Forecasting

5.1 Introduction

With the implementation of the National Water Model (NWM), the National Weather Service (NWS) has made a step-change advance in operational water forecasting by enabling high-resolution (1 hr, 1 km for land surface and 250 m for routing) hydrologic modeling across the US (NWS 2020). As a part of the NWM initiative, the NWS has been mandated to provide forecasts at even higher spatiotemporal resolutions when and where such information is demanded such as in large urban areas for flood warning, and areas of high-value infrastructure, susceptible to landslides, or impacted by forest fires (Graziano et al., 2017). The value of high-resolution products and services depends not only on the hydrologic and hydraulic models but also on the quality of the forcings, model parameters, initial conditions (IC) and boundary conditions at the commensurate resolutions. In the DFW area, the Collaborative Adaptive Sensing of the Atmosphere (CASA) Program operates a network of X-band radars to provide a suite of meteorological, hydrometeorological and hydrologic products for severe weather and flash flood monitoring and prediction (Chandrasekar et al., 2013). The network currently consists of 7 radars located at Addison, Arlington, Cleburne, Denton, Fort Worth, Mesquite and Midlothian, TX. A salient feature of the above operation is that the radar rainfall data are available at a very high resolution of 500 m and 1 min. The CASA quantitative precipitation estimates (QPE) are currently input to the NWS Hydrology Laboratory-Research Distributed Hydrologic Model (HL-RDHM, (Koren et al., 2004; NWS, 2008) to produce a suite of hydrologic products at the same resolution in real time (Rafieeinasab et al., 2015a; Habibi et al., 2016; Habibi and Seo, 2018). The characteristic spatial scale of natural and man-made physiographic features in the study area suggests that a further increase in hydrologic model resolution may improve the information content of the model output (Habibi et al., 2019). There is also an ever increasing demand for higher resolution hydrologic products for enhanced spatio-temporal specificity. The purpose of this work is to assess using WRF-Hydro how increasing the resolution of hydrologic modeling, calibration of selected model parameters and assimilating locally-available observations of precipitation and streamflow may improve flood modeling and prediction toward high-resolution water forecasting in urban areas.

Real-time continuous operation of high-resolution models is computationally very expensive particularly for large areas (Habibi et al., 2019). A more practical approach is likely to be event-based operation with robust initialization. As such, our assessment is carried out in the context of event-based modeling and prediction. The event-based paradigm meant that most conventional calibration methods, which rely on time-continuous observations of precipitation and streamflow, and sequential DA methods, which employ recursive state updating, may not be applicable or desirable. To that end, we employ multi-event averaging of event-specific parameter optimization results for calibration and reduced-rank fixed-lag smoothing for DA. The new contributions of this chapter are: selective calibration of WRF-Hydro for urban flood modeling and prediction, improving simulation of highly peaked hydrographs with the addition of a conditional bias (CB) penalty, and assessment of the impacts of different spatio-temporal resolutions of rainfall-runoff and routing models, of ICs and land cover, and of assimilation of streamflow observations for initialization of WRF-Hydro toward event-based operation of high-resolution urban flood prediction. This chapter is an adaptation of Kim et al. (2020) and is organized as follows. In Section 5.2, we describe the study area, data used and the hydrologic models used. Section 5.3 describes the methods used in the experiment design, calibration and DA. Section 5.4 describes the experiments and presents the results. Section 5.5 provides the conclusions and future research recommendations.

5.2 Study area, data and hydrologic models used

Here we describe the study area, data used and hydrologic models used.

5.2.1 Study area

The study area comprises the Johnson (40.2 km^2), Cottonwood (32.3 km^2) and Fish (54.6 km^2) Creek Catchments in the Cities of Arlington and Grand Prairie in the Dallas-Fort Worth

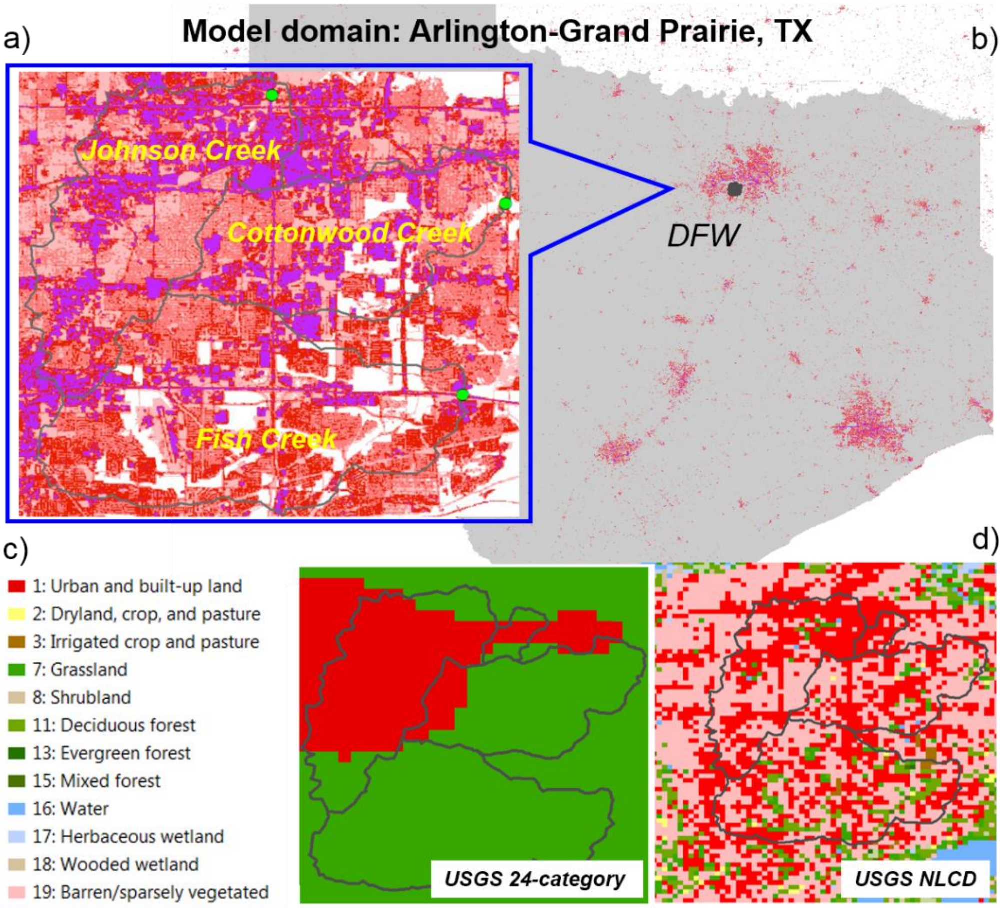


Figure 5-1 a) The 3-basin study area with commercial impervious (purple) and high-density developed (red) areas in the background. b) State-wide view of the study area. c) USGS 24-category and d) NLCD land cover in the study area

(DFW) area of TX (see Figure 5-1a,b). These basins have been used in previous studies of high-resolution hydrologic modeling and sensing (Rafieeinasab et al., 2015a; Habibi et al., 2016, 2019; Norouzi, 2016). The Johnson, Cottonwood and Fish Creek Catchments, referred to herein as JC, CC and FC, respectively, are highly urbanized with impervious fractions of 0.48, 0.37 and 0.31, respectively (Habibi et al., 2019), see Figure 5-1a). Hydroclimatologically, the study basins are particularly challenging for hydrologic modeling and prediction due to very short memory in the surface and soil water storages. Recent assessment of the streamflow prediction skill of the NWS operational hydrologic models indicates that the study region has the smallest predictability among the 138 basins assessed in 8 different River Forecast Centers' (RFC) service areas across large sections of the US (Alizadeh et al., 2020).

5.2.2 Data used

The CASA QPE products have been extensively evaluated (Chandrasekar et al., 2012; Cifelli et al., 2018). Comparative evaluation of different radar-based QPE products (Rafieeinasab et al., 2014a, 2015b) showed that the CASA QPE is generally more accurate for larger precipitation amounts in the study area whereas the Multisensor Precipitation Estimator (MPE) (Seo et al., 2010) estimates do better for smaller amounts. The CASA QPE operation recently began fusing the QPE

Table 5-1 List of rainfall events used.

Event	Event total mean areal rainfall (mm)			Period of record	Duration
	JC[1]	CC[2]	FC[3]		
Jan 2017	75.8	90.8	71.6	00:00Z 01/16/2017 - 23:59Z 01/17/2017	48 hrs
Feb 2018	95.2	93.7	100.5	00:00Z 02/20/2018 - 07:59Z 02/21/2018	32 hrs
Sep 2018	97.6	103.1	131.9	12:00Z 09/21/2018 - 19:59Z 09/22/2018	32 hrs
Apr 2019	31.5	33.5	27.1	00:00Z 04/17/2019 – 11:28Z 04/18/2018	35 hrs
May 2019	56.5	60.1	62.5	00:00Z 05/08/2019 – 03:43Z 05/09/2019	28 hrs

[1] Johnson Creek Catchment
[2] Cottonwood Creek Catchment
[3] Fish Creek Catchment

from the X-band radar network with that from the WSR-88D in Burleson, TX (Chen and Chandrasekar, 2015). The rainfall estimates used in this study are the resulting fused QPE product. For details, the reader is referred to Chandrasekar (2017).

Because the CASA network has been in continuous operation only for several years, a long period of time-continuous data is not available. In this study, we used the 5 recent events of varying magnitude listed in Table 5-1. Figure 5-2 shows the total rainfall maps for the 4 largest events. All

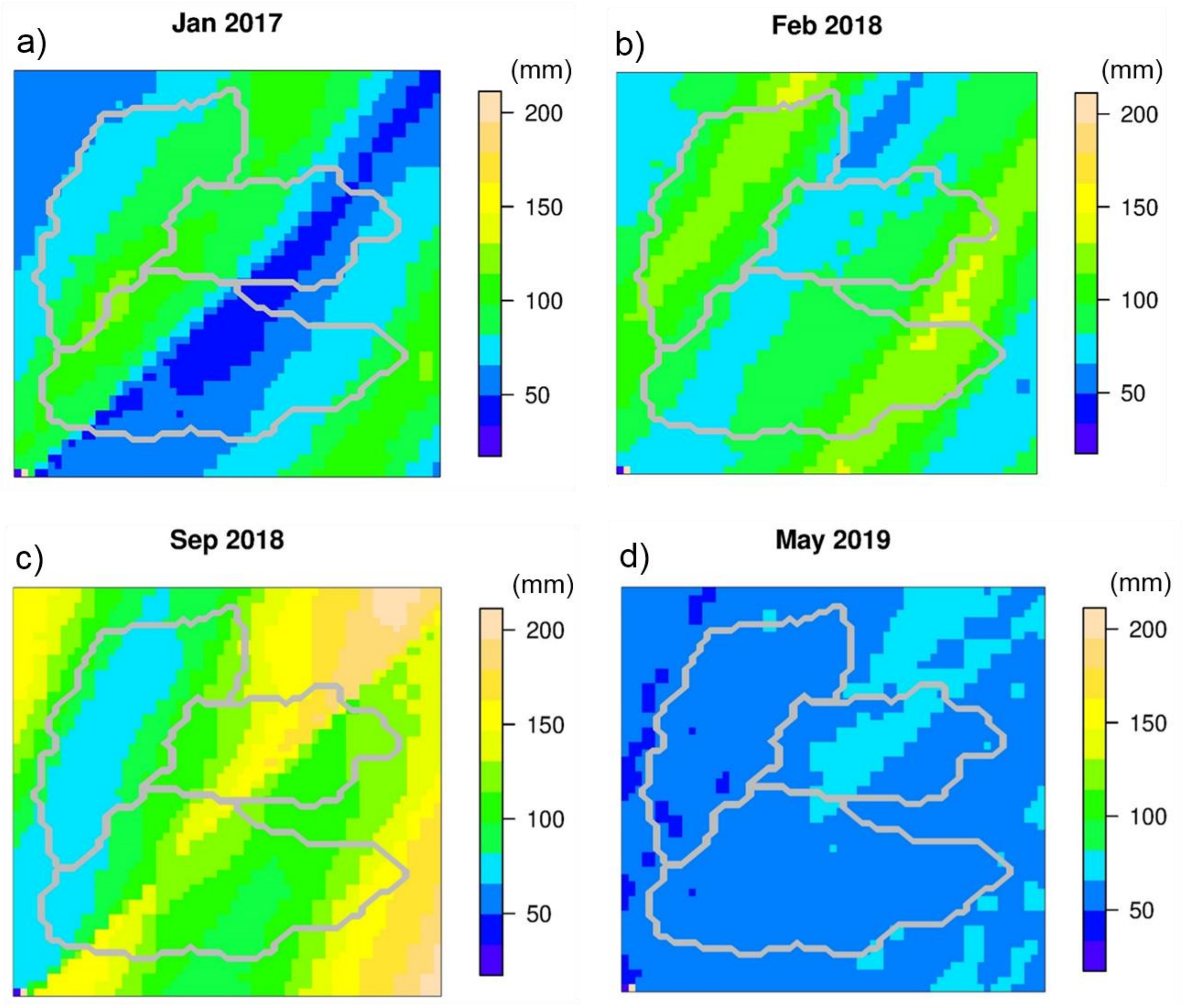


Figure 5-2 Event total rainfall maps (in mm) for the a) Jan 2017, b) Feb 2018, c) Sep 2018 and d) May 2019 events

other forcings for WRF-Hydro are from the near real-time North American Land Data Assimilation System (NLDAS) Phase 2 forcing and model output produced operationally at the Environmental Modeling Center of the NOAA/NWS/National Centers for Environmental Prediction (Cosgrove et al., 2003). Networks of ALERT sensors operated by the Cities of Arlington and Grand Prairie provide water level observations in the study area including at the catchment outlets. The observations are based on pressure transducers located at the channel bottom. To estimate discharge from stage observations, we used rating curves derived by Norouzi (2016) at the outlets of the 3 catchments (see Figure 5-1a) based on the numerical modeling approach of Kean and Smith (2004, 2005, 2010).

5.2.3 Hydrologic model used

The hydrologic model used is WRF-Hydro Version 5.0.2. (Gochis et al., 2018). For urban flood modelling, the most important components are the rainfall-runoff, terrain, or hillslope, routing and channel routing models. Below, we describe only the core model dynamics that are directly relevant to the development of this work.

5.2.4 Rainfall-runoff model

The rainfall-runoff option used in this work is the Simple Water Balance model (SWB) of Schaake et al. (1996) which is used by the NWM also. As in Moore (1985) and the SCS curve number method (USDA, 1986), the SWB models the average runoff depth over a grid box or a catchment, Q_s, as (Schaake et al., 1996):

$$Q_s = \frac{P_x^2}{(P_x + I_c)^2} \tag{5.1}$$

where P_x and I_c denote the average precipitation depth and infiltration capacity over the grid box. The infiltration capacity, I_c, in Eq. (5.1) is modeled as (Schaake et al., 1996):

$$I_c = D_x(1 - e^{-kt}) \tag{5.2}$$

Where D_x denotes the maximum water holding capacity of the soil column, k denotes the decay coefficient and t denotes the time elapsed. Eq. (5.2) is analogous to the potential infiltration depth, F, of the Horton infiltration model (Horton, 1941) without the constant infiltration rate due to gravity:

$$F = \frac{f_o}{k}(1 - e^{-kt}) \tag{5.3}$$

where f_0 denotes the initial potential infiltration rate due to suction pressure and k denotes the decay rate. One may hence interpret the maximum soil water holding capacity, D_x, as representing f_0/k in Eq. (5.2) where $1/k$ represents the time scale of decay of potential infiltration rate. The maximum water holding capacity D_x in Eq. (5.2) is modeled as (Schaake et al., 1996):

$$D_x = \sum_{i=1}^{4} \Delta Z_i(\theta_{sat} - \theta_i) \tag{5.4}$$

where ΔZ_i denotes the thickness of the i-th soil layer, θ_{sat} denotes the saturation soil water content (i.e., porosity) and θ_i denotes the initial soil water content in the i-th soil layer. Eq. (45.) is analogous to the total infiltration depth in the Green-Ampt infiltration equation (Green and Ampt, 1911):

$$F = Z_f(\theta_{sat} - \theta_{init}) \tag{5.5}$$

where Z_f denotes the depth to the wetting front and θ_{init} denotes the vertically uniform initial soil water content. As shown above, the surface runoff component of the SWB may be considered as a combination of the SCS method for runoff ratio and the Horton infiltration equation without the gravity term for time decay in potential infiltration rate in which the maximum water holding capacity is prescribed by the depth-integrated soil pore space given the antecedent soil water content. The study area is highly urbanized. Accurate high-resolution depiction of land cover is hence very important (Rafieeinasab et al., 2015a; Habibi et al., 2016; Norouzi, 2016). WRF-Hydro uses the United States Geological Survey's (USGS) 24-category land cover product (Loveland et al., 1995), see Figure 5-1c) to parameterize the Land Surface Model (LSM). In this work, we use the USGS's National Land Cover Database (Wickham et al., 2020) for higher resolution depiction (see Figure 5-1d) and compare with the USGS 24-category land cover.

