The Quest for Model Uncertainty Quantification: A Hybrid Ensemble and Variational Data Assimilation Framework

Peyman Abbaszadeh[1] (ID), Hamid Moradkhani[1] (ID), and Dacian N. Daescu[2]

[1]Center for Complex Hydrosystems Research, Department of Civil, Construction and Environmental Engineering, University of Alabama, Tuscaloosa, AL, USA, [2]Fariborz Maseeh Department of Mathematics and Statistics, Portland State University, Portland, OR, USA

Abstract This article presents a novel approach to couple a deterministic four-dimensional variational (4DVAR) assimilation method with the particle filter (PF) ensemble data assimilation system, to produce a robust approach for dual-state-parameter estimation. In our proposed method, the Hybrid Ensemble and Variational Data Assimilation framework for Environmental systems (HEAVEN), we characterize the model structural uncertainty in addition to model parameter and input uncertainties. The sequential PF is formulated within the 4DVAR system to design a computationally efficient feedback mechanism throughout the assimilation period. In this framework, the 4DVAR optimization produces the maximum a posteriori estimate of state variables at the beginning of the assimilation window without the need to develop the adjoint of the forecast model. The 4DVAR solution is then perturbed by a newly defined prior error covariance matrix to generate an initial condition ensemble for the PF system to provide more accurate and reliable posterior distributions within the same assimilation window. The prior error covariance matrix is updated from one cycle to another over the main assimilation period to account for model structural uncertainty resulting in an improved estimation of posterior distribution. The premise of the presented approach is that it (1) accounts for all sources of uncertainties involved in hydrologic predictions, (2) uses a small ensemble size, and (3) precludes the particle degeneracy and sample impoverishment. The proposed method is applied on a nonlinear hydrologic model and the effectiveness, robustness, and reliability of the method is demonstrated for several river basins across the United States.

1. Introduction

Soil moisture and streamflow are among those key environmental variables that greatly affect flood forecasting, drought monitoring, and agricultural production that all collectively control the land and atmospheric system. Although, theoretically, these quantities can be estimated through hydrologic modeling, in practice they are often biased or erroneous due to the presence of uncertainties in all layers of hydrologic predictions. Data assimilation (DA) has been well received in the hydrologic community as one of the most effective methods in characterizing the aforementioned uncertainties while estimating parameters, prognostic, and diagnostic variables (Abbaszadeh et al., 2018; Clark et al., 2008; Moradkhani, Sorooshian, et al., 2005; Moradkhani et al., 2018; Pathiraja et al., 2016; Vrugt et al., 2006).

Generally, DA is defined as the application of Bayes' theorem to probabilistically condition the states of a dynamical model on observations (Moradkhani et al., 2018). A plethora of techniques is available to assimilate observations into a model for better initialization of the system and quantification of model parameter uncertainties. They all have some overlapping features making it difficult to define a clear-cut classification. Bayesian data assimilation seeks probabilistic estimates of state variables of interest in order to characterize their uncertainties. These probability distributions are sequentially adjusted according to the Bayes' theorem to better match the observations. In the hydrologic community, the best known and ubiquitous Bayesian approach is the ensemble Kalman filter (EnKF; Crow & Wood, 2003; De Lannoy et al., 2007; Moradkhani, Hsu, et al., 2005; Reichle et al., 2002). Despite the widespread use of the EnKF and its different variants in numerous hydrologic applications, it is subject to some inherent limitations that result in suboptimal performance of this technique. These include (1) the linear updating rule, (2) Gaussian assumption of errors in observations, and (3) violation of water balance (e.g., DeChant & Moradkhani, 2012; Matgen et al., 2010; Noh et al., 2011; Plaza et al., 2012). PF as an effective alternative to EnKF has emerged for applications in nonlinear and non-Gaussian systems (Abbaszadeh et al., 2018; DeChant & Moradkhani,

2011; Dong et al., 2015; Moradkhani et al., 2012, Pathiraja, Moradkhani, et al., 2018; Yan et al., 2015; Yan et al., 2017). While successfully used in numerous applications, PFs may be subject to particle degeneracy and sample impoverishment. Particle degeneracy occurs when a few of the particles close to the measurement receive significant weights while others are discarded (ensemble collapse to a single point). Although resampling alleviates the degeneracy problem, it may diminish the diversity of particles and increase the number of repeated particles (known as sample impoverishment). To mitigate these issues, researchers have combined PF with several procedures, such as the Markov chain Monte Carlo (MCMC) algorithm (Andrieu et al., 2010; Moradkhani et al., 2012), metaheuristic techniques (Han et al., 2011; Kwok et al., 2005; Yin et al., 2015), and a combination of these two (Abbaszadeh et al., 2018; Liang & Wong, 2001; Zhu et al., 2018). The successful application of Bayesian data assimilation techniques owes to their stochastic nature as they enable the uncertainty quantification in forecasting systems (Verkade & Werner, 2011; Zhu et al., 2002).