5.2.5 Terrain routing model

The terrain, or hillslope, routing option used in this work is the diffusive wave model. The mass balance equation is given by:

$$\frac{\partial h}{\partial t} + \frac{\partial q_x}{\partial x} + \frac{\partial q_y}{\partial y} = i_e \tag{5.6}$$

where h denotes the water depth, q_x and q_y denote the specific discharge along the x- and y-directions, respectively, and i_e denotes the excess precipitation, or surface runoff depth, given by the rainfall-runoff model. Though expressed as a 2D model, Eq. (5.6) is solved only along the steepest-descending direction, referred to as the D8 option in WRF-Hydro (Gochis et al., 2018). The momentum balance equation is given by:

$$-\frac{\partial h}{\partial x} + S_{ox} = S_{fx} = \left(\frac{n_{ov} q_x}{h^{\frac{5}{3}}}\right)^2 \tag{5.7}$$

where S_{ox} denotes the terrain or channel bed slope, S_{fx} denotes the friction slope and n_{ov} denotes the Manning's friction coefficient for the hillslope. The last equality in Eq. (5.7) follows from the Manning's equation under the wide channel assumption (Akan and Houghtalen, 2013). In WRF-Hydro, S_{ox} is calculated based on the DEM data and n_{ov} is prescribed according to land cover. As such, the choice of the land cover data impacts terrain routing.

5.2.6 Channel routing model

The channel routing option used in this work is the gridded diffusive wave model which solves the following mass and momentum balance equations:

$$\frac{\partial A}{\partial t} + \frac{\partial Q}{\partial x} = q_l \tag{5.8}$$

$$-\frac{\partial h}{\partial x} + S_o = S_f = \left(\frac{nQ}{AR^{\frac{2}{3}}}\right)^2 \tag{5.9}$$

where A denotes the wetted channel cross-sectional area, Q denotes the flow rate, q_l denotes the lateral inflow from Eqs. (5.6) and (5.7), h denotes the water depth, S_o denotes the channel bed slope, n denotes the Manning's roughness coefficient for the channel bed and R denotes the hydraulic radius of the channel cross section. The resulting finite difference equation is solved iteratively using the Newton-Raphson method (Gochis et al., 2018). The channels are delineated based on the National Hydrographic Dataset Plus Version 2 (NHDPlusV2, (Moore et al., 2019). The channel routing model assumes trapezoidal cross section for which two additional parameters, the channel bottom width and side slope, are necessary:

$$Q = \frac{1}{n} AR^{\frac{2}{3}} S_f^{\frac{1}{2}} = \frac{1}{n} \frac{\left((B_w + zh)h\right)^{\frac{5}{3}}}{\left(B_w + 2h\sqrt{1+z^2}\right)^{\frac{2}{3}}} S_f^{\frac{1}{2}} \tag{5.10}$$

where B_w denotes the channel bottom width, z denotes the channel side slope and h denotes the water depth. WRF-Hydro prescribes the above parameters stream order-specifically, i.e., channels of the same Strahler stream order share the same parameter values for channel routing (Gochis et al., 2018).

5.3 Methods

To assess how the resolution of hydrologic modeling, calibration, and DA may impact urban flood modeling and prediction using WRF-Hydro, we designed and carried out a set of simulation experiments. In this section, we describe the experiment design, calibration and DA.

5.3.1 Design of experiments

Table 5-2 shows the combinations of resolutions considered in this work. The CASA QPE is available at 500 m 1 min resolution. Rafieeinasab et al. (2015a) report that a resolution of 500 m and 15 min or higher is necessary for streamflow prediction at the outlets of the study basins using CASA QPE and HL-RDHM (Koren et al., 2004). To assess how higher spatial resolution of hydrologic modeling may improve flood simulation in the study area, we disaggregate the 500 m QPE to QPEs at nominal resolutions of 250 m and 125 m by remapping the CASA QPE on a lat-lon grid to a Lambert conformal conic grid for ingest by WRF-Hydro. For the remapping, we used the conserve method available for ESMF (NCAR 2020). In addition, to assess possible gains from higher temporal resolution modeling, we aggregated the native resolution 1 min CASA QPE to 10 min accumulations. With the above choices, the LSM was run at 3 different spatial resolutions of 500, 250 and 125 m with a common native temporal QPE resolution for the spatial resolution experiment, and at two different temporal resolutions of 1 and 10 min with a common spatial resolution of 250 m for the temporal resolution experiment. In the above experiments, the resolution of the routing models was fixed at 250 m. The limited number of combinations of resolutions

Table 5-2 Combinations of spatio-temporal resolutions used.

	QPE	Rainfall-runoff	Terrain and channel routing
Spatial	125, 250, 500 m (all at 1 min resolution)	125, 250 m	50, 125, 250 m
Temporal	1, 10 min (both at 250 m resolution)	1 min timestep	15 sec timestep

represent a compromise between the computational requirements and the range of resolutions that are most likely to be of operational interest in the study area.

It was observed in the early stages of the spatial resolution experiment that the mean areal precipitation (MAP) calculated at 500 m resolution is significantly different from that at 250 m or 125 m. The differences were traced to the coarseness of 500 m grid boxes in delineating small catchments in WRF-Hydro. Significant errors in precipitation volume often translate into significant errors in peak flow and time-to-peak flow. As such, we excluded 500 m resolution from further consideration. For routing, we initially considered 25 m resolution as well. It was discovered in the early stages, however, that the number of stream segments at this resolution for the study domain exceeds the maximum allowed by WRF-Hydro. For this reason, we excluded 25 m from further consideration for routing. Though limited in number, the resulting combinations allow comparisons of the LSM resolutions of 250 m and 125 m given the common routing model resolution of 125 m and of the routing model resolutions of 250 m, 125 m and 50 m given the common LSM resolution of 125 m.

5.3.2 Calibration

WRF-Hydro employs a large number of parameters for rainfall-runoff and routing modeling. Most of them are modeled as spatially-varying and specified by spatial maps or lookup tables of the relevant physiographic variables. Due to the computational cost, it is impractical to calibrate a large number of parameters. The approach taken in this work is to identify only the most influential and adjust them up or down with multiplicative scaling factors over the entire catchment, thus maintaining the prescribed spatial variations and physiographic relationships (Gupta et al., 2003). Examination of the model physics described in Eqs. (5.1) through (5.10) indicates the most influential parameters for the rainfall-runoff and routing models are likely to be the potential infiltration rate decay coefficient k in Eq. (5.2), the Manning's friction coefficient for overland flow, n_{ov}, in Eq. (5.7) and the 4 channel routing parameters of the Manning's friction coefficient n, the bottom width, B_w, the side slope z, and the initial water depth, h. The above 6 parameters, k, n_{ov}, n, B_w, z and h, are denoted in WRF-Hydro as $refdk$, sfc_rough, $rmannn$, bw, $chsslp$ and $hlink$, respectively, which are used below. Extensive sensitivity analysis involving all rainfall-runoff and routing parameters confirm the above choices. The decay coefficient k in Eq.(5.2) is coded in WRF-Hydro as:

$$k = \left(\mathrm{REFKDT}\frac{DKSAT}{REFDK}\right) \cdot \left(\frac{DT}{86400}\right) \tag{5.11}$$

where $DKSAT$ denotes the saturated hydraulic conductivity, $REFDKDT$ and $REFDK$ are parameters for surface runoff (Gochis et al. 2018), and DT denotes the time step in seconds. Both $REFKDT$ and $REFDK$ are calibratable parameters. Because adjusting $REFDKT$ has the same effect as adjusting $REFDK^{-1}$ for k, it is not necessary in practice to calibrate both. As such, we calibrate only $REFDK$ in this work. Note in Eq. (5.11) that, if $REFDK$ increases or decreases, k decreases or increases and hence the infiltration capacity decreases or increases given the maximum water holding capacity, D_x, respectively. Accordingly, one may consider $REFDK$ as controlling the runoff ratio. All other parameters in the LSM are set to the WRF-Hydro default (Gochis et al., 2018).

For the terrain routing model, n_{ov} is by far the most important. In WRF-Hydro, n_{ov} is prescribed according to the USGS 24-category land cover (Loveland et al., 1995). In this work, we use the National Land Cover Database (NLCD, (Wickham et al., 2020) and the same default land cover-dependent values of n_{ov}. In the calibration process, we apply a single multiplicative adjustment factor to the spatially varying n_{ov} for the entire catchment. Calibration of channel routing parameters presents a particular challenge as elaborated below. There are a total of 4 parameters, B_w, z and n, and the initial condition, h, to be determined in the calibration process whereas the only source of information available is observed streamflow at the catchment outlet. For most natural channels, the cross sections are not trapezoidal. It is hence difficult to prescribe B_w and z externally based on physiographic information particularly for small streams. Given the above picture, we opted to assess first the impact of changes in each channel routing parameter via a series of idealized sensitivity analysis using the recently developed general analytical solution for nonlinear reservoir (Nazari and Seo, 2020). In this analysis, we prescribe an impulse as the upstream hydrograph and route it through a nonlinear reservoir which is modeled as a hydraulically-equivalent trapezoidal channel as in WRF-Hydro. We then visually examine the shape of the downstream hydrographs and assess the impact of changes in each of the 4 parameters to the downstream hydrograph. The results indicate that changes in each of the 4 routing parameters often produce similar effects, that the shape of the outlet hydrograph is least sensitive to changes in z and that, in addition to n, both B_w and h shape the outlet hydrograph to a significant degree, in particular, the upper and lower parts of the falling limb. The above findings suggest that one may be able to prescribe z externally and calibrate only the other three. In this work, we chose to calibrate all 4 parameters to assess empirically the degree of under-determinedness in each.

For calibration, we initially considered the Shuffled Complex Evolution (SCE) (Duan et al. 1992) and the Stepwise Line Search (SLS) (Kuzmin et al., 2008). Due to excessive computational requirement of SCE, however, we chose SLS as the main calibration technique (see Kuzmin et al. 2008 for comparison). Once the parameter space is defined, we use Latin Hypercube sampling (LHS) (Tang, 1993) to run WRF-Hydro with the randomly-sampled parameter values from which a small number of best-performing parameter sets is retained. We then run SLS using the parameter sets retained above as starting points, visually examine the resulting hydrographs and choose the best. The original SLS minimizes the multi-scale objective function consisting of normalized root mean square error of simulated flow at multiple time scales of aggregation such as hourly, daily, weekly, monthly, etc. The hydrologic response time of the study basins, on the other hand, is sub-daily for which the multiscale objective function is not necessary. A second modification to SLS deals with the objective function itself as elaborated below. Arguably the two most important variables for urban flood prediction are the peak flow and time-to-peak flow, i.e., the time until the peak flow occurs relative to some reference time of user’s interest. The hydrographs for the study basins are often characterized by high degrees of peakedness due to fast surface runoff over urban and semi-dry land surfaces. Commonly used objective functions for calibration such as the mean squared error (MSE) of simulated flow or its variable transform is not very effective in simulating very sharp peaks due to the typically very small number of observations associated with peak flows. To address the above, we combine the mean error (ME), MSE and Type II conditional bias (CB) for the objective function as follows the last of which is specifically to improve simulation of peaked hydrographs:

$$J = \left(\frac{1}{n}\Sigma_{i=1}^{n} O_i - \frac{1}{n}\Sigma_{i=1}^{n} S_i\right)^2 + \frac{1}{n}\sum\nolimits_{i=1}^{n}(O_i - S_i)^2$$

$$+\alpha\frac{1}{n}\sum\nolimits_{k}^{K} n_k\left\{O_k^{mid} - \frac{1}{n_k}\sum\nolimits_{i=1}^{n_k}\left(S_i \middle| O_k^{min} \le O_i \le O_k^{max}\right)\right\}^2 \qquad (5.12)$$

where O_i and S_i denote the observed and simulated flows at timestep i, n denotes the total number of $\{O_i, S_i\}$ pairs in the calibration period, α denotes the weight given to the conditional bias penalty term, K denotes the number of subintervals dividing the range of observed flow, O_k^{min} and O_k^{max} denote the lower and upper bounds of the k-th subinterval, n_k denotes the number of observed flow within the k-th subinterval, O_k^{mid} denotes the mid-point between O_k^{min} and O_k^{max}, i.e., $O_k^{mid} = O_k^{min} + (O_k^{max} - O_k^{min})/2$, and $S_i|O_k^{min} \leq O_i \leq O_k^{max}$ denotes the i-th simulated flow for which the verifying observed flow falls in the k-th subinterval. The three terms in Eq. (5.12) represent the ME, the MSE and the mean of the Type-II CB squared, respectively. The first term may appear redundant in that reducing CB is a sufficient condition for reducing ME. In practice, however, the CB penalty may not be effective across all ranges of flow due to small sample size in certain sub-ranges. Our experience indicates that a sub-range of 10 (cms) and $\alpha = 2$ generally yield satisfactory results for the study basins. We note here that the last two terms in Eq. (5.12) represent a sample statistic for the objective function used in CB-penalized optimal linear estimation for improved estimation of extremes (Seo, 2012; Brown and Seo, 2013; Seo et al., 2014, 2018a, 2018b; Kim et al., 2018; Jozaghi et al., 2019; Lee et al., 2019; Shen et al., 2019).

Though the number of parameters calibrated is small, it is still computationally too expensive to perform resolution-specific calibration for all combinations of resolutions (see Table 5-2). The alternative strategy adopted in this work is to calibrate using SLS-LHS at the lowest spatial resolution, i.e., 250 m for both the LSM and routing models, and use the resulting parameter values as the starting point for calibration at the next higher-resolution using SLS only. For the routing model resolution of 50 m, however, the above strategy could not be used due to excessively large computational requirements (see Table 5-3). Instead, we borrow the calibration results at 250 m LSM and 125 m routing models and assess parameter transferability from 125 m to 50 m for routing. Event-specific calibration is bound to overfit the specific event at hand. To avoid dependent evaluation based on overfitted parameters, we averaged the middle 3 parameter values out of the

Table 5-3 Wall clock times (in sec) for a 32-hr WRF-Hydro simulation[1]

Resolution (m)		Number of threads			
LSM	**Routing models**	**4**	**8**	**16**	**32**
250	250	32	18	13	11
250	125	63	37	26	22
250	50	1043	637	386	264
125	125	150	79	48	43

[1] On Intel(R) Xeon(R) Gold 6152 CPU @ 2.10GHz 44 CPU core (2 threads/core) Linux computer

5 from event-specific calibration. The rationale for dropping the largest and the smallest values is to avoid large biases arising from possible extremes. The average parameter values thus obtained are referred to as the non-event-specific calibration results.

5.3.3 Assimilation of streamflow observations

Hydrologic and hydraulic processes are heavily controlled by complex local physiography. The models may not capture the fixed boundary conditions, the ICs or the physical processes occurring over certain ranges of scale. In addition, the precipitation input may have significant systematic or random errors, or the hydrologic model may lack adequate calibration. In such situations, the model states may become too unrealistic to produce skillful predictions especially when the hydrometeorological or hydrologic conditions depart from the historically observed. For this reason, some form of state updating, manual or automatic, is generally necessary for real-time flood forecasting (WMO, 1992). With high-resolution models, however, manual DA is not viable due to the very large dimensionality (Lee et al., 2011; Lee and Seo, 2014). In this work, we assess how assimilating streamflow observations at the catchment outlet may be used to initialize WRF-Hydro for event-based prediction. For the DA method, we use the fixed-lag formulation (Seo et al., 2003, 2009) of the ensemble Kalman filter (EnKF) (Evensen, 1994, 2003). The motivation for the fixed-lag smoother is to support forecaster-supervised on-demand initialization of WRF-Hydro whether DA was previously run or not. We note here that EnKF is implemented in OpenDA (Rakovec et al., 2015; van Velzen et al., 2016) which is integrated with the NWS's Community Hydrologic Prediction System (Roe et al., 2010), the main operational river forecast system at the RFCs. As such, there already exists an operational tool for implementation of the proposed method.

The control variables, i.e., the variables to be updated or adjusted via DA, include the multiplicative adjustment factor, β_P, to precipitation, P_x, applicable uniformly to the precipitation over the entire catchment P_x, and over the entire assimilation window (see Eq. (5.13)), and the multiplicative adjustment factor, β_θ, to soil moisture, β_θ, applicable uniformly to all 4 soil moisture layers $\theta_i, i = 1,..,4$, and valid at the beginning of the assimilation window(see Eq. (5.14)):

$$Q_s = \frac{(\beta_P P_x)^2}{(P_x + I_c)^2},\ \beta_P \geq 0 \tag{5.13}$$

$$D_x = \sum_{i=1}^{4} \Delta Z_i(\theta_{sat} - \beta_\theta \theta_i),\ \beta_\theta \geq 0, i = 1,..,4 \tag{5.14}$$

The simulated streamflow observations are then augmented to the state vector to render the observation equation linear (Lorentzen and Naevdal, 2011; Rafieeinasab et al., 2014b; Lee et al., 2019). As formulated above, the DA problem amounts to solving for the two adjustment factors in each assimilation cycle such that the simulated streamflow at the catchment outlet tracks the observed. If sequential estimation is desired, the control variables may be propagated from one assimilation cycle to the next based, e.g., on the first-order autogressive-1 model (Lee et al. 2019). Different variations of the above DA approach have been used successfully with both lumped and distributed hydrologic models in both operational and research settings in the US and elsewhere (Seo et al., 2003, 2009; Lee et al., 2011, 2012, 2015, 2016; Kim et al., 2014; Lee and Seo, 2014; Rafieeinasab et al., 2014b; Riazi et al., 2016; Mazzoleni et al., 2018; Noh et al., 2018).

An important difference between the above formulation and the previous formulations of fixed lag smoothing is that the former does not include additive errors to runoff, i.e., later inflow into channels. The reason for this departure is that the addition requires modifications to the WRF-Hydro source code. Because there is no guarantee a priori that the model dynamics admit the error-added flows, the above modifications may produce numerical instabilities that are difficult to diagnose or control. The lack of additive error in the control vector means that the DA formulation is strongly-constrained rather than weakly-constrained (Lee et al., 2016), and hence more likely to render the smoother more susceptible to model structural or parametric errors. In addition to the assimilation window length and ensemble size, it is necessary to prescribe several uncertainty parameters for the smoother: the observation error variances for precipitation and streamflow, and mean and variance (or, alternatively, median and coefficient of variation) of each of β_θ and β_p. In this work, the above DA parameters were prescribed following Lee et al. (2019) using the homoscedastic model and lognormal distribution for β_θ and β_p, and were estimated based on limited sensitivity analysis (Rafieeinasab et al., 2015a; Lee et al., 2019). Due to the strongly-constrained nature of the DA formulation, however, the performance of DA is likely to benefit significantly from more rigorous estimation of the DA parameters.

5.4 Results

Our assessment consisted of the 5 experiments described below. We use peak flow and time-to-peak flow errors as the primary performance measures, by far the two most important for urban flood prediction (Liu et al., 2011; Rafieeinasab et al., 2014b).

5.4.1 Experiment 1: Event-specific vs. non-event-specific calibration

Figure 5-3 shows examples of event-specific (black) vs. non-event-specific (red) calibration results at 250 m resolution for both the LSM and routing models. Additional results are presented in Figure 10 in the context of DA. The temporal resolution of QPE is 1 min. The event-

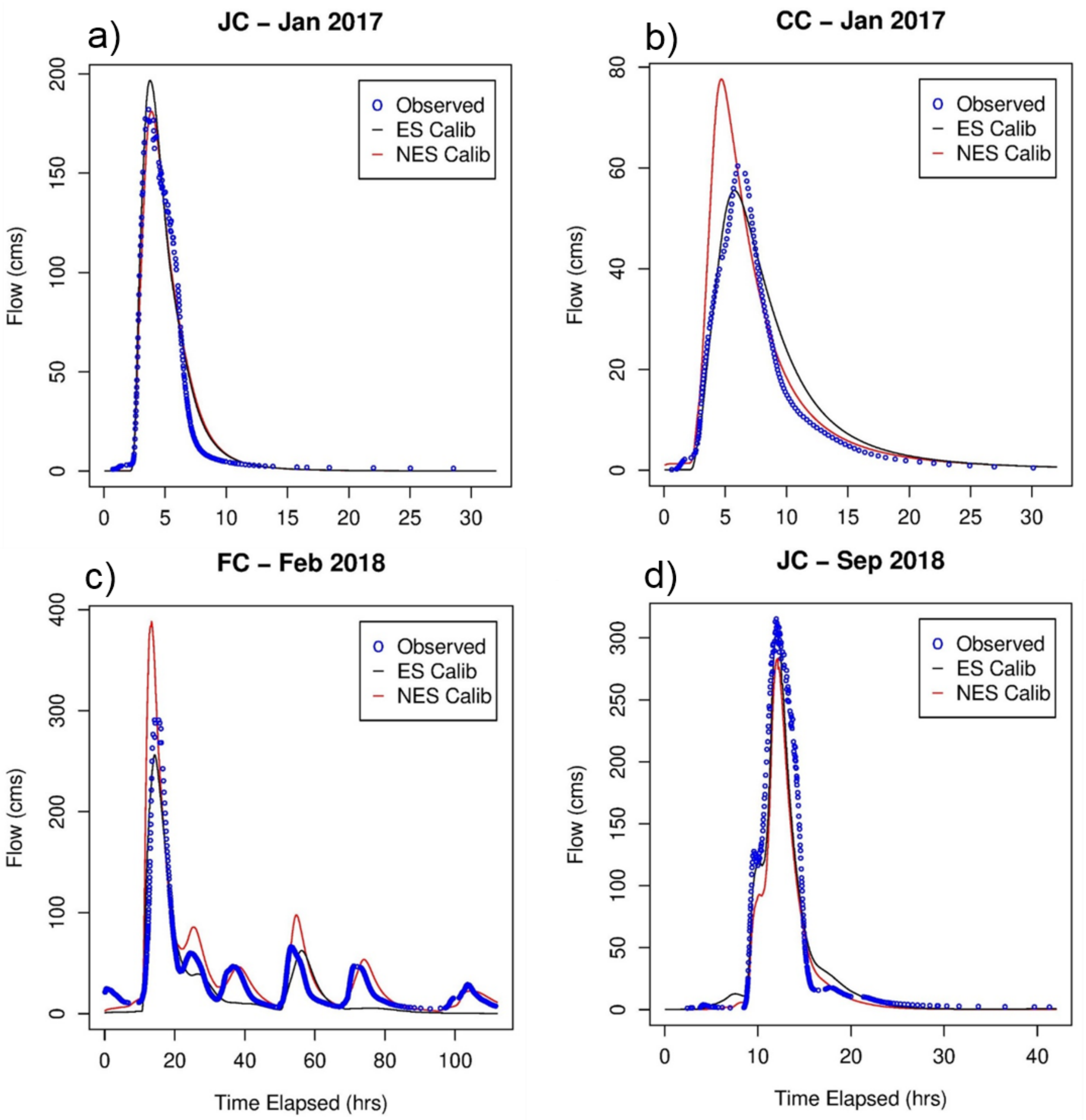


Figure 5-3 Simulation results from event-specific (black) and non-event-specific (red) calibration vs. the observed (blue empty circles) for the a) JC Jan 2017, b) CC Jan 2017, c) FC Feb 2018 and d) JC Sep 2018 cases

specific results are based on calibrating the 6 parameters specifically for each event. The non-event-specific results are based on dropping the largest and smallest values from the 5 event-specific results and averaging only the middle 3. It is important to point out that, in event-specific calibration, $rekft$ reflects the soil moisture ICs. Note in Eqs. (2) and (4) that changing $refdk$ has effects similar to changing the maximum water holding capacity of the soil, D_x, which is a function of the initial soil water content. Event-specific calibration of $refdk$ is hence subject to event-to-event variability of antecedent soil moisture conditions. The averaging of the 3 middle parameter values from the event-specific results is an attempt to dampen or average out this variability in the ICs. To illustrate, Figure 5-4a shows the event-specific result for the multiplicative factor to $rmannn$, or fac_rmannn. Significant event-to-event variations are seen particularly for less impervious CC and FC (see Figure 1). Figure 5-5b shows the non-event-specific result from averaging the middle 3 parameter values in Figure 5-5a. Note that JC, which has the largest impervious fraction (see Figure 5-1), has significantly smaller $rmannn$ than CC and FC, and that little adjustment from the WRF-Hydro default was needed for the least impervious FC.

The event-specific results indicate that the calibration strategy is mostly successful in simulating hydrographs for the most important rising limbs. For a number of cases, however, the simulated hydrographs do not recede as quickly as the observed. A likely contributing factor is that WRF-Hydro does not model storm drains. While the impact of storm drains is not very significant for large events (Rafieeinasab et al., 2015a), in lower flow conditions, the impact is likely to be larger (Habibi and Seo, 2018). Of the 15 cases (i.e., from 5 events for 3 basins), significant differences were observed for 10 cases between the event-specific and non-even-specific results. Comparison of the parameter values between the two indicates that significant differences exist most

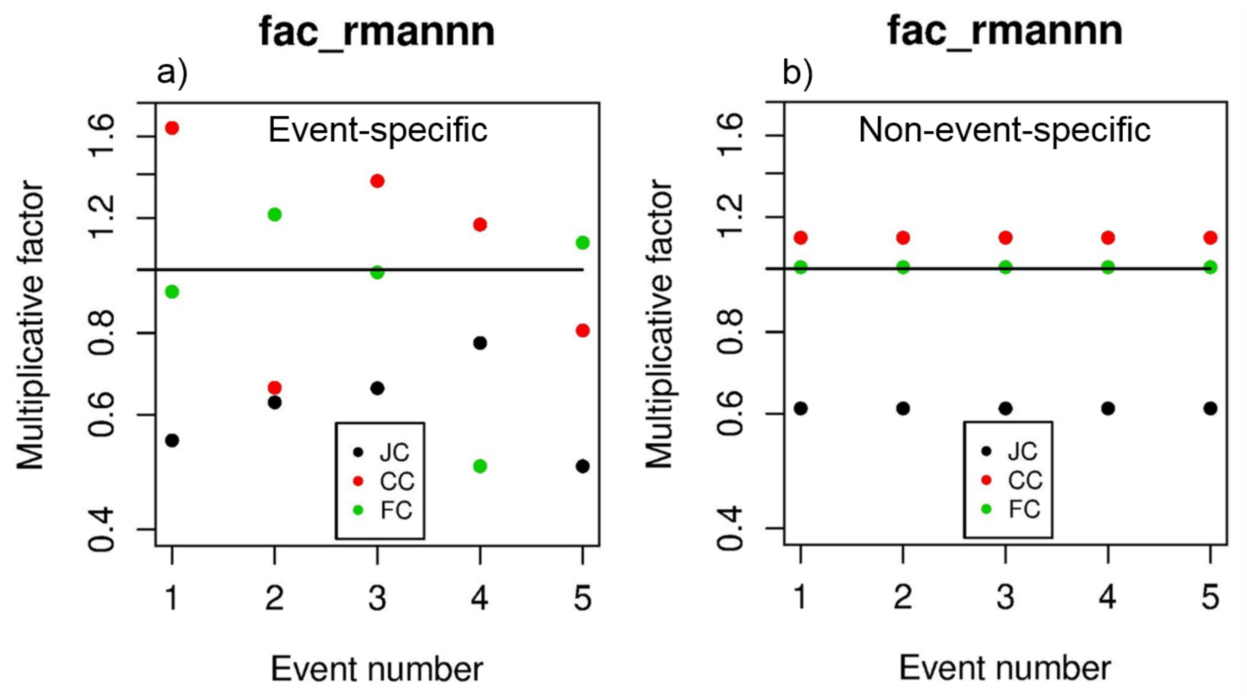


Figure 5-4 a) Multiplicative factors to Manning's n for channel routing obtained from event-specific calibration. b) Non-event-specific estimates of Manning's n for channel routing obtained from averaging for each catchment the middle 3 of the 5 values in a)

often in $refdk$ followed by $rmannn$ and sfc_rough. For bw, $hlink$ and $chsslp$, significant differences were observed only in a few cases. The large event-to-event variability of $refdk$ is not surprising in that in event-specific calibration this parameter can effectively control dynamically-varying runoff ratio as explained above. Of the 15 non-event-specific cases, 6 and 3 cases show over- and under-simulation of runoff volume resulting in over- and under-simulation of peak flows and too early and late rises to peak flows, respectively. Figure 5-5a shows the simulated peak flows from event-specific (black) and non-event-specific (red) calibration vs. the observed. Figure 5b shows the associated time-to-peak flow since the beginning of the rising limb vs. the observed. In Figure 5-5b, the absolute magnitude of the time-to-peak flow is of little importance because the beginning of the rising limb can be anywhere, and only the departure of the time-to-peak flow from the diagonal is of interest. In Figure 5-5, the JC Feb 2018 event was excluded due to lack of observed peak flow. Shown for reference in Figure 5a and Figure 5b are the lines of 10, 20 and 30 percent errors in peak flow and of 1, 2 and 3 hr errors in time-to-peak flow, respectively. Harmel et al. (2006) report streamflow measurement errors of 42%, 19%, 10%, 6% and 3% for small watersheds for the worst, typical maximum, typical average, typical minimum, and the best case scenarios, respectively. Di Baldassarre and Montanari (2009) report that the overall error affecting river discharge observations ranges from 6.2% to 42.8%, at the 95% confidence level, with an average value of 25.6%. The 10 to 30 percent error lines in Figure 5a hence provide a sense of the magnitude of the errors in simulated peak flow relative to possible observational errors. Empirical unit hydrographs for JC, CC and FC show time-to-peak values of 0.75, 3 and 2.75 hrs, respectively (Rafieeinasab et al., 2015a). An error in time-to-peak flow on the order of the time-to-peak values hence indicates poor performance. Figure 5-5 indicates that most case-specific calibration results have less than 10% error in peak flow and less than an hour of time-to-peak flow error, but that, for about 5 cases, the non-event-specific results suffer from significantly larger errors. All 5 cases of excessively large peak flow or time-to-peak flow errors are associated with significant volume errors except for the FC May 2019 case for which a less than accurate simulation of the rising limb is responsible for the large time-to-peak flow error. The above results indicate that high-quality initialization is necessary for event-based urban flood prediction using WRF-Hydro. In Experiment 5, we assess how DA may help address the situation.