However, variational data assimilation approaches are more similar to traditional and standard calibration procedures that primarily rely on nonlinear least squares optimization (Efstratiadis & Koutsoyiannis, 2010). Variational data assimilation can be considered as a Bayesian method (in the case of a quadratic cost function resulting from Gaussian statistics), which seeks to minimize a cost function defined as the departures of the simulated values from the observations within an assimilation window (Reichle et al., 2001). This class of data assimilation includes different variants, such as three-dimensional variational data assimilation (3DVAR), which estimates the state of a system at a particular time using the information at that time only, and 4DVAR, where information is propagated both forward and backward across a specified window, known as the assimilation window, to best estimate the initial conditions at the beginning of that time period. By linearizing the model's dynamics, 4DVAR efficiently finds the optimum solution in the convex search space by the use of gradient descent methods. While this component has made 4DVAR the most far-reaching approach in meteorological science, it has been less popular in hydrology due to difficulties in linearizing the hydrologic model (Liu & Gupta, 2007). In addition to this problem, in variational data assimilation, the adjoint models require increased software development efforts and are often difficult to maintain given the increased complexity incorporated in hydrologic models. Despite these obstacles, some research efforts have gone to the 4DVAR assimilation of hydrologic variables, such as soil moisture, into the land surface models (Jones et al., 2004; Lee et al., 2011; Seo et al., 2009; Tian et al., 2009).

In contrast to the PF that works in a sequential manner, the 4DVAR approach operates in a batch-processing manner by using all the observations simultaneously within the assimilation interval. Unlike Bayesian data assimilation, the variational data assimilation most often does not provide estimates of the predictive uncertainty of estimated model states and parameters, which considerably limits the value of the estimated variables in decision-making processes (Abdolghafoorian & Farhadi, 2016; Reichle et al., 2001). Given these concerns, integrating Bayesian and variational assimilation schemes seems a logical step to constrain the deficiencies of each method while benefiting from the strength of each (Atkins & Morzfeld, 2013; Chorin et al., 2010; Chorin & Tu, 2009; Hernández & Liang, 2018; Morzfeld et al., 2012; Slivinski et al., 2015; van Leeuwen, 2015; Zhu et al., 2016). Bannister (2017) presented a holistic review of these hybrid techniques and explained thoroughly their necessities in the data assimilation area.

In this paper, we present an approach to integrate the PF and 4DVAR data assimilation methods to leverage the best advantages of both in a single framework. More specifically, we seek to address the following questions that have been the main concerns in many hydrologic data assimilation studies: (1) In data assimilation, mischaracterization of errors leads to suboptimal model performance, and in the worst case, to degraded estimates, even compared to the open loop model run (Pathiraja, Moradkhani, et al., 2018; Salamon & Feyen, 2009). While many efforts have gone to characterizing the model parameter and forcing uncertainties, less attention has been given to account for the model structural uncertainty in hydrologic studies. Ignoring model structural uncertainty leads to an inaccurate and biased simulation of the hydrologic processes (Parrish et al., 2012; Pathiraja, Moradkhani, et al., 2018). Therefore, the first question is how we can benefit from the advantage of the 4DVAR method to better account for the model structural uncertainty when particle filtering is in operation. (2) Under a perfect model assumption, the observation error covariance R is the only matrix to be estimated. However, for imperfect model scenarios, the prior (background) state error covariance B should also be prescribed (Zhu

et al., 2017). In standard particle filtering, this matrix does not play a role during the updating process except for initialization of state variables. To reduce the uncertainty in initial conditions, in many geoscience applications the prior state error covariance B is inflated. This is done through either deterministic additive means with an inflation factor (Anderson & Anderson, 1999) or multiplicative inflation of the B matrix (Anderson, 2001). Therefore, the second objective of this research is to introduce a new approach to dynamically estimate and update this error covariance matrix throughout the assimilation process and thus generate more accurate and reliable posterior distributions. It is noted that the idea of using dynamic B matrix has been previously discussed in several studies (e.g., Lorenc, 2017; Wang et al., 2008; Wu et al., 2013).

The remainder of this paper is organized as follows: Section 2 briefly describes the PF and 4DVAR data assimilation approaches. In this section, we elaborate on the formulation of our proposed approach, which integrates both methods. Section 3 presents the experimental design, including the study areas, multiple performance measures used in this study, and DA settings for both synthetic and real case studies. Section 4 discusses the performance of the proposed hybrid approach over a synthetic experiment and several real case studies. Section 5 concludes the results and presents some recommendations and suggestions for future work.

2. Methodology

In this study, instead of using standard PF, we use the enhanced version of this method, the evolutionary particle filter with MCMC (EPFM), recently developed by the authors of this paper (Abbaszadeh et al., 2018). EPFM is a data assimilation technique that uses the evolutionary concept of the Genetic Algorithm (GA) combined with MCMC technique to effectively shuffle the particles before the resampling step of the filtering to produce a more complete representation of the posterior distribution for both state variables and parameters. EPFM significantly mitigates the particle degeneracy and sample impoverishment, the issues that mainly result in the suboptimal performance of the particle filtering and sometimes even the failure of this approach. The proposed method is a combination of the EPFM and 4DVAR methods. It should be noted that although we use EPFM approach as a sequential data assimilation method in this study to introduce our hybrid approach, any variant of PF can be used as well. Section 2.2 presents a comprehensive description of 4DVAR data assimilation approach. Linearizing the observation operator h and forecast model operator $\mathcal{M}$ often are not achievable or feasible for hydrologic models. To circumvent this issue, we formulate our hybrid model in such a way that relaxes the linearization constraints of 4DVAR method in hydrologic applications. This issue will be further discussed in section 2.3, where we explain the proposed HEAVEN (Hybrid Ensemble And Variational Data Assimilation Method for ENvironmental Systems) method.