5.4.2 Experiment 2: Impact of temporal resolution of precipitation

In this experiment, we assess how the temporal resolution of precipitation input may impact the quality of streamflow simulation by forcing the LSM with 1-min average of 10-min QPE vs. the native 1-min QPE. For 10 min QPE, we aggregate the 1-min CASA QPE to 10 min accumulations and run the LSM at 1 min timestep using the 1-min average over each 10 min period. For comparison, we also ran the LSM at 10 min timestep using 10-min QPE. In this experiment, we

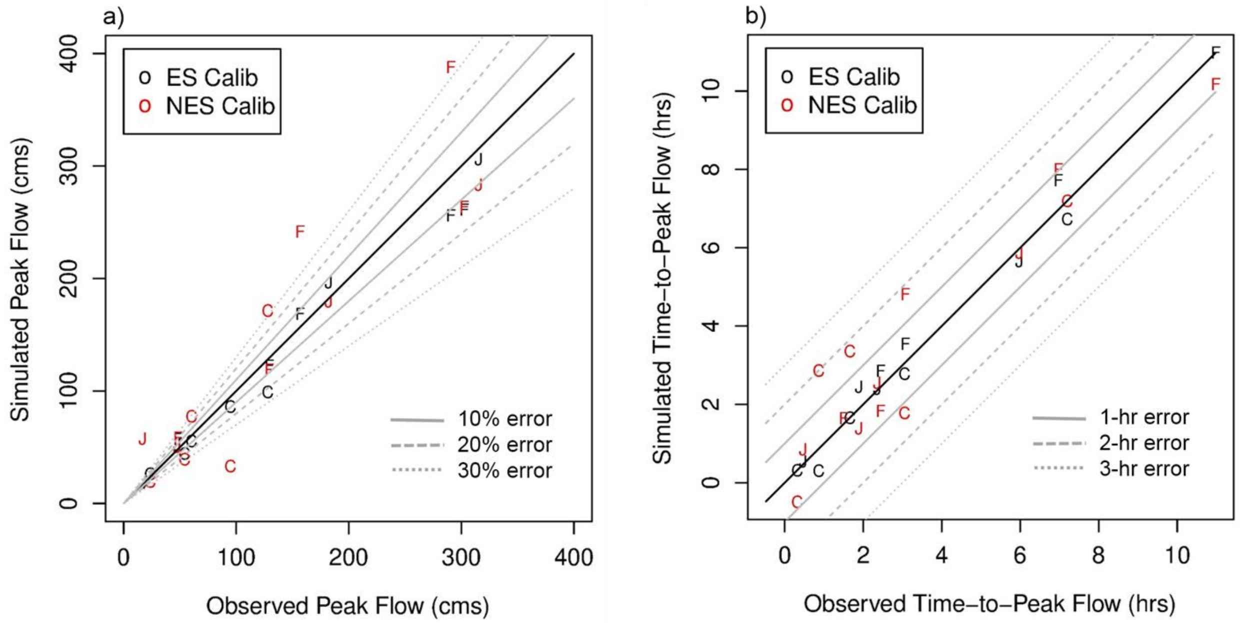


Figure 5-5 a) Comparison of simulated peak flow from event-specific (black) and non-event-specific (red) calibration vs. the observed for all 15 cases except for the JC Feb 2018 case. The symbols "J", "C" and "F" denote the JC, CC and FC results, respectively. The solid, dashed and dotted gray lines represent ±10, 20 and 30% errors. b) Same as a) but for time-to-peak flow. The solid, dashed and dotted gray lines represent ±1, 2 and 3-hr errors

use the parameter values obtained from the event-specific calibration to reduce hydrologic uncertainty. The common spatial resolution used is 250 m for both the LSM and routing models. Examination of the results for all cases indicates that the differences in simulated hydrographs due to 1 min vs. 10 min QPE are very small except for the May 2019 event which we elaborate below. Figure 5-6 shows the simulated vs. observed hydrographs at the outlet of JC for the May 2019 event. The simulation of the second rise for this double-peaked event is cut short due to missing CASA QPE. To identify possible causes for the disparate response in simulated streamflow, we examined the MAP time series for all cases. It is observed that the MAP values for the second peak of the May 2019 event are significantly smaller than those for all other events. Because runoff generation may be considered as thresholding rainfall such that little runoff occurs for rain rate below some threshold and almost all excess rainfall runs off for rain rate above the threshold (see Subsection 5.2.4, (Norouzi et al., 2019), one may look for a threshold rain rate above and below which the runoff response is very different. Examination of the MAP hyetographs and the associated hydrographs for the May 2019 event points to a threshold of about 0.5 mm. For this event, the maximum 1 min MAP associated with the second peak was well above 0.5 mm for all three basins. The maximum 1 min-average of 10 min MAP, on the other hand, was well below 0.5 mm for JC and CC, and stayed above 0.5 mm only for a single 10 min period for FC. The above findings indicate that the SWB used for rainfall-runoff modelling in WRF-Hydro is sensitive to the temporal resolution of precipitation for moderate precipitation amounts due to the increased nonlinearity in runoff generation (see Eqs. (5.1), (5.2) and (5.11)).

5.4.3 Experiment 3: Impact of spatial resolutions of rainfall-runoff modelling and routing

In this experiment, we compare the quality of the outlet simulations for peak flow and time-to-peak flow among the resolutions of 250 m, 125 m and 50 m for routing with a common LSM

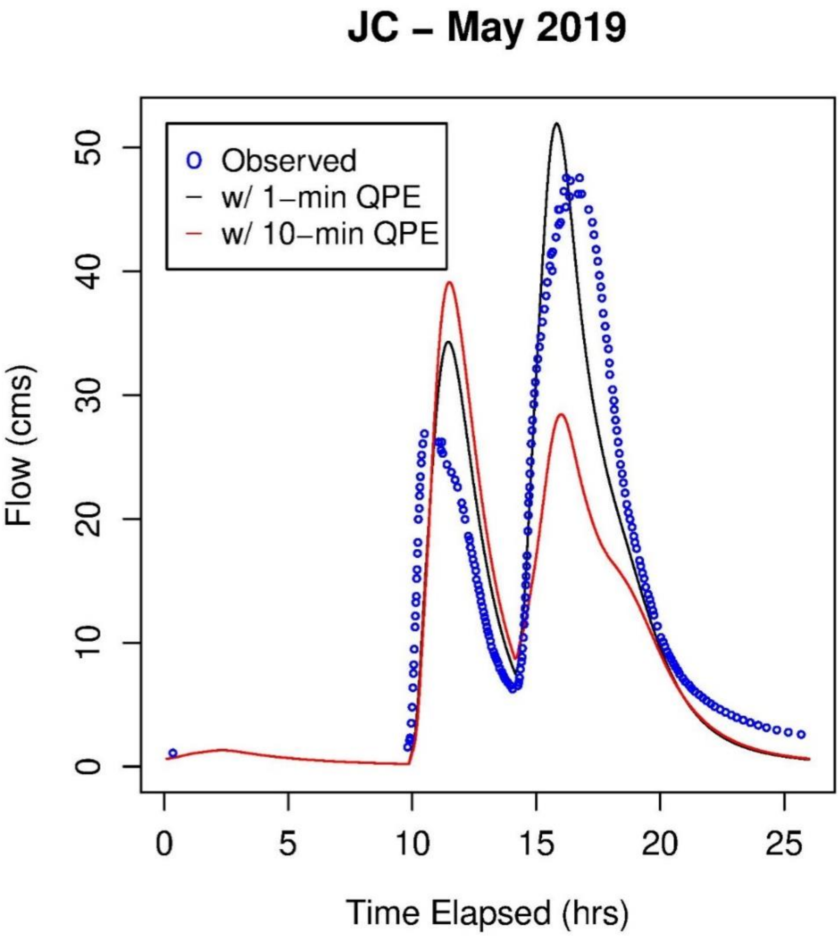


Figure 5-6 Comparison of simulated hydrographs forced by 1-min (black) and 1-min average of 10-min (red) CASA QPE vs. the observed (blue empty circles) for the JC May 2019 case

resolution of 125 m, and between the resolutions of 250 m and 125 m for LSM with a common routing model resolution of 125 m. The 250 m LSM and 250 m routing model simulations, referred to herein as the 250m-250m results, are based on event-specific calibration using SLS with LHS. One may hence consider the above calibration as based on quasi-global optimization. The 250 m LSM and 125 m routing simulations, referred to herein as the 250m-125m results, are based on event-specific calibration using only SLS in which the local search is started with the 250m-250m results. One may hence consider the above calibration as local optimization of a priori parameter values from a coarser resolution. As mentioned in Section 3, it was not possible to calibrate at the 250 m LSM and 50 m routing resolution due to excessive computational requirements (see Table 5-3). The 250 m LSM and 50 m routing simulations, referred to herein as the 250m-50m results, are based on the parameter values borrowed from the 250m-125m results. One may hence consider the above results as based solely on a priori parameter values transferred from a coarser resolution. Because the level of calibration is different from one resolution to another, it is not very meaningful to compare the non-event-specific results. For this reason, we focus below on the event-specific results only.

Figure 5-7a and 5-7b show the simulated peak flow vs. the observed, and the simulated time-to-peak flow vs. the observed, respectively. As in Figure 5-5, we overlay the 10, 20 and 30 percent error lines in Figure 5-7a and of 1, 2 and 3 hours of timing error lines in Figure 5-7b to help assess the magnitude of the errors. Figure 5-7 indicates that the 250m-250m and 250m-125m results, both of which are calibrated scale-specifically, are very similar, and that for a number of events the 250m-50m results are not as good as the above two. The above observations are perhaps not very surprising in that one may expect scale-specific calibration to perform better than using parameter values borrowed from a lower resolution. The magnitude of the errors in the 250m-50m results, however, is surprisingly large for a number of events. To trace the potential sources of the error, we examined the spatially-distributed channel routing parameters, including the channel grid, flow accumulation, flow direction and stream order at all resolutions. It is seen that, whereas the differences between 250 and 125 m are relatively small, there are large differences between 50 m and the coarser resolutions. To illustrate, Figures 5-8a and 5-8b show the histograms of the stream order in the model domain at resolutions of 125 m and 50 m, respectively. The histogram at 250 m is similar to that at 125 m. In the figure, the frequency for the stream order of zero represents the number of grid boxes that do not contain any channel segments. As one may expect, at 50 m resolution, the channel network is much denser and has more higher-order streams. WRF-Hydro prescribes the channel routing parameters according to the stream order. As such, changes in the channel density or stream order are very likely to change the conveyance characteristics of the

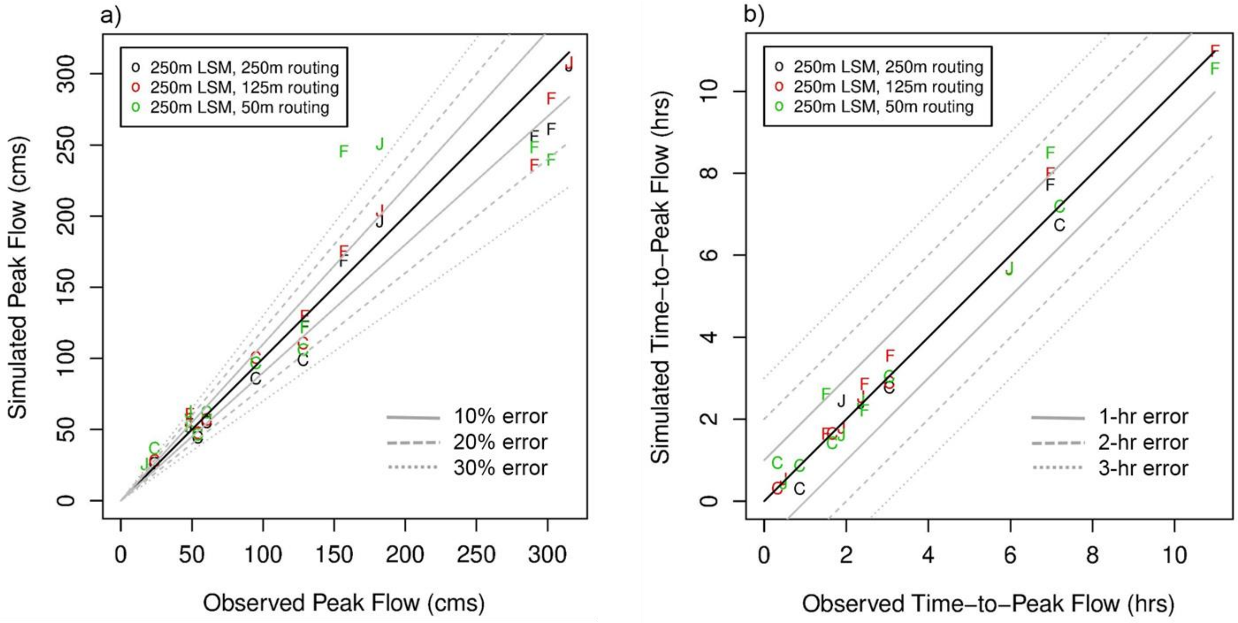


Figure 5-7 Same as Figure 5 but the comparison is among the 250 m LSM and 250 m routing (black), 250 m LSM and 125 m routing (red) and 250 m LSM and 50 m routing (green) results

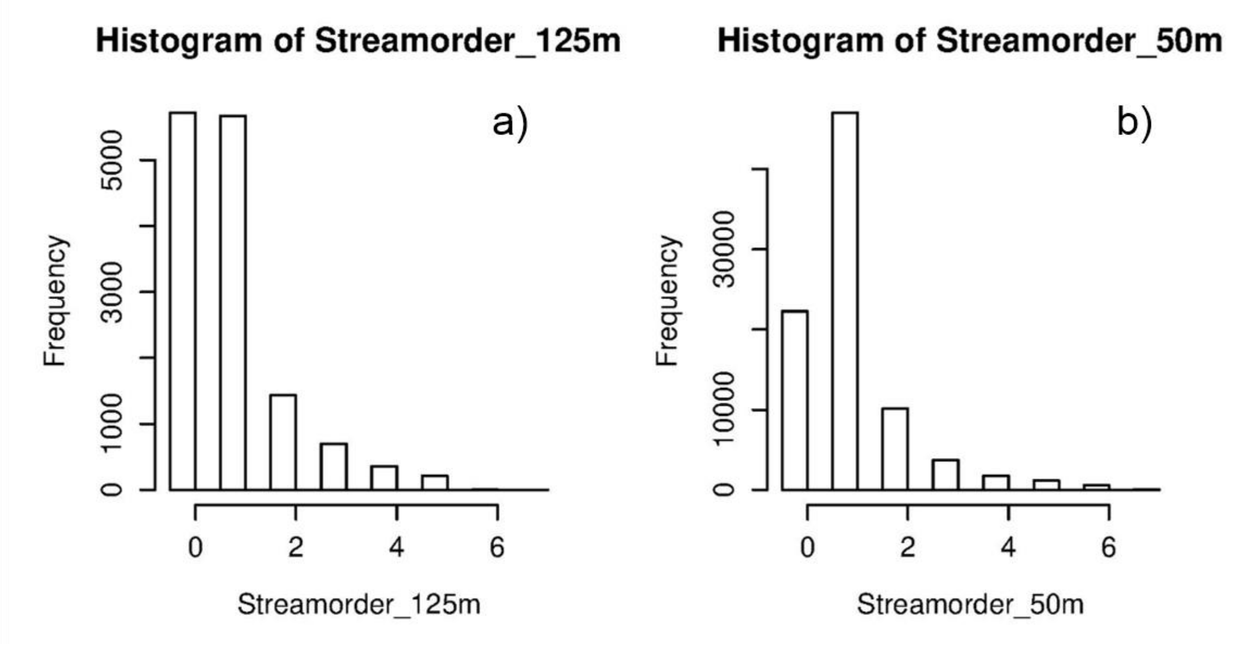


Figure 5-8 Histograms of stream order as modeled at resolutions of a) 125 m and b) 50 m

channel network. The above findings suggest that a combination of resolution-specific prescription of the channel routing parameters and their calibration is likely to be necessary to benefit from very high-resolution modeling using WRF-Hydro. We also compared the 250m-125m results with the 125m-125m to assess the impact of increasing the LSM resolution. As with the 250m-125m results, the 125m-125m results are based on scale-specific local optimization using SLS in which the parameter values from the 250m-125m results are used as the starting point. The comparison indicates that the 125m-125m results improve the peak flow prediction over the 250m-125m for the study basins but only marginally.