The following two equations describe the generic nonlinear dynamic system.

$$x_t = \mathcal{M}(x_{t-1}, u_t, \theta) + \omega_t \qquad \omega_t \sim N(0, Q_t) \tag{1}$$

$$y_t = h(x_t) + \upsilon_t \qquad \upsilon_t \sim N(0, R_t) \tag{2}$$

where $x_t \in \mathbb{R}^n$ and $\theta \in \mathbb{R}^d$ are vectors of uncertain state variables and model parameters, respectively. u_t represents the uncertain forcing data, $y_t \in \mathbb{R}^m$ indicates a vector of observation data, ω_t and υ_t are the model and measurement errors, respectively. More often, ω_t and υ_t are assumed to be independent and follow white noises with mean zero and covariance Q_t and R_t, respectively (Gaussian distribution). Based on these definitions, the EPFM, 4DVAR, and HEAVEN approaches are formulated, respectively, in the following subsections.

2.1. EPFM Data Assimilation Method

The EPFM is a sequential data assimilation approach built based on recently developed PF-MCMC method (Moradkhani et al., 2012) and GA-MCMC approach (Abbaszadeh et al., 2018). GA-MCMC is an evolutionary Monte Carlo approach that can be used within any particle filtering algorithm to preprocess the ensemble members toward enhancing the assimilation results. Such an operation simultaneously alleviates the particle degeneration phenomenon by intensifying the particles' diversity and generating more accurate and

reliable estimation of posterior distributions. In this method, we use the following formulas which are based on Bayes theorem to calculate the posterior distribution of the state variables at time t.

$$p(x_t|y_{1:t}) = p(x_t|y_{1:t-1}, y_t) = \frac{p(y_t|x_t)p(x_t|y_{1:t-1})}{p(y_t|y_{1:t-1})} = \frac{p(y_t|x_t)p(x_t|y_{1:t-1})}{\int p(y_t|x_t)p(x_t|y_{1:t-1})dx_t} \tag{3}$$

$$p(x_t|y_{1:t-1}) = \int p(x_t, x_{t-1}|y_{1:t-1})\mathrm{d}x_{t-1} = \int p(x_t|x_{t-1})p(x_{t-1}|y_{1:t-1})\mathrm{d}x_{t-1} \tag{4}$$

where $p(y_t|x_t)$ is the likelihood for time step t, $p(x_t|y_{1:t-1})$ is the prior distribution, and $p(y_t|y_{1:t-1})$ is the normalization factor. The marginal likelihood function $p(y_{1:t})$ can be computed as

$$p(y_{1:t}) = p(y_1)\prod p(y_t|y_{1:t-1}) \tag{5}$$

where the normalization factor $p(y_t|y_{1:t-1})$ is calculated as follows:

$$p(y_t|y_{1:t-1}) = \int p(y_t, x_t|y_{1:t-1})\mathrm{d}x_t = \int p(y_t|x_t)p(x_t|y_{1:t-1})\mathrm{d}x_t \tag{6}$$

Since only in special cases the analytical solution of equation (3) is available, such as those with linear systems and white Gaussian noise, for practical reasons, the posterior distribution is not calculated directly by this equation, and instead, it is approximated using a set of particles with associated weights.

$$p(x_t|y_{1:t}) \approx \sum_{i=1}^{N} w^{i+}\delta(x_t - x_t^i) \tag{7}$$

where w^{i+}, δ, and N denote the posterior weight of the ith particle, Dirac delta function, and the ensemble size, respectively. The posterior weight is normalized as follows:

$$w^{i+} = \frac{w^{i-}.p(y_t|x_t^i, \theta_t^i)}{\sum_{i=1}^{N} w^{i-}.p(y_t|x_t^i, \theta_t^i)} \tag{8}$$

where w^{i-} is the prior particle weights, and the $p(y_t|x_t^i, \theta_t^i)$ can be computed from the likelihood $L(y_t|x_t^i, \theta_t^i)$. To calculate this, for simplicity, a Gaussian likelihood is used as follows:

$$L(y_t|x_t^i, \theta_t^i) = \frac{1}{\sqrt{(2\pi)^m|R_t|}} \exp\left[-\frac{1}{2}(y_t - h(x_t^i))^T R_t^{-1}(y_t - h(x_t^i))\right] \tag{9}$$

It is noted that the EPFM benefits from MCMC technique twice, before the resampling step, when it is combined with GA (earlier called as GA-MCMC) to shuffle the particles and generate more appropriate proposal state distribution, and after the resampling step for parameter updating. This feature of EPFM method improves the assimilation results while avoiding the particle degeneracy and sample impoverishment problems, with no need to increase the ensemble size.

Here the implementation of GA-MCMC method, as an effective tool to shuffle the particles in importance sampling step of the filtering, is explained. Weights of the particles are considered as the fitness value f_t^i.

$$f_t^i = w_t^{i+} \tag{10}$$

Next, the particles are sorted in descending order of their fitness values to implement the roulette wheel selection method. This approach is applied based on a crossover probability (ρ_c) to select a portion of particles for crossover operation. Basically, ρ_c specifies how many particles are to be selected in this step. Each pair of these selected particles undergoes an arithmetic crossover operator to generate a new pair of particles. This is formulated by the following relationships:

$$x_{t-1}^{i'j'} = \xi.x_{t-1}^{i,j'} + (1-\xi).x_{t-1}^{i,j''} \qquad j' \neq j'' \tag{11}$$

$$x_{t-1}^{i'j''} = (1-\xi).x_{t-1}^{i,j'} + \xi.x_{t-1}^{i,j''} \qquad j' \neq j'' \tag{12}$$

where $x_{t-1}^{i,j'}$ and $x_{t-1}^{i,j''}$ are the selected particles from the main ensemble $\{x_{t-1}^i\}$ by the roulette wheel selection method before the resampling step at time $t-1$, $x_{t-1}^{i'j'}$ and $x_{t-1}^{i'j''}$ is a new pair of particles at the same time, and ξ is a uniform random number which varies between 0 and 1. The parameters i and i' refer to the ensemble pool before and after crossover operator. Parameters j' and j'' indicate two particles. To further increase the diversity of the particles, a mutation strategy is also designed. The mutation operation is executed with the appropriate mutation probability (ρ_m). It is important to note that the performance of the EPFM approach is not sensitive to the choice of ρ_c and ρ_m (see Abbaszadeh et al., 2018 for more information). The parameter ρ_m specifies how many of the crossovered particles are selected for mutation operation. One of the state variables of the selected particle is then randomly altered by the following equation:

$$x_{t-1}^{i',k} = x_{t-1}^{i,k} + \eta \qquad x_{t-1}^{i,k} \in \left\{x_{t-1}^{i'j'}, x_{t-1}^{i'j''}\right\} \qquad \eta \sim N\left(0, \psi \text{Var}\left(x_{t-1}^i\right)\right) \tag{13}$$

where $N\left(0, \psi \text{Var}\left(x_{t-1}^i\right)\right)$ denotes a random sample drawn from a Gaussian distribution with mean zero and variance of $\psi \text{Var}\left(x_{t-1}^i\right)$, where $\text{Var}\left(x_{t-1}^i\right)$ is the variance of the states at time $t-1$, and ψ is a small tuning parameter, which was set to 0.01 in this approach through ad hoc process. Therefore, the new proposal state $x_{t-1}^{i,p}$ is generated through the crossover and mutation operators. The MCMC approach is then used to decide which new particles from $x_{t-1}^{i,p}$ should remain or be replaced with the old corresponding particle. This acceptance/rejection step is necessary as it ensures that an appropriate prior state distribution is constructed in each time step, and thus, a more desirable posterior distribution is estimated. To know whether to accept or reject the proposed states, the metropolis acceptance ratio α is calculated as follows:

$$\alpha = \min\left(1, \frac{p\left(x_t^{i,p}, \theta_t^{i-}|y_{1:t}\right)}{p\left(x_t^{i-}, \theta_t^{i-}|y_{1:t}\right)}\right) = \min\left(1, \frac{p\left(y_{1:t}|x_t^{i,p}, \theta_t^{i-}\right).p\left(x_t^{i,p}|\theta_t^{i-}, y_{1:t-1}\right)}{p\left(y_{1:t}|x_t^{i-}, \theta_t^{i-}\right).p\left(x_t^{i-}|\theta_t^{i-}, y_{1:t-1}\right)}\right) \tag{14}$$

where $p\left(x_t^{i,p}, \theta_t^{i-}|y_{1:t}\right)$ is the proposed joint probability distribution.

$$p\left(x_t^{i,p}, \theta_t^{i-}|y_{1:t}\right) \propto p\left(y_t|x_t^{i,p}, \theta_t^{i-}\right).p\left(x_t^{i,p}|\theta_t^{i-}, y_{1:t-1}\right).p\left(\theta_t^{i-}|y_{1:t-1}\right) \tag{15}$$

$$x_t^{i,p} = \mathscr{M}\left(x_{t-1}^{i,p}, u_t^i, \theta_t^{i-}\right) \tag{16}$$

where $p\left(y_t|x_t^{i,p}, \theta_t^{i-}\right)$ is computed similar to equation (9) and the proposal state Probability Density Function (PDF) $p\left(x_t^{i,p}|\theta_t^{i-}, y_{1:t-1}\right)$ is calculated based on an assumption that the proposal states follow the marginal Gaussian distributions with mean μ_t (equation (18)) and variance σ_t^2 (equation (19)). Although, a joint distribution is a perfect match to this scenario, we prefer to select the marginal priors due to the nonlinear nature of state variables that makes fitting a joint distribution difficult. To calculate the proposal PDF, weighted mean and variance of the Gaussian distribution are calculated as follows:

$$x_t^{i-} = \mathscr{M}\left(x_{t-1}^{i+}, u_t^i, \theta_t^{i-}\right) \tag{17}$$