5.4.4 Experiment 4: Impact of quality of ICs

In this experiment, we assess how the quality of the ICs of the rainfall-runoff model may impact the accuracy of streamflow prediction. A potential source of the ICs in real-time event-based operation of WRF-Hydro is the warm states of the NWM. A direct use in this experiment of the NWM warm states, however, is not likely to allow clear attribution at least for two reasons.

The first is that the USGS 24-category land cover (see Figure 5-1c) and the MRMS QPE (Zhang et al., 2011, 2016) used in NWM are of coarser resolution than those used in this work. The second is that the model parameter values used in the NWM (Gochis et al., 2019) are not the same as those used in WRF-Hydro in this work. As such, the ICs from the NWM analysis are not likely to transfer cleanly to WRF-Hydro as implemented in this work as evidenced in Experiments 1 through 3 above. As a compromise, we emulate the NWM analysis by running WRF-Hydro using the USGS 24-category land cover and NLDAS precipitation (Cosgrove et al., 2003) in place of the NLCD land cover and CASA QPE, respectively. The NLDAS precipitation has a much lower resolution than the 1 km 1 hr MRMS QPE used by the NWM. It is hence possible that the results from this experiment may somewhat inflate the positive impact of higher resolution precipitation. The above experiment design nonetheless completely removes all model-parametric uncertainties and hence makes possible unambiguous attribution.

In this experiment, we start running WRF-Hydro at least several hours before the prediction time using the NLDAS precipitation and USGS 24-category land cover where the lower resolution NLDAS precipitation is disaggregated uniformly in space and time to a resolution of 250 m and 1 min. The prediction time is chosen where the observed hydrograph begins to rise. This is also when streamflow response is most sensitive to the ICs. At the prediction time, we switch to the CASA QPE and NLCD land cover for simulation over the forecast horizon. For the above comparison run, we assume average soil moisture conditions for the LSM and pre-storm conditions for the hillslopes and channel routing models as obtained from event-specific calibration (see Subsection 4.1). In the baseline run, we run the model at 250 m 1 min resolution using the CASA QPE and NLCD land cover for the entire simulation period. Any differences in the two simulated hydrographs over the forecast horizon are hence due solely to the ICs valid at the prediction time. Figure

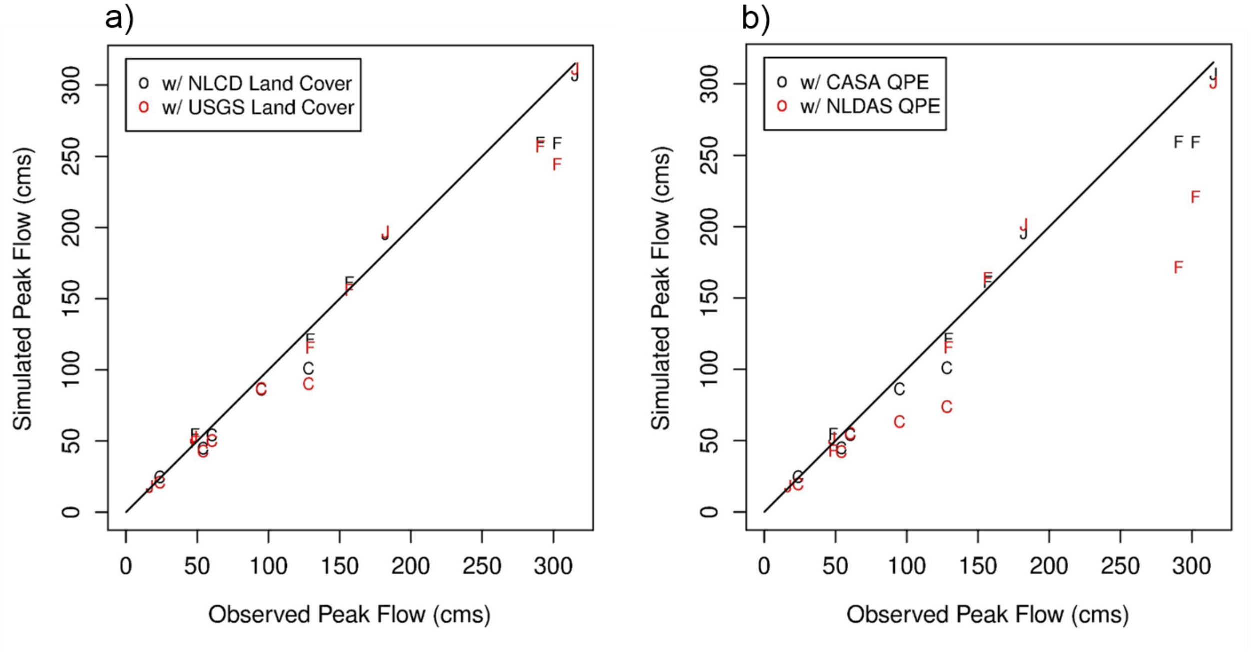


Figure 5-9 Same as Figure 5-5a but the comparison is for a) the NLCD (black) vs. the USGS 24-category (red) land cover results, and b) the CASA QPE (black) vs. the NLDAS QPE (red) results

5-9a shows the simulated vs. observed peak flow for the NLCD (black) and USGS 24-category (red) land cover. All other conditions are the same as in the baseline 250m-250m simulation. The positive impact of higher-resolution land cover is readily seen. Note that the differences are the smallest for JC which is identified mostly as urban by the USGS 24-category land cover in agreement with the NLCD (see Figure 5-1c,d). Figure 5-9b shows the simulated vs. observed peak flow for the CASA (black) and NLDAS (red) QPE-forced ICs. All other conditions are the same as in the baseline 250m-250m simulation. Note the very significant positive impact of higher-resolution QPE, particularly for CC and FC for Feb 2018 and Sep 2018, the two largest events among the five (see Figure 5-2 and Table 5-1). Examination of timing errors associated with Figs 5-9a and 5-9b shows similarly positive impact of higher-resolution QPE and, to a lesser extent, land cover.

5.4.5 Experiment 5: Impact of updating ICs via DA

In this Experiment, we assess how DA may potentially be used to initialize WRF-Hydro for event-based prediction. In the real world, it is generally not possible to schedule pre-storm warmup runs as described in the 4th Experiment. Instead, it is necessary to be able to initialize the model on demand often without the aid of any a priori information. The fixed-lag smoother, solved using EnKF in this work, is aimed at supporting such an operation. For high-resolution runs, EnKF is computationally expensive. In this work, all ensemble runs were made at the coarsest spatial resolution of 250 m for both the LSM and routing models. Limited sensitivity analysis suggests that a small ensemble size of 12 is generally acceptable for ensemble mean prediction owing to the very low dimensionality of the DA formulation. We then use the non-event-specific calibration results to emulate realistic model-parametric uncertainty and predict streamflow with and without DA. Due to the small sample size, quantitative verification was not possible. Instead, we critically examine the DA-aided predictions for those 5 cases for which the non-event-specific calibration results compare least favorably with the event-specific in Experiment 1 (see Figure 5-3). By far the largest potential value of DA in urban flood prediction is improving peak flow and time-to-peak flow predictions when the streams first respond to rainfall. Accordingly, we focus specifically on DA-aided predictions when the hydrograph begins to rise. This is also the time when the degrees of freedom for signal for DA (Rodgers, 2000) is greatly reduced due to the generally reduced predictive skill of rainfall-runoff and routing models, and hence streamflow observations carry larger information content relative to the model prediction (Zupanski et al., 2007; Zupanski, 2009).

Figure 5-10 shows the streamflow predictions without DA (red), DA-aided ensemble predictions (cyan), the associated ensemble mean predictions (blue), ensemble streamflow analysis from DA (green) and the verifying observed hydrographs (empty blue circles) for 4 of the 5 cases for which non-event-specific calibration produced very poor simulations in Experiment 1. The case not shown in Figure 5-10 due to space limitations is JC Apr 2019 which is by far the smallest

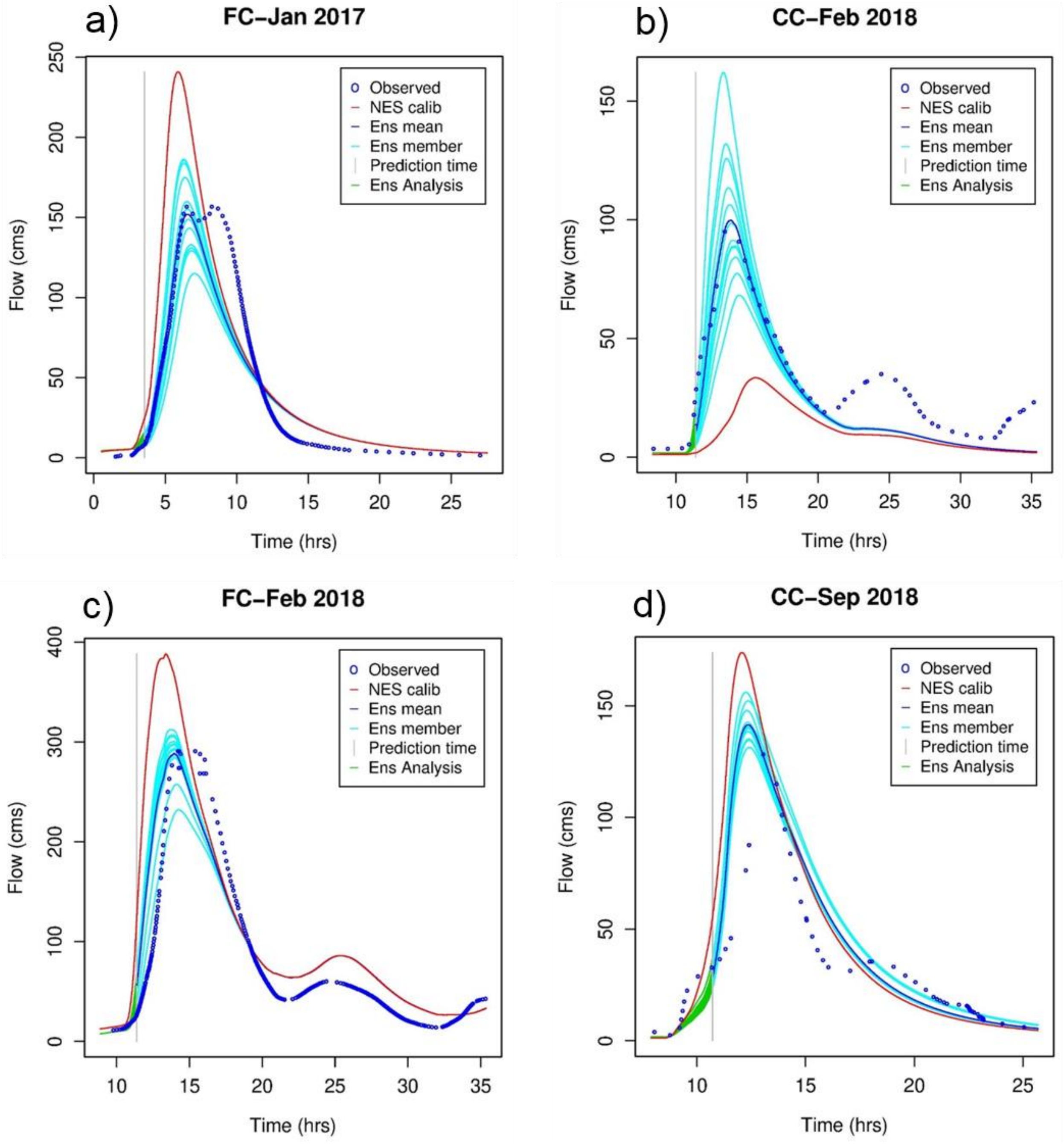


Figure 5-10 DA-aided ensemble predictions (cyan), ensemble mean prediction (blue) and DA-unaided base predictions based on non-event-specific calibration (red) vs. the observed (blue empty circles) for the a) FC Jan 2017, b) CC Feb 2018, c) FC Feb 2018 and d) CC Sep 2018 cases. The green and black lines show the ensemble DA analysis within the assimilation window and the prediction time, respectively

event of the 5 and is hence of lesser interest. In the figure, the vertical gray line indicates the prediction time which also marks the end of the assimilation window. The horizontal extent of the ensemble analysis (green) shows the size of the assimilation window. All streamflow and precipitation observations valid within the assimilation window are assimilated in these runs to update the soil moisture states valid at the prediction time. All DA results are based on single assimilation cycles to emulate on-demand operation without the potential benefit of any previous DA cycles. The results indicate that DA improves prediction for all 5 cases over the DA-unaided base predictions. For the FC Jan 2017 and CC Feb 2018 events, for which non-event-specific calibration very significantly over- and under-predict, respectively, DA greatly improves prediction. As noted in Section 3, the primary source of error in peak flow or time-to-peak flow is the error in runoff volume. The results indicate that DA is largely able to reduce runoff volume errors by providing WRF-Hydro with high quality ICs. Figure 10 shows, however, that the ensembles are significantly underspread in the recession limb due to lack of accounting of structural and parametric uncertainties, and that WRF-Hydro is not able to reproduce the bimodal or attenuated peaks, or the fast-receding falling limbs in FC Jan 2017 (Figure 5-10a) and FC Feb 2018 (Figure 5-10c). The above results indicate that, overall, the fixed-lag smoother is very effective in reducing runoff volume errors and hence errors in peak flow and time-to-peak flow.

5.5 Conclusions and future research recommendations

We assess the impact of increasing the resolution of hydrologic modeling, calibration of selected model parameters and assimilation of streamflow observations toward event-based high-resolution urban flood modeling and prediction using WRF-Hydro in the Dallas-Fort Worth area (DFW). We use quantitative precipitation estimates (QPE) at 500-m 1-min resolution from the Collaborative Adaptive Sensing of the Atmosphere (CASA) operation for observed rainfall, the

Stepwise Line Search for calibration, and ensemble Kalman filter (EnKF) implementation of fixed-lag smoothing for data assimilation (DA). The model domain is a 144.6 km^2 area comprising 3 urban catchments in the Cities of Arlington and Grand Prairie in the middle of DFW. The main findings, conclusions and recommendations follow below.

Event-specific calibration of the 6 WRF-Hydro parameters identified in this work is largely successful in simulating hydrographs in the study area, in particular, the most important rising limbs. It is less successful, however, for attenuated peaks or fast-receding falling limbs. A novel element in the above calibration is the inclusion of a conditional bias penalty in the objective function to improve simulation specifically of highly peaked hydrograph. A spatial resolution of at least 250 m is necessary for the land surface model (LSM) to delineate small catchments and hence to capture catchment-wide rainfall with acceptable accuracy. Increasing the resolution of the LSM from 250 m to 125 m showed marginal improvement. The same resolution increase for the routing models showed little improvement. Increasing the routing resolution further to 50 m using parameter values borrowed from 125 m, on the other hand, increased errors for a number of cases due to large changes in channel grid and stream order. The above findings suggest that, to benefit from very high-resolution modeling using WRF-Hydro, a combination of resolution-specific prescription and calibration of the channel routing parameters is likely to be necessary. The high-resolution CASA QPE and the National Land Cover Database (NLCD) land cover showed very significant and significant positive impact on streamflow simulation compared to the lower-resolution North American Land Data Assimilation System (NLDAS) QPE and USGS 24-category land cover, respectively. The above points out the importance of resolution-consistent high-quality initialization of WRF-Hydro for event-based operation. The EnKF implementation of fixed-lag smoother significantly reduced peak flow errors under realistic parametric uncertainty for predictions made when streams first respond to rainfall. The DA-aided ensemble predictions are, however, significantly underspread in the recession limb due to lack of accounting of structural and parametric uncertainties. The overall results suggest that, in the absence of resolution-specific prescription and calibration of channel routing parameters, a resolution of 250 m for both the LSM and routing models is a good choice in terms of performance and computational requirements. Recall that the National Water Model currently runs routing at 250 m over the continental US. The results also suggest that, in the absence of high-quality calibration and continuous simulation of streamflow, DA is necessary to initialize WRF-Hydro for event-based operation for high-resolution urban flood prediction.