$$\mu_t = \sum w_{t-1}^{i+} x_t^{i-} \tag{18}$$

$$\sigma_t^2 = \sum w_{t-1}^{i+}\left(x_t^{i-} - \mu_t\right)^2 \tag{19}$$

Up to this point, we obtained the appropriate prior state variables $\left\{x_{t-1}^{i'}\right\}$. The next step is to use this updated ensemble to recalculate the posterior weights (w^{i+}) using equation (8). For this, we use those posterior particle weights computed before GA operator implementation as prior particle weights (w^{i-}). It is important to note that in the EPFM approach we draw samples from a proposal distribution $q(x|y)$, which is generated using the GA and MCMC steps, not from the original prior as traditionally done in particle filtering applications. For more detailed information about proposal distributions, we refer the readers to Doucet and Johansen (2009), who provided a thorough description of particle filtering and its different variants. Those particles were also used for assigning fitness values (f^i) in GA operation. After this, we will do resampling to generate posterior state variables.

To do resampling, the sampling importance resampling (SIR) algorithm is suggested. In this method, we resample those particles whose probabilities are greater than the uniform probability. This approach discards the particles with smaller weights while retaining those whose weights are larger. At the end of each time step, the particle weights are again set to $1/N$. Apparently, the particles with larger weights are more likely to be drawn multiple times during the resampling step, leading to a loss of diversity in particles known as sample impoverishment. To avoid this, the resampled parameters are perturbed and new proposal distribution is formed as follows:

$$\theta_t^{i,p} = \theta_t^{i+} + \varepsilon_t^i \qquad \varepsilon_t^i \sim N\left(0, s_t \mathrm{Var}\left(\theta_t^{i-}\right)\right) \tag{20}$$

where θ_t^{i+} shows the parameter after resampling at time t, $\mathrm{Var}\left(\theta_t^{i-}\right)$ is the variance of the prior parameters at time t, and s_t is a small time variant tuning parameter. To accept or reject the proposal parameters $\theta_t^{i,p}$ we calculate the metropolis acceptance ratio α as follows:

$$\alpha = \min\left(1, \frac{p\left(x_t^{i,p'}, \theta_t^{i,p} | y_{1:t}\right)}{p\left(x_t^{i+}, \theta_t^{i+} | y_{1:t}\right)}\right) \tag{21}$$

$$p\left(x_t^{i,p'}, \theta_t^{i,p} | y_{1:t}\right) \propto p\left(y_{1:t} | x_t^{i,p'}, \theta_t^{i,p}\right) . p\left(x_t^{i,p'} | \theta_t^{i,p}, y_{1:t-1}\right) . p\left(\theta_t^{i,p} | y_{1:t-1}\right) \tag{22}$$

$$x_t^{i,p'} = \mathscr{M}\left(x_{t-1}^{i+}, u_t^{i+}, \theta_t^{i,p}\right) \tag{23}$$

where $p\left(x_t^{i,p'}, \theta_t^{i,p} | y_{1:t}\right)$ is the proposed joint probability distribution, $x_t^{i,p'}$ is a sample drawn from the proposal state distribution at time step t, u_t^{i+} is the perturbed forcing data associated with the ith particle, and x_{t-1}^{i+} is the posterior of state variables generated from $\left\{x_{t-1}^i\right\}$ in the resampling step. It should be noted that to calculate the proposed state $x_t^{i,p'}$ we use the perturbed forcing data u_t^{i+}. The optimal tuning parameter s_t in equation (20) is unknown and needs to be estimated automatically as a time-variant quantity. If s is too small, the posterior distribution is overconfident. On the contrary, if s is too large, the resampled parameters poorly characterize the posterior. To overcome this issue and calculate the s parameter, we used the variable variance multiplier method developed by Leisenring and Moradkhani (2012) and further modified by Moradkhani et al. (2012).

Up to this point, we explained the particle filtering approach used in this study. However, we should note that the implementation of PF relies on the assumption of Gaussian likelihood function given unavailability of information on the type of likelihood function. In addition, in the calculation of the metropolis acceptance ratio in the MCMC step, it is assumed that the proposal state follows the marginal Gaussian distribution. These may imply an approximation to the full Bayesian solution, which is unavoidable in hydrology for practical reasons. For future studies, one may want to explore the feasibility of space–time particle filter that characterizes the transition proposal density in the MCMC step to avoid the need to approximate the actual PDF.

2.2. The 4DVAR Data Assimilation Method

The 4DVAR approach provides an estimate (analysis) to a time-distributed sequence of state variables by minimizing a cost function defined as

$$\begin{aligned}
J(x_0, \ \dots \ , \ x_K) &= J^b + J^o + J^q \\
&= \frac{1}{2}\left(x_0 - x_{0,b}\right)^T B^{-1}\left(x_0 - x_{0,b}\right) + \frac{1}{2}\sum_{k=0}^{K}\left(y_k - h_k(x_k)\right)^T R_k^{-1}\left(y_k - h_k(x_k)\right) \\
&\quad + \frac{1}{2}\sum_{k=1}^{K}\left(x_k - M_{k-1\to k}(x_k)\right)^T Q_k^{-1}\left(x_k - M_{k-1\to}(x_{k-1})\right)
\end{aligned} \tag{24}$$

where $x_{0,b} \in \mathbb{R}^n$, $y_k \in \mathbb{R}^m$, $B \in \mathbb{R}^{n \times n}$, $R_k \in \mathbb{R}^{m \times m}$, and $Q_k \in \mathbb{R}^{n \times n}$. Error covariance matrices are the same as those defined in section 2.1. The optimal solution is the joint maximum likelihood estimate of the state variables in the interval $[t_0, t_K]$ given the observations.

The computational cost may be reduced by neglecting the model error (perfect model scenario). By imposing the model equations as strong constraint,

$$x_k = \mathcal{M}_{k-1 \to k}(x_{k-1}) \qquad k = 1 : K \tag{25}$$

where $\mathcal{M}_{k-1 \to k}$ denotes the nonlinear forecast model from t_{k-1} to t_k, the state at time t_k is expressed as

$$x_k = \mathcal{M}_{0 \to k}(x_0) \qquad k = 1 : K \tag{26}$$

In this context, the cost functional is expressed in terms of the initial state x_0 as

$$J(x_0) = J^b + J^o = \frac{1}{2}\left(x_0 - x_{0,b}\right)^T B^{-1}\left(x_0 - x_{0,b}\right) + \frac{1}{2}\sum_{k=0}^{K}\left(y_k - h_k(x_k)\right)^T R_k^{-1}\left(y_k - h_k(x_k)\right) \tag{27}$$

The optimal solution (analysis) is obtained through an iterative method that typically relies on linearized versions of the model and observational operator to obtain a quadratic approximation to the cost J (outer iteration) and adjoint modeling for gradient information. We refer the interested readers to these publications (Talagrand & Courtier, 1987; Tremolet, 2006; Zupanski, 1997), which provide more detailed information of fundamentals of the 4DVAR methods.

2.3. EPFM+4DVAR Data Assimilation Method (HEAVEN)

Here, we develop the integration of EPFM and 4DVAR data assimilation methods to create an improved methodology for which all sources of uncertainties involved in the hydrologic predictions are accounted. The main idea behind the HEAVEN method is that it provides the possibility that both sequential and variational assimilation approaches can effectively feed each other in a single framework in order to generate a more comprehensive representation of posterior distributions. HEAVEN is a framework that enables the particle filter to account for the hydrologic model structural uncertainty by using 4DVAR approach without a need to linearize the model and observation operators. In this approach, we operate the EPFM filter within the assimilation window for which the best initial condition is estimated by 4DVAR method. In doing so, the question arises as to how to use the deterministic (single) initial condition achieved by 4DVAR method to initialize the EPFM filter, which is an ensemble-based approach. To cope with this problem, we define a prior error covariance B, which involves two components: dynamic (B_d) and static (B_s) prior error covariances, to perturb the deterministic solution of 4DVAR approach and generate best initial condition for the EPFM filter. This error covariance matrix is updated and propagated from cycle-to-cycle during the whole assimilation period. Dynamic error covariance matrix B_d is calculated within the EPFM framework while static error covariance matrix B_s comes from the 4DVAR approach. In this study, we calculate B_d using the equation presented by Shaw and Daescu (2016), who developed the formula to derive the model error ensemble for the weak-constraint 4DVAR method. In the remainder of this section, we first present a schematic of the HEAVEN in Figure 1, then elaborate on the approach for better understanding and implementation. It is worth mentioning that both 4DVAR and PF (steps 1 and 4 shown in Figure 1) components of HEAVEN follow the Bayes' theorem. Steps 5 and 6 are based on the statistical principles described by Shaw and Daescu (2016), while step 2 is an ad hoc procedure. Step 3 is not considered as an individual operation as it represents a transition phase from steps 2 to 4.

For the time period of T and assimilation window size K ($[t_0, t_{k=K}]$), the number of assimilation cycles in the HEAVEN becomes $^T/_K$. For example, for a 1-year analysis period of $T = 365$ days, with the assumption of $K = 5$ days, 73 assimilation cycles or windows are defined. In each assimilation cycle, k ranges between 0 and K, where $k = 0$ indicates the initial time step.

Initialization. Three error covariance matrices, B, R_k, and Q_k should be prescribed. Observation error covariance matrix R_k at each time step can be specified as