Chapter 6. Conclusions and Future Research Recommendations

Floods are associated with extreme states of the hydrologic system which can only be observed with significant uncertainty. As such, data assimilation (DA) for flood forecasting is subject to conditional bias (CB) which impacts optimal information fusion. While existing solutions such as CBPKF and CBEnKF can drastically improve the DA performance over the tails, they do not improve unconditional performance. In this work, we proposed CB-aware DA techniques, which adaptively prescribe the weight for CB penalty, thereby optimally combines the information content in observation and model prediction.

Chapter 2 presents adaptive CBPKF and demonstrate its capability of improving improve unconditional performance by adaptively prescribing the weight for the CB penalty. For evaluation, we carried out synthetic experiments using linear systems with varying degrees of dynamical model uncertainty, observational uncertainty, and predictability. We also introduced VIKF, which is an approximation of CBPKF that is much less computationally intensive.

Chapter 3 describes conditional bias-penalized ensemble Kalman filter (CBEnKF) which is then used as the foundation for adaptive CBEnKF presented in Chapter 4. We evaluate adaptive CBEnKF against widely applied ensemble Kalman filter (EnKF) for 6 headwater basins in Texas using the NWS operational lumped hydrologic models. CB-aware DA and the degrees of freedom for signal are then used to quantify the marginal information content in the observations. The main findings, conclusions and future research recommendations follow below.

CB arises very frequently in varying magnitude when assimilating streamflow observations during the catchment’s response to precipitation and drainage. In general, the larger the discharge is, the larger the CB is. Adaptive CEnKF improves over EnKF consistently and significantly regardless of the level of hydrologic uncertainty modeling or the severity of timing errors. The improvement is particularly large during sharp rises of the outlet hydrograph with large peak flow. The flow-dependent marginal information content in the observations varies with the streamflow response of the catchment and the magnitude of CB, and tends to decrease and increase in the rising and falling phases of the hydrograph, respectively. The work summarized herein indicates that CB-aware DA with information content analysis offers an objective approach for dynamically balancing the flow-dependent predictive skill of hydrologic models, quality and frequency of hydrologic observations, and scheduling of the DA cycle toward cost-effective improvement of operational flood forecasting. Addressing timing errors effectively remain a very significant challenge in streamflow DA. Multi-scale bias correction approaches (Noh et al., 2019), which are largely phase error-agnostic, warrant further attention. Additional research is also needed toward computationally less expensive and algorithmically simpler approximations to CBPKF/CBEnKF such as variance-inflated KF (VIKF)/EnKF (Shen et al., 2019). In this work, α was optimized in a brute-force manner. Additional research is necessary to utilize the closed-form expression for the derivative of the quadratic cost function with respect to α available in VIKF for gradient-based minimization.

BIBLIOGRAPHY

Afshar, M.H., Yilmaz, M.T., Crow, W.T., 2019. Impact of Rescaling Approaches in Simple Fusion of Soil Moisture Products. Water Resour. Res. 55, 7804–7825. https://doi.org/10.1029/2019WR025111

Akan, A.O., Houghtalen, R.J., 2013. Urban Hydrology, Hydraulics, and Stormwater Quality: Engineering Applications and Computer Modeling. Wiley, Hoboken.

Alizadeh, B., Limon, R.A., Seo, D.-J., Lee, H., Brown, J., 2020. Multiscale Postprocessor for Ensemble Streamflow Prediction for Short to Long Ranges. Journal of Hydrometeorology 21, 265–285. https://doi.org/10.1175/JHM-D-19-0164.1

Anagnostou, E.N., Krajewski, W.F., 1999. Real-Time Radar Rainfall Estimation. Part I: Algorithm Formulation. J. Atmos. Oceanic Technol. 16, 189–197. https://doi.org/10.1175/1520-0426(1999)016<0189:RTRREP>2.0.CO;2

Antoniou, C., Ben-Akiva, M., Koutsopoulos, H.N., 2007. Nonlinear Kalman Filtering Algorithms for On-Line Calibration of Dynamic Traffic Assignment Models. IEEE Transactions on Intelligent Transportation Systems 8, 661–670. https://doi.org/10.1109/TITS.2007.908569

Bhotto, M.Z.A., Bajić, I.V., 2015. Constant Modulus Blind Adaptive Beamforming Based on Unscented Kalman Filtering. IEEE Signal Processing Letters 22, 474–478. https://doi.org/10.1109/LSP.2014.2362932

Bocher, M., Fournier, A., Coltice, N., 2018. Ensemble Kalman filter for the reconstruction of the Earth’s mantle circulation. Nonlinear Processes in Geophysics 25, 99–123. https://doi.org/10.5194/npg-25-99-2018

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

Brock, F.V., Crawford, K.C., Elliott, R.L., Cuperus, G.W., Stadler, S.J., Johnson, H.L., Eilts, M.D., 1995. The Oklahoma Mesonet: A Technical Overview. Journal of Atmospheric and Oceanic Technology 12, 5–19. https://doi.org/10.1175/1520-0426(1995)012<0005:TOMATO>2.0.CO;2

Brown, J.D., He, M., Regonda, S., Wu, L., Lee, H., Seo, D.-J., 2014. Verification of temperature, precipitation, and streamflow forecasts from the NOAA/NWS Hydrologic Ensemble Forecast Service (HEFS): 2. Streamflow verification. Journal of Hydrology 519, 2847–2868. https://doi.org/10.1016/j.jhydrol.2014.05.030

Brown, J.D., Seo, D.-J., 2013. Evaluation of a nonparametric post-processor for bias correction and uncertainty estimation of hydrologic predictions. Hydrological Processes 27, 83–105. https://doi.org/10.1002/hyp.9263

Brown, J.D., Seo, D.-J., Brown, J.D., Seo, D.-J., 2010. A Nonparametric Postprocessor for Bias Correction of Hydrometeorological and Hydrologic Ensemble Forecasts. Journal of Hydrometeorology 11, 642–665. https://doi.org/10.1175/2009JHM1188.1

Budyko, M.I., 1974. Climate and Life. Elsevier Science, Oxford.

Burnash, R.J.C., Ferral, R.L., McGuire, R.A., 1973. A generalized streamflow simulation system: conceptual modeling for digital computers. U. S. Dept. of Commerce, National Weather Service, Sacramento, California.

Carpenter, T.M., Georgakakos, K.P., 2004. Continuous streamflow simulation with the HRCDHM distributed hydrologic model. Journal of Hydrology, The Distributed Model Intercomparison Project (DMIP) 298, 61–79. https://doi.org/10.1016/j.jhydrol.2004.03.032

Carpenter, T.M., Georgakakos, K.P., Sperfslagea, J.A., 2001. On the parametric and NEXRAD-radar sensitivities of a distributed hydrologic model suitable for operational use. Journal of Hydrology 253, 169–193. https://doi.org/10.1016/S0022-1694(01)00476-0

Catherine Loader, 2013. locfit: Local Regression, Likelihood and Density Estimation.

Chandrasekar, V., 2017. The CASA Dallas-Fort Worth urban radar network for flood monitoring: accomplishments and lessons learned after 4 years of operation. Presented at the Weather Radar and Hydrology, Seoul, Korea.

Chandrasekar, V., Chen, H., Philips, B., Seo, D., Junyent, F., Bajaj, A., Zink, M., Mcenery, J., Sukheswalla, Z., Cannon, A., Lyons, E., Westbrook, D., 2013. The CASA Dallas-Fort Worth remote sensing network ICT for urban disaster mitigation 15, EGU2013-6351.

Chandrasekar, V., Wang, Y., Chen, H., 2012. The CASA quantitative precipitation estimation system: a five year validation study. Natural Hazards and Earth System Sciences 12, 2811–2820. https://doi.org/10.5194/nhess-12-2811-2012

Chen, F., Crow, W.T., Ryu, D., 2014. Dual Forcing and State Correction via Soil Moisture Assimilation for Improved Rainfall–Runoff Modeling. J. Hydrometeor. 15, 1832–1848. https://doi.org/10.1175/JHM-D-14-0002.1

Chen, H., Chandrasekar, V., 2015. The quantitative precipitation estimation system for Dallas–Fort Worth (DFW) urban remote sensing network. Journal of Hydrology, Hydrologic Applications of Weather Radar 531, 259–271. https://doi.org/10.1016/j.jhydrol.2015.05.040

Chen, W., Shen, H., Huang, C., Li, X., 2017. Improving Soil Moisture Estimation with a Dual Ensemble Kalman Smoother by Jointly Assimilating AMSR-E Brightness Temperature and MODIS LST. Remote Sensing 9, 273. https://doi.org/10.3390/rs9030273

Chow, V.T., Maidment, D.R., Mays, L.W., 1988. Applied hydrology, McGraw-Hill series in water resources and environmental engineering. McGraw-Hill, New York.

Ciach, G.J., Morrissey, M.L., Krajewski, W.F., Ciach, G.J., Morrissey, M.L., Krajewski, W.F., 2000. Conditional Bias in Radar Rainfall Estimation. Journal of Applied Meteorology 39, 1941–1946. https://doi.org/10.1175/1520-0450(2000)039<1941:CBIRRE>2.0.CO;2

Cifelli, R., Chandrasekar, V., Chen, H., Johnson, L.E., 2018. High resolution radar quantitative precipitation estimation in the San Francisco Bay area: rainfall monitoring for the urban environment. Journal of the Meteorological Society of Japan 96A, 141–155. https://doi.org/10.2151/jmsj.2018-016

Clark, M.P., Rupp, D.E., Woods, R.A., Zheng, X., Ibbitt, R.P., Slater, A.G., Schmidt, J., Uddstrom, M.J., 2008. Hydrological data assimilation with the ensemble Kalman filter: Use of streamflow observations to update states in a distributed hydrological model. Advances in Water Resources 31, 1309–1324. https://doi.org/10.1016/J.ADVWATRES.2008.06.005

Clark, M.P., Schaefli, B., Schymanski, S.J., Samaniego, L., Luce, C.H., Jackson, B.M., Freer, J.E., Arnold, J.R., Moore, R.D., Istanbulluoglu, E., Ceola, S., 2016. Improving the theoretical underpinnings of process-based hydrologic models. Water Resour. Res. 52, 2350–2365. https://doi.org/10.1002/2015WR017910

Cosgrove, B.A., Lohmann, D., Mitchell, K.E., Houser, P.R., Wood, E.F., Schaake, J.C., Robock, A., Marshall, C., Sheffield, J., Duan, Q., Luo, L., Higgins, R.W., Pinker, R.T., Tarpley, J.D., Meng, J., 2003. Real-time and retrospective forcing in the North American Land Data Assimilation System (NLDAS) project. Journal of Geophysical Research: Atmospheres 108, 2002JD003118. https://doi.org/10.1029/2002JD003118

Cotter, J., 2015. USACE Flood Risk and Water Supply Activities in the Dallas Fort Worth Area. Sustainable Urban Water Workshop, The University of Texas at Arlington, Arlington, TX.

Crow, W.T., Ryu, D., 2009. A new data assimilation approach for improving runoff prediction using remotely-sensed soil moisture retrievals. Hydrology and Earth System Sciences 13, 1–16. https://doi.org/10.5194/hess-13-1-2009

Crow, W.T., van den Berg, M.J., Huffman, G.J., Pellarin, T., 2011. Correcting rainfall using satellite-based surface soil moisture retrievals: The Soil Moisture Analysis Rainfall Tool (SMART). Water Resour. Res. 47. https://doi.org/10.1029/2011WR010576

Curran, J., 2007. Modeling future flows in the Blanco river watershed under various development and rainfall scenarios. Texas Journal of Science 59, 209–232.

Das, N.N., Entekhabi, D., Dunbar, R.S., Chaubell, M.J., Colliander, A., Yueh, S., Jagdhuber, T., Chen, F., Crow, W., O'Neill, P.E., Walker, J.P., Berg, A., Bosch, D.D., Caldwell, T., Cosh, M.H., Collins, C.H., Lopez-Baeza, E., Thibeault, M., 2019. The SMAP and Copernicus Sentinel 1A/B microwave active-passive high resolution surface soil moisture product. Remote Sensing of Environment 233, 111380. https://doi.org/10.1016/j.rse.2019.111380

de Wit, A.J.W., van Diepen, C.A., 2007. Crop model data assimilation with the Ensemble Kalman filter for improving regional crop yield forecasts. Agricultural and Forest Meteorology 146, 38–56. https://doi.org/10.1016/j.agrformet.2007.05.004

Di Baldassarre, G., Montanari, A., 2009. Uncertainty in river discharge observations: a quantitative analysis. Hydrology and Earth System Sciences 6. https://doi.org/10.5194/hessd-6-39-2009

Dong, Z., You, Z., 2006. Finite-horizon robust Kalman filtering for uncertain discrete time-varying systems with uncertain-covariance white noises. IEEE Signal Processing Letters 13, 493–496. https://doi.org/10.1109/LSP.2006.873148

Dooge, J.C.I., 1959. A general theory of the unit hydrograph. J. Geophys. Res. 64, 241–256. https://doi.org/10.1029/JZ064i002p00241

Evensen, G., 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. Journal of Geophysical Research: Oceans 99, 10143–10162. https://doi.org/10.1029/94JC00572

Evensen, G., 2003. The ensemble Kalman filter: theoretical formulation and practical implementation. Ocean Dynamics 53, 343–367. https://doi.org/10.1007/s10236-003-0036-9

Evensen, G., van Leeuwen, P.J., 2000. An Ensemble Kalman Smoother for Nonlinear Dynamics. Mon. Wea. Rev. 128, 1852–1867. https://doi.org/10.1175/1520-0493(2000)128<1852:AEKSFN>2.0.CO;2

Fiebrich, C.A., Grimsley, D.L., McPherson, R.A., Kesler, K.A., Essenberg, G.R., 2006. The Value of Routine Site Visits in Managing and Maintaining Quality Data from the Oklahoma Mesonet. Journal of Atmospheric and Oceanic Technology 23, 406–416. https://doi.org/10.1175/JTECH1852.1

Fisher, M., 2003. Estimation of entropy reduction and degrees of freedom for signal for large variational analysis systems. https://doi.org/10.21957/2BEC9M38O

Frost, C., Thompson, S.G., 2000. Correcting for regression dilution bias: comparison of methods for a single predictor variable. J Royal Statistical Soc A 163, 173–189. https://doi.org/10.1111/1467-985X.00164

Fuller, W.A., 1987. Measurement error models. Wiley.

Fulton, R.A., Breidenbach, J.P., Seo, D.-J., Miller, D.A., O'Bannon, T., 1998. The WSR-88D Rainfall Algorithm. Wea. Forecasting 13, 377–395. https://doi.org/10.1175/1520-0434(1998)013<0377:TWRA>2.0.CO;2

Furl, C., Sharif, H., Zeitler, J.W., Hassan, A.E., Joseph, J., 2018. Hydrometeorology of the catastrophic Blanco river flood in South Texas, May 2015. Journal of Hydrology: Regional Studies 15, 90–104. https://doi.org/10.1016/j.ejrh.2017.12.001

Gao, Z., Shen, W., Zhang, H., Ge, M., Niu, X., 2016. Application of Helmert Variance Component Based Adaptive Kalman Filter in Multi-GNSS PPP/INS Tightly Coupled Integration. Remote Sensing 8, 553. https://doi.org/10.3390/rs8070553

Gochis, D., Barlage, M., Dugger, A., FitzGerald, K., Karsten, L., McAllister, M., McCreight, J., Mills, J., Rafieei Nasab, A., Read, L., Sampson, K., Yates, D., Yu, W., 2018. The WRF-Hydro modeling system technical description, (Version 5.0). Available online at: https://ral.ucar.edu/sites/default/files/public/WRF-HydroV5TechnicalDescription.pdf (last accessed: Jul 07, 2020).