$$R_k = \left[\max\{(\lambda \times \text{Obs}_k), 1\}\right]^2 \tag{28}$$

where λ is the error percentage in observations and Obs_k is observation at time k. Prior error covariance matrix B can be specified as
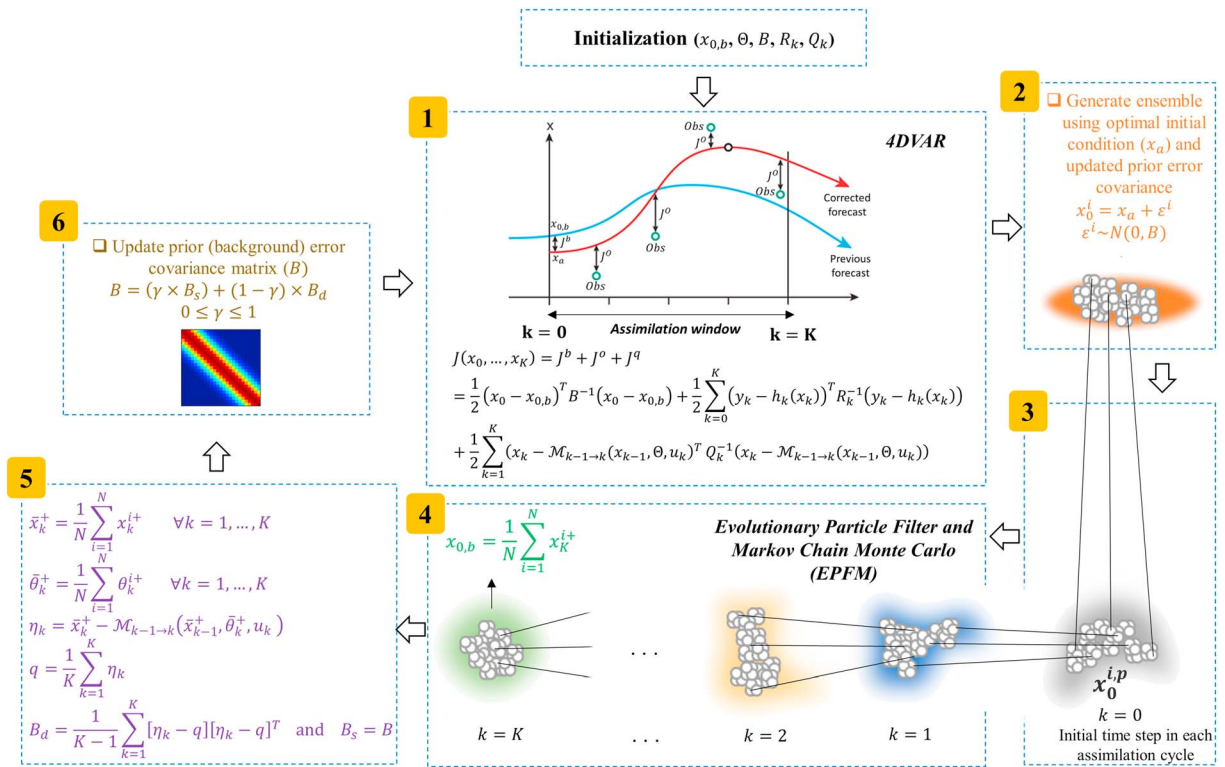
Figure 1. Schematic of the proposed HEAVEN method in a reciprocal form. B_s and B_d are, respectively, the static and dynamic prior error covariance matrices. Steps 1 through 6 indicate the procedure for one assimilation cycle. k, K, N, and i show time step in each assimilation window, assimilation window size, ensemble size, and particle index, respectively. B, R_k, and Q_k specify prior, observation, and model error covariance matrices, respectively. Initial deterministic guess for state variables and parameters are also, respectively, represented by $x_{0,b}$ and Θ. HEAVEN = Hybrid Ensemble And Variational data assimilation method for ENvironmental systems; 4DVAR = four-dimensional variational; EPFM = evolutionary particle filter with MCMC.

$$B = \text{diag}\left(\left[\Omega \times x_{0,b}\right]^2\right) \tag{29}$$

where Ω is the error percentage in initial state variables and $x_{0,b}$ is the deterministic initial guess for state variables. Similarly, the model error covariance can be calculated by the following formula:

$$Q_k = \Gamma \times \text{diag}\left(\left[\pi \times x_{0,b}\right]^2\right) \tag{30}$$

where π is the error percentage in model structure and Γ is the model error covariance inflation ($\Gamma \geq 1$) or deflation factor ($\Gamma \leq 1$). It should be noted that in this study we use the model error covariance matrix in a static form, which means it is once defined in the initialization step and used for the entire assimilation cycles. Therefore, in our formulation Q_k becomes Q. Since in this study, a dual state–parameter estimation of hydrologic model is presented, it is necessary to assign a deterministic initial guess for the model parameter Θ. Using Latin hypercube sampling, the initial ensemble of parameter θ^i for the assimilation cycle τ is generated. Since 4DVAR cost function is executed deterministically, we use the ensemble mean of θ^i as follows:

$$\Theta = \frac{1}{N} \sum_{i=1}^{N} \theta^i \tag{31}$$

Step 1: The 4DVAR finds an initial condition by which the model forecasts best fit the observations within the assimilation interval. We specify the model parameters Θ at each time step within the assimilation interval. We then find the best initial state variables (also known as analysis) x_a by minimizing the 4DVAR cost function. For each of the perfect and imperfect model scenarios, we respectively use

strong- and weak-constraint 4DVAR methods in order to be consistent with their inherent definitions. Figure 1 demonstrates weak-constraint 4DVAR cost function, but as mentioned earlier, it can be simplified to strong-constraint 4DVAR as well. In either formulation, the best initial condition (analysis) x_a can be computed through the minimization of the cost function. Therefore, we use cost functions (32) and (33) for real and synthetic case studies, respectively. To run both weak-constraint and strong-constraint 4DVAR, forecast state variables within the assimilation interval are required. This can be achieved by the model forward run.