Gochis, D., Yates, D., Sampson, K., Dugger, A., McCreight, J., Barlage, M., Rafieeinasab, A., Karsten, L., Read, L., Zhang, Y., McAllister, M., Cabell, R., FitzGerald, K., 2019. Overview of national water model calibration general strategy & optimization.

Gourley, J.J., Vieux, B.E., 2005. A Method for Evaluating the Accuracy of Quantitative Precipitation Estimates from a Hydrologic Modeling Perspective. Journal of Hydrometeorology 6, 115–133. https://doi.org/10.1175/JHM408.1

Graziano, T., Clark, E., Cosgrove, B., Gochis, D., 2017. Transforming National Oceanic and Atmospheric Administration (NOAA) water resources prediction. Presented at the 97th American Meteorological Society Annual Meeting, AMS, Seattle, WA.

Green, W.H., Ampt, G.A., 1911. Studies on Soil Phyics. The Journal of Agricultural Science 4, 1–24. https://doi.org/10.1017/S0021859600001441

Greene, K., 2013. Texas Statewide Mesonet Network (TexMesonet).

Gupta, H.V., Sorooshian, S., Hogue, T.S., Boyle, D.P., 2003. Advances in automatic calibration of watershed models, in: Calibration of Watershed Models. American Geophysical Union (AGU), pp. 9–28. https://doi.org/10.1029/WS006p0009

Gupta, N., Hauser, R., 2007. Kalman Filtering with Equality and Inequality State Constraints. arXiv:0709.2791 [physics].

Habib, E., Qin, L., Seo, D.-J., Ciach, G.J., Nelson, B.R., 2013. Independent Assessment of Incremental Complexity in NWS Multisensor Precipitation Estimator Algorithms. J. Hydrol. Eng. 18, 143–155. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000638

Habibi, H., Dasgupta, I., Noh, S., Kim, S., Zink, M., Seo, D.-J., Bartos, M., Kerkez, B., 2019. High-resolution hydrologic forecasting for very large urban areas. Journal of Hydroinformatics 21, 441–454. https://doi.org/10.2166/hydro.2019.100

Habibi, H., Rafieeinasab, A., Norouzi, A., Nazari, B., Seo, D.-J., Muttiah, R., Davis, C., 2016. High resolution flash flood forecasting for the Dallas-Fort Worth metroplex. Journal of Water Management Modeling. https://doi.org/10.14796/JWMM.C401

Habibi, H., Seo, D.-J., 2018. Simple and modular integrated modeling of storm drain network with gridded distributed hydrologic model via grid-rendering of storm drains for large urban areas. Journal of Hydrology 567, 637–653. https://doi.org/10.1016/j.jhydrol.2018.10.037

Halko, N., Martinsson, P.G., Tropp, J.A., 2011. Finding Structure with Randomness: Probabilistic Algorithms for Constructing Approximate Matrix Decompositions. SIAM Rev. 53, 217–288. https://doi.org/10.1137/090771806

Harmel, R.D., Cooper, R.J., Slade, R.M., Haney, R.L., Arnold, J.G., 2006. Cumulative uncertainty in measured streamflow and water quality data for small watersheds. Transactions of the ASABE 49, 689–701. https://doi.org/10.13031/2013.20488

Hausman, J., 2001. Mismeasured Variables in Econometric Analysis: Problems from the Right and Problems from the Left. The Journal of Economic Perspectives 15, 57–67.

Herman, G.R., Schumacher, R.S., 2016. Extreme Precipitation in Models: An Evaluation. Weather and Forecasting 31, 1853–1879. https://doi.org/10.1175/WAF-D-16-0093.1

Hersbach, H., 2000. Decomposition of the Continuous Ranked Probability Score for Ensemble Prediction Systems. Wea. Forecasting 15, 559–570. https://doi.org/10.1175/1520-0434(2000)015<0559:DOTCRP>2.0.CO;2

Hoerl, A.E., 1962. Application of ridge analysis to regression problems. Chemical Engineering Progress 58, 54–59.

Hoerl, A.E., Kennard, R.W., 1970. Ridge Regression: Biased Estimation for Nonorthogonal Problems. Technometrics 12, 55–67. https://doi.org/10.1080/00401706.1970.10488634

Horton, R.E., 1941. An Approach Toward a Physical Interpretation of Infiltration-Capacity. Soil Science Society of America Journal 5, 399–417. https://doi.org/10.2136/sssaj1941.036159950005000C0075x

Houtekamer, P.L., Zhang, F., 2016. Review of the Ensemble Kalman Filter for Atmospheric Data Assimilation. Monthly Weather Review 144, 4489–4532. https://doi.org/10.1175/MWR-D-15-0440.1

Jacobs, J.M., Mohanty, B.P., Hsu, E.-C., Miller, D., 2004. SMEX02: Field scale variability, time stability and similarity of soil moisture. Remote Sensing of Environment, 2002 Soil Moisture Experiment (SMEX02) 92, 436–446. https://doi.org/10.1016/j.rse.2004.02.017

Jain, A., Krishnamurthy, P.K., 2016. Phase Noise Tracking and Compensation in Coherent Optical Systems Using Kalman Filter. IEEE Communications Letters 20, 1072–1075. https://doi.org/10.1109/LCOMM.2016.2550429

Jiang, Y., Liao, M., Zhou, Z., Shi, X., Zhang, L., Balz, T., 2016. Landslide Deformation Analysis by Coupling Deformation Time Series from SAR Data with Hydrological Factors through Data Assimilation. Remote Sensing 8, 179. https://doi.org/10.3390/rs8030179

Jolliffe, I.T., Stephenson, D.B., 2003. Forecast Verification: A Practitioner’s Guide in Atmospheric Science. John Wiley & Sons.

Jozaghi, A., Nabatian, M., Noh, S., Seo, D.-J., Tang, L., Zhang, J., 2019. Improving multisensor precipitation estimation via adaptive conditional bias–penalized merging of rain gauge data and remotely sensed quantitative precipitation estimates. Journal of Hydrometeorology 20, 2347–2365. https://doi.org/10.1175/JHM-D-19-0129.1

Jozaghi, A., Shen, H., Ghazvinian, M., Seo, D.-J., Zhang, Y., Welles, E., Reed, S., 2020. Multi-Model Streamflow Prediction Using Conditional Bias-Penalized Multiple Linear Regression. Stochastic Environmental Research and Risk Assessment.

Kalman, R.E., 1960. A New Approach to Linear Filtering and Prediction Problems. Journal of Basic Engineering 82, 35. https://doi.org/10.1115/1.3662552

Kean, J.W., Smith, J.D., 2004. Flow and boundary shear stress in channels with woody bank vegetation. Riparian vegetation and fluvial geomorphology 8, 237–252.

Kean, J.W., Smith, J.D., 2005. Generation and verification of theoretical rating curves in the Whitewater River basin, Kansas. Journal of Geophysical Research: Earth Surface 110. https://doi.org/10.1029/2004JF000250

Kean, J.W., Smith, J.D., 2010. Calculation of stage-discharge relations for gravel bedded channels. Journal of Geophysical Research: Earth Surface 115. https://doi.org/10.1029/2009JF001398

Kim, B., Seo, D.-J., Noh, S.J., Prat, O.P., Nelson, B.R., 2018. Improving multisensor estimation of heavy-to-extreme precipitation via conditional bias-penalized optimal estimation. Journal of Hydrology 556, 1096–1109. https://doi.org/10.1016/j.jhydrol.2016.10.052

Kim, S., Seo, D.-J., Riazi, H., Shin, C., 2014. Improving water quality forecasting via data assimilation – Application of maximum likelihood ensemble filter to HSPF. Journal of Hydrology 519, 2797–2809. https://doi.org/10.1016/J.JHYDROL.2014.09.051

Koren, V., Reed, S., Smith, M., Zhang, Z., Seo, D.-J., 2004. Hydrology laboratory research modeling system (HL-RMS) of the US national weather service. Journal of Hydrology 291, 297–318. https://doi.org/10.1016/j.jhydrol.2003.12.039

Krishnamoorthy, A., Menon, D., 2013. Matrix Inversion Using Cholesky Decomposition. arXiv:1111.4144 [cs].

Krzysztofowicz, R., 1999. Bayesian theory of probabilistic forecasting via deterministic hydrologic model. Water Resources Research 35, 2739–2750. https://doi.org/10.1029/1999WR900099

Kurtz, W., Franssen, H.-J.H., Vereecken, H., 2012. Identification of time-variant river bed properties with the ensemble Kalman filter. Water Resources Research 48. https://doi.org/10.1029/2011WR011743

Kuzmin, V., Seo, D.-J., Koren, V., 2008. Fast and efficient optimization of hydrologic model parameters using a priori estimates and stepwise line search. Journal of Hydrology 353, 109–128. https://doi.org/10.1016/j.jhydrol.2008.02.001

Lee, H., S. J. Noh, S. Kim, H. Shen, D.-J. Seo, Y. Zhang, 2018. Improving Flood Forecasting Using Conditional Bias-Penalized Ensemble Kalman Filter. Journal of Hydrology.

Lee, H., Seo, D.-J., 2014. Assimilation of hydrologic and hydrometeorological data into distributed hydrologic model: Effect of adjusting mean field bias in radar-based precipitation estimates. Advances in Water Resources 74, 196–211. https://doi.org/10.1016/j.advwatres.2014.09.002

Lee, H., Seo, D.-J., Koren, V., 2011. 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. Advances in Water Resources 34, 1597–1615. https://doi.org/10.1016/j.advwatres.2011.08.012

Lee, H., Seo, D.-J., Liu, Y., Koren, V., McKee, P., Corby, R., 2012. Variational assimilation of streamflow into operational distributed hydrologic models: effect of spatiotemporal scale of adjustment. Hydrology and Earth System Sciences 16, 2233–2251. https://doi.org/10.5194/hess-16-2233-2012

Lee, H., Seo, D.-J., Noh, S.J., 2016. A weakly-constrained data assimilation approach to address rainfall-runoff model structural inadequacy in streamflow prediction. Journal of Hydrology 542, 373–391. https://doi.org/10.1016/j.jhydrol.2016.09.009

Lee, H., Shen, H., Noh, S.J., Kim, S., Seo, D.-J., Zhang, Y., 2019. Improving flood forecasting using conditional bias-penalized ensemble Kalman filter. Journal of Hydrology 575, 596–611. https://doi.org/10.1016/j.jhydrol.2019.05.072

Lee, H., Zhang, Y., Seo, D.-J., Xie, P., 2015. Utilizing satellite precipitation estimates for streamflow forecasting via adjustment of mean field bias in precipitation data and assimilation of streamflow observations. Journal of Hydrology 529, 779–794. https://doi.org/10.1016/j.jhydrol.2015.08.057

Liu, Y., Brown, J., Demargne, J., Seo, D.-J., 2011. A wavelet-based approach to assessing timing errors in hydrologic predictions. Journal of Hydrology 397, 210–224. https://doi.org/10.1016/j.jhydrol.2010.11.040

Liu, Y., Gupta, H.V., 2007. Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework: HYDROLOGIC DATA ASSIMILATION. Water Resour. Res. 43. https://doi.org/10.1029/2006WR005756

Liu, Y., Weerts, A.H., Clark, M., Hendricks Franssen, H.-J., Kumar, S., Moradkhani, H., Seo, D.-J., Schwanenberg, D., Smith, P., van Dijk, A.I.J.M., van Velzen, N., He, M., Lee, H., Noh, S.J., Rakovec, O., Restrepo, P., 2012. Advancing data assimilation in operational hydrologic forecasting: progresses, challenges, and emerging opportunities. Hydrology and Earth System Sciences 16, 3863–3887. https://doi.org/10.5194/hess-16-3863-2012

Lorentzen, R.J., Naevdal, G., 2011. An iterative ensemble kalman filter. IEEE Transactions on Automatic Control 56, 1990–1995. https://doi.org/10.1109/TAC.2011.2154430

Loveland, T.R., Merchant, J.W., Brown, J.F., Ohlen, D.O., Reed, B.C., Olson, P., Hutchinson, J., 1995. Seasonal Land-Cover Regions of the United States. Annals of the Association of American Geographers 85, 339–355. https://doi.org/10.1111/j.1467-8306.1995.tb01797.x-i1

Lu, X., Wang, L., Wang, H., Wang, X., 2016. Kalman filtering for delayed singular systems with multiplicative noise. IEEE/CAA Journal of Automatica Sinica 3, 51–58. https://doi.org/10.1109/JAS.2016.7373762

Lv, H., Qi, F., Zhang, Y., Jiao, T., Liang, F., Li, Z., Wang, J., 2016. Improved Detection of Human Respiration Using Data Fusion Basedon a Multistatic UWB Radar. Remote Sensing 8, 773. https://doi.org/10.3390/rs8090773

Ma, R., Zhang, L., Tian, X., Zhang, J., Yuan, W., Zheng, Y., Zhao, X., Kato, T., 2017. Assimilation of Remotely-Sensed Leaf Area Index into a Dynamic Vegetation Model for Gross Primary Productivity Estimation. Remote Sensing 9, 188. https://doi.org/10.3390/rs9030188

Mao, Y., Crow, W.T., Nijssen, B., 2019. A Framework for Diagnosing Factors Degrading the Streamflow Performance of a Soil Moisture Data Assimilation System. Journal of Hydrometeorology 20, 79–97. https://doi.org/10.1175/JHM-D-18-0115.1

Margulis, S.A., McLaughlin, D., Entekhabi, D., Dunne, S., 2002. Land data assimilation and estimation of soil moisture using measurements from the Southern Great Plains 1997 Field Experiment. Water Resour. Res. 38, 35-1-35–18. https://doi.org/10.1029/2001WR001114

Mazzoleni, M., Noh, S.J., Lee, H., Liu, Y., Seo, D.-J., Amaranto, A., Alfonso, L., Solomatine, D.P., 2018. Real-time assimilation of streamflow observations into a hydrological routing model: effects of model structures and updating methods. Hydrological Sciences Journal 63, 386–407. https://doi.org/10.1080/02626667.2018.1430898

Mohanty, B.P., Skaggs, T.H., Famiglietti, J.S., 2000. Analysis and mapping of field-scale soil moisture variability using high-resolution, ground-based data during the Southern Great Plains 1997 (SGP97) Hydrology Experiment. Water Resour. Res. 36, 1023–1031. https://doi.org/10.1029/1999WR900360

Moore, R.B., McKay, L.D., Rea, A.H., Bondelid, T.R., Price, C.V., Dewald, T.G., Johnston, C.M., 2019. User’s guide for the national hydrography dataset plus (NHDPlus) high resolution (USGS Numbered Series No. 2019–1096), User’s guide for the national hydrography dataset plus (NHDPlus) high resolution, Open-File Report. U.S. Geological Survey, Reston, VA. https://doi.org/10.3133/ofr20191096

Moore, R.J., 1985. The probability-distributed principle and runoff production at point and basin scales. Hydrological Sciences Journal 30, 273–297. https://doi.org/10.1080/02626668509490989

Moradkhani, H., Sorooshian, S., Gupta, H.V., Houser, P.R., 2005. Dual state–parameter estimation of hydrological models using ensemble Kalman filter. Advances in Water Resources 28, 135–147. https://doi.org/10.1016/J.ADVWATRES.2004.09.002

Muñoz-Sabater, J., 2015. Incorporation of Passive Microwave Brightness Temperatures in the ECMWF Soil Moisture Analysis. Remote Sensing 7, 5758–5784. https://doi.org/10.3390/rs70505758

Nair, A., Indu, J., 2016. Enhancing Noah Land Surface Model Prediction Skill over Indian Subcontinent by Assimilating SMOPS Blended Soil Moisture. Remote Sensing 8, 976. https://doi.org/10.3390/rs8120976

Nazari, B., Seo, D.-J., 2020. Analytical solution for nonlinear reservoir routing with power-law storage function, under review. ASCE Journal of Hydraulic Engineering.

Noh, S., Lee, J.-H., Lee, S., Seo, D.-J., 2019. Retrospective Dynamic Inundation Mapping of Hurricane Harvey Flooding in the Houston Metropolitan Area Using High-Resolution Modeling and High-Performance Computing. Water 11, 597. https://doi.org/10.3390/w11030597

Noh, S.J., Weerts, A.H., Rakovec, O., Lee, H., Seo, D.-J., 2018. Assimilation of streamflow observations, in: Duan, Q., Pappenberger, F., Thielen, J., Wood, A., Cloke, H.L., Schaake, J.C. (Eds.), Handbook of Hydrometeorological Ensemble Forecasting. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1–36. https://doi.org/10.1007/978-3-642-40457-3_33-2

Norouzi, A., 2016. Improving Hydrologic Prediction for Large Urban Areas Through Stochastic Analysis of Scale-Dependent Runoff Response, Advanced Sensing and High-Resolution Modeling. University of Texas Arlington.

Norouzi, A., Habibi, H., Nazari, B., Noh, S.J., Seo, D.-J., Zhang, Y., 2019. Toward parsimonious modeling of frequency of areal runoff from heavy-to-extreme precipitation in large urban areas under changing conditions: a derived moment approach. Stoch Environ Res Risk Assess 33, 1263–1281. https://doi.org/10.1007/s00477-019-01698-8

NRC, 2012. Challenges and Opportunities in the Hydrologic Sciences. National Academies Press, Washington, D.C. https://doi.org/10.17226/13293

NWS, 2008. HYDROLOGY LABORATORY-RESEARCH DISTRIBUTED HYDROLOGIC MODEL (HL-RDHM) USER MANUAL V. 2.4.2.

Press, W.H. (Ed.), 2007. Numerical recipes: the art of scientific computing, 3rd ed. ed. Cambridge University Press, Cambridge, UK ; New York.

Rafieeinasab, A., Norouzi, A., Kim, S., Habibi, H., Nazari, B., Seo, D.-J., Lee, H., Cosgrove, B., Cui, Z., 2015a. Toward high-resolution flash flood prediction in large urban areas – Analysis of sensitivity to spatiotemporal resolution of rainfall input and hydrologic modeling. Journal of Hydrology, Hydrologic Applications of Weather Radar 531, 370–388. https://doi.org/10.1016/j.jhydrol.2015.08.045

Rafieeinasab, A., Norouzi, A., Mathew, T., Seo, D.J., Chen, H., Chandrasekar, V., Rees, P., Nelson, B., 2014a. Comparative evaluation of multiple radar-based QPEs for North Texas, in: International Symposium Weather Radar and Hydrology. Reston, VA, pp. 7–10.

Rafieeinasab, A., Norouzi, A., Nelson, B., 2015b. Improving high-resolution quantitative precipitation estimation via fusion of multiple radar-based precipitation products. Journal of Hydrology 531, 320–336. https://doi.org/10.1016/J.JHYDROL.2015.04.066

Rafieeinasab, A., Seo, D.-J., Lee, H., Kim, S., 2014b. Comparative evaluation of maximum likelihood ensemble filter and ensemble Kalman filter for real-time assimilation of streamflow data into operational hydrologic models. Journal of Hydrology 519, 2663–2675. https://doi.org/10.1016/J.JHYDROL.2014.06.052

Rakovec, O., Weerts, A.H., Hazenberg, P., Torfs, P.J.J.F., Uijlenhoet, R., 2012. State updating of a distributed hydrological model with Ensemble Kalman Filtering: effects of updating frequency and observation network density on forecast accuracy. Hydrol. Earth Syst. Sci. 16, 3435–3449. https://doi.org/10.5194/hess-16-3435-2012

Rakovec, O., Weerts, A.H., Sumihar, J., Uijlenhoet, R., 2015. Operational aspects of asynchronous filtering for flood forecasting. Hydrology and Earth System Sciences 19, 2911–2924. https://doi.org/10.5194/hess-19-2911-2015

Reichle, R.H., 2008. Data assimilation methods in the Earth sciences. Advances in Water Resources 31, 1411–1418. https://doi.org/10.1016/j.advwatres.2008.01.001

Reichle, R.H., McLaughlin, D.B., Entekhabi, D., 2002. Hydrologic Data Assimilation with the Ensemble Kalman Filter. Mon. Wea. Rev. 130, 103–114. https://doi.org/10.1175/1520-0493(2002)130<0103:HDAWTE>2.0.CO;2

Riazi, H., Kim, S., Seo, D.-J., Shin, C., Kim, K., 2016. Improving operational water quality forecasting with ensemble data assimilation. JWMM. https://doi.org/10.14796/JWMM.C413

Rodgers, C.D., 2000. Inverse Methods for Atmospheric Sounding: Theory and Practice, Series on Atmospheric, Oceanic and Planetary Physics. World Scientific. https://doi.org/10.1142/3171

Roe, J., Dietz, C., Restrepo, P., Halquist, J., Hartman, R., Horwood, R., Olsen, B., Opitz, H., Shedd, R., Welles, E., 2010. NOAA’s community hydrologic prediction system, in: Proceedings from the 4th Federal Interagency Hydrologic Modeling Conference.

Saghafian, B., 2006. Nonlinear transformation of unit hydrograph. Journal of Hydrology 596–603. https://doi.org/10.1016/j.jhydrol.2006.04.026

Schaake, J.C., Koren, V.I., Duan, Q.-Y., Mitchell, K., Chen, F., 1996. Simple water balance model for estimating runoff at different spatial and temporal scales. Journal of Geophysical Research: Atmospheres 101, 7461–7475. https://doi.org/10.1029/95JD02892

Schaefer, G.L., Cosh, M.H., Jackson, T.J., 2007. The USDA Natural Resources Conservation Service Soil Climate Analysis Network (SCAN). Journal of Atmospheric and Oceanic Technology 24, 2073–2077. https://doi.org/10.1175/2007JTECHA930.1

Schweppe, F.C., 1973. Uncertain dynamic systems. Prentice-Hall.

Seber, G.A.F. (George A.F., Wild, C.J. (Christopher J., 1989. Nonlinear regression. Wiley, New York.

Seo, D.-J., 1998. Real-time estimation of rainfall fields using radar rainfall and rain gage data. Journal of Hydrology 208, 37–52. https://doi.org/10.1016/S0022-1694(98)00141-3

Seo, D.-J., 2012. Conditional bias-penalized kriging (CBPK). Stoch Environ Res Risk Assess 27, 43–58. https://doi.org/10.1007/s00477-012-0567-z

Seo, D.-J., Breidenbach, J.P., Johnson, E.R., 1999. Real-time estimation of mean field bias in radar rainfall data. Journal of Hydrology 223, 131–147. https://doi.org/10.1016/S0022-1694(99)00106-7

Seo, D.-J., Cajina, L., Corby, R., Howieson, T., 2009. Automatic state updating for operational streamflow forecasting via variational data assimilation. Journal of Hydrology 367, 255–275. https://doi.org/10.1016/j.jhydrol.2009.01.019

Seo, D.-J., Herr, H.D., Schaake, J.C., 2006. A statistical post-processor for accounting of hydrologic uncertainty in short-range ensemble streamflow prediction. Hydrology and Earth System Sciences Discussions 3, 1987–2035. https://doi.org/10.5194/hessd-3-1987-2006

Seo, D.-J., Koren, V., Cajina, N., 2003. Real-time variational assimilation of hydrologic and hydrometeorological data into operational hydrologic forecasting. Journal of Hydrometeorology 4, 627–641. https://doi.org/10.1175/1525-7541(2003)004<0627:RVAOHA>2.0.CO;2

Seo, D.-J., Saifuddin, M.M., Lee, H., 2018a. Conditional bias-penalized Kalman filter for improved estimation and prediction of extremes. Stochastic Environmental Research and Risk Assessment 32, 183–201. https://doi.org/10.1007/s00477-017-1442-8

Seo, D.-J., Saifuddin, M.M., Lee, H., 2018b. Correction to: Conditional bias-penalized Kalman filter for improved estimation and prediction of extremes. Stoch Environ Res Risk Assess 32, 3561–3562. https://doi.org/10.1007/s00477-018-1626-x

Seo, D.-J., Seed, A., Delrieu, G., 2010. Radar-based rainfall estimation, in: AGU Book Volume on Rainfall: State of the Science, F. Testik and M. Gebremichael, Editors., Geophysical Monograph Series.

Seo, D.-J., Siddique, R., Zhang, Y., Kim, D., 2014. Improving real-time estimation of heavy-to-extreme precipitation using rain gauge data via conditional bias-penalized optimal estimation. Journal of Hydrology 519, 1824–1835. https://doi.org/10.1016/j.jhydrol.2014.09.055

Shen, H., Lee, H., Seo, D.-J., 2019. Adaptive conditional bias-penalized Kalman filter for improved estimation of extremes and its approximation for reduced computation. arXiv:1908.00482 [eess].

Smith, J.A., Krajewski, W.F., 1991. Estimation of the Mean Field Bias of Radar Rainfall Estimates. J. Appl. Meteor. 30, 397–412. https://doi.org/10.1175/1520-0450(1991)030<0397:EOTMFB>2.0.CO;2

Smith, J.A., Seo, D.J., Baeck, M.L., Hudlow, M.D., 1996. An Intercomparison Study of NEXRAD Precipitation Estimates. Water Resour. Res. 32, 2035–2045. https://doi.org/10.1029/96WR00270

Sorooshian, S., Dracup, J.A., 1980. Stochastic parameter estimation procedures for hydrologie rainfall-runoff models: Correlated and heteroscedastic error cases. Water Resources Research 16, 430–442. https://doi.org/10.1029/WR016i002p00430

Stisen, S., Sandholt, I., 2010. Evaluation of remote-sensing-based rainfall products through predictive capability in hydrological runoff modelling. Hydrol. Process. 24, 879–891. https://doi.org/10.1002/hyp.7529

Tang, B., 1993. Orthogonal array-based latin hypercubes. Journal of the American Statistical Association 88, 1392–1397. https://doi.org/10.1080/01621459.1993.10476423

Tibshirani, R., 1996. Regression Shrinkage and Selection Via the Lasso. Journal of the Royal Statistical Society: Series B (Methodological) 58, 267–288. https://doi.org/10.1111/j.2517-6161.1996.tb02080.x

Tibshirani, R., 1997. The Lasso Method for Variable Selection in the Cox Model. Statistics in Medicine 16, 385–395. https://doi.org/10.1002/(SICI)1097-0258(19970228)16:4<385::AID-SIM380>3.0.CO;2-3

Tikhonov, A.N., Arsenin, V.I., 1977. Solutions of ill-posed problems, Scripta series in mathematics. Winston ; distributed solely by Halsted Press, Washington : New York.

Tikhonov, A.N., Goncharsky, A.V., Stepanov, V.V., Yagola, A.G., 1995. Numerical Methods for the Solution of Ill-Posed Problems. Springer Netherlands, Dordrecht. https://doi.org/10.1007/978-94-015-8480-7

USDA, 1986. Urban hydrology for small watersheds. 2nd edition (Technical Report). U.S. Dept. of Agriculture, Soil Conservation Service, Engineering Division.

van Velzen, N., Altaf, M.U., Verlaan, M., 2016. OpenDA-NEMO framework for ocean data assimilation. Ocean Dynamics 66, 691–702. https://doi.org/10.1007/s10236-016-0945-z

Wallace, L., Lucieer, A., Watson, C., Turner, D., 2012. Development of a UAV-LiDAR System with Application to Forest Inventory. Remote Sensing 4, 1519–1543. https://doi.org/10.3390/rs4061519

Wang, D., Chen, Y., Cai, X., 2009. State and parameter estimation of hydrologic models using the constrained ensemble Kalman filter. Water Resour. Res. 45. https://doi.org/10.1029/2008WR007401

Weerts, A.H., El Serafy, G.Y.H., 2006. Particle filtering and ensemble Kalman filtering for state updating with hydrological conceptual rainfall-runoff models: PARTICLE AND ENSEMBLE KALMAN FILTERING. Water Resour. Res. 42. https://doi.org/10.1029/2005WR004093

Wi, S., Yang, Y.C.E., Steinschneider, S., Khalil, A., Brown, C.M., 2015. Calibration approaches for distributed hydrologic models in poorly gaged basins: implication for streamflow projections under climate change. Hydrology and Earth System Sciences 19, 857–876. https://doi.org/10.5194/hess-19-857-2015

Wickham, J., Stehman, S.V., Neale, A.C., Mehaffey, M., 2020. Accuracy assessment of NLCD 2011 percent impervious cover for selected USA metropolitan areas. International Journal of Applied Earth Observation and Geoinformation 84, 101955. https://doi.org/10.1016/j.jag.2019.101955

Wilks, D.S., 2011. Statistical methods in the atmospheric sciences, 3rd ed. ed, International geophysics series. Elsevier/Academic Press, Amsterdam ; Boston.

WMO (Ed.), 1992. Simulated real-time intercomparison of hydrological models, Operational hydrology report. World Meteorological Organization, Geneva, Switzerland.

Yan, M., Tian, X., Li, Z., Chen, E., Wang, X., Han, Z., Sun, H., 2016. Simulation of Forest Carbon Fluxes Using Model Incorporation and Data Assimilation. Remote Sensing 8, 567. https://doi.org/10.3390/rs8070567

Yilmaz, K.K., Gupta, H.V., Wagener, T., 2008. A process-based diagnostic approach to model evaluation: Application to the NWS distributed hydrologic model: PROCESS-BASED DIAGNOSTIC EVALUATION OF HYDROLOGIC MODEL. Water Resour. Res. 44. https://doi.org/10.1029/2007WR006716

Young, C.B., Bradley, A.A., Krajewski, W.F., Kruger, A., Morrissey, M.L., 2000. Evaluating NEXRAD Multisensor Precipitation Estimates for Operational Hydrologic Forecasting. J. Hydrometeor. 1, 241–254. https://doi.org/10.1175/1525-7541(2000)001<0241:ENMPEF>2.0.CO;2

Yu, K.K.C., Watson, N.R., Arrillaga, J., 2005. An adaptive Kalman filter for dynamic harmonic state estimation and harmonic injection tracking. IEEE Transactions on Power Delivery 20, 1577–1584. https://doi.org/10.1109/TPWRD.2004.838643

Zhang, J., Howard, K., Langston, C., Kaney, B., Qi, Y., Tang, L., Grams, H., Wang, Y., Cocks, S., Martinaitis, S., Arthur, A., Cooper, K., Brogden, J., Kitzmiller, D., 2016. Multi-Radar Multi-Sensor (MRMS) quantitative precipitation estimation: initial operating capabilities. Bulletin of the American Meteorological Society 97, 621–638. https://doi.org/10.1175/BAMS-D-14-00174.1

Zhang, J., Howard, K., Langston, C., Vasiloff, S., Kaney, B., Arthur, A., Van Cooten, S., Kelleher, K., Kitzmiller, D., Ding, F., Seo, D.-J., Wells, E., Dempsey, C., 2011. National Mosaic and Multi-Sensor QPE (NMQ) system: description, results, and future plans. Bulletin of the American Meteorological Society 92, 1321–1338. https://doi.org/10.1175/2011BAMS-D-11-00047.1

Zhou, H., Huang, H., Zhao, H., Zhao, X., Yin, X., 2017. Adaptive Unscented Kalman Filter for Target Tracking in the Presence of Nonlinear Systems Involving Model Mismatches. Remote Sensing 9, 657. https://doi.org/10.3390/rs9070657

Zupanski, D., 2009. Information measures in ensemble data assimilation, in: Park, S.K., Xu, L. (Eds.), Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications. Springer, Berlin, Heidelberg, pp. 85–95. https://doi.org/10.1007/978-3-540-71056-1_4

Zupanski, D., Hou, A.Y., Zhang, S.Q., Zupanski, M., Kummerow, C.D., Cheung, S.H., 2007. Applications of information theory in ensemble data assimilation. Quarterly Journal of the Royal Meteorological Society 133, 1533–1545. https://doi.org/10.1002/qj.123