$$
\begin{aligned}
J(x_0, \; ..., \; x_K) &= J^b + J^o + J^q \\
&= \frac{1}{2}(x_0 - x_{0,b})^T B^{-1}(x_0 - x_{0,b}) + \frac{1}{2}\sum_{k=0}^{K}(y_k - h_k(x_k))^T R_k^{-1}(y_k - h_k(x_k)) \\
&\quad + \frac{1}{2}\sum_{k=1}^{K}(x_k - M_{k-1 \to k}(x_{k-1}, \Theta, \; u_k))^T Q^{-1}(x_k - M_{k-1 \to k}(x_{k-1}, \Theta, u_k))
\end{aligned}
\tag{32}
$$

$$
\begin{aligned}
J(x_0) &= J^b + J^o \\
&= \frac{1}{2}(x_0 - x_{0,b})^T B^{-1}(x_0 - x_{0,b}) \\
&\quad + \frac{1}{2}\sum_{k=0}^{K}(y_k - h_k(x_k))^T R_k^{-1}(y_k - h_k(x_k))
\end{aligned}
\tag{33}
$$

As briefly discussed before, variational data assimilation methods require linearization of observation h and model $\mathcal{M}$ operators. This significantly hampers their utilization in hydrological applications where such linearizations are not often achievable. To circumvent this problem, we use a derivative-free optimization method to minimize the 4DVAR cost function. Derivative-free optimization methods are those in mathematical optimization discipline that do not use derivatives and finite differences in the classical sense to find optimal solutions (Rios & Sahinidis, 2013). In this paper, we use Nelder-Mead algorithm to find the optimal solution x_a of the 4DVAR cost functions. This method requires no derivatives to be computed and the objective function does not need to be smooth. Nelder-Mead algorithm is very fast, simple, and effective for the problems with small dimensions (Gao & Han, 2012). For higher-dimensional problems, there are other techniques, such as Metaheuristic Algorithms (MAs); however, their implementations are computationally intensive and, therefore, their use may require high-performance computing approach. In this study, we used *fminsearch* in MATLAB to use Nelder-Mead algorithm and find the optimal solution of the 4DVAR cost function. The algorithm first makes a simplex around the initial guess ($x_{0, b}$) and then updates the simplex repeatedly to minimize the objective functions (equations (32) and (33)). The ideal scenario for implementing the proposed HEAVEN approach is to use forecast model adjoint for the minimization of 4DVAR cost function. This, however, may not be practical in hydrologic applications, and therefore we employed the Nelder-Mead algorithm, which was sufficient for the model used in this study. If the model adjoint is not achievable, we encourage the readers to select the most appropriate optimization technique depending on their used model.

Step 2: In this step, we generate initial state ensemble x_0^i for particle filtering from the deterministic solution of 4DVAR approach x_a. This can be done by equation (34). It should be reminded that the prior error covariance matrix B is the one used in the 4DVAR cost function in step 1. It should be mentioned that here we used background error covariance matrix to generate the ensemble around the 4DVAR solution (x_a) knowing that the analysis error covariance matrix is not available.

$$
x_0^i = x_a + \varepsilon^i \qquad \varepsilon^i \sim N(0, B)
\tag{34}
$$

Step 3: To ensure that an appropriate initial condition x_0^i is replicated for cycle τ, which later results in better estimation of the posterior distributions in that window interval, we run the forward model for cycle τ using two initial ensemble scenarios: (1) x_0^i and (2) state posterior distribution obtained in the last time step ($k = K$) of assimilation cycle $\tau - 1$ (x_K^i) (particles depicted with the green glow color in Figure 1). Under these two initial conditions, we calculate y_k^i for ensemble members within the assimilation interval $[t_0, t_K]$, and based on their discrepancies from the observations Obs_k, one can

decide to preserve the particles x_0^i or replace them with those already available from the previous cycle $\tau - 1$. Here, we present a pseudo-code to further clarify the implementation of step 3 in the HEAVEN framework.

Step 3

$x_0^i = x_a + \varepsilon^i$ and $x_0^i = x_K^i$ ▷ initialization

for $i = 1 : N$ **do**   ▷ for each ensemble member

    for $k = 1 : K$ **do**   ▷ within the assimilation interval

        $x_k^i = \mathcal{M}_{k-1 \to k}\left(x_{k-1}^i, u_k, \theta^i\right) + \omega_k$ $\omega_k \sim N(0, Q_k)$ ▷ run model

        $y_k^i = h_k\left(x_k^i\right) + v_k$ $v_k \sim N(0, R_k)$

        $E_k^i = \left|y_k^i - Obs_k\right|$ ▷ absolute error

    end for

    $e^i = \sum_{k=1}^K \frac{E_k^i}{\sqrt{R_k}}$ ▷ discrepancy of each ensemble

end for

$\left\{e_1^i\right\}$ ▷ initialized with $x_a + \varepsilon^i$

$\left\{e_2^i\right\}$ ▷ initialized with x_K^i

for $i = 1 : N$ **do**   ▷ for each ensemble member

    if $e_1^i < e_2^i$ **then**

        $x_0^{i,p} = x_a^i$

    else

        $x_0^{i,p} = x_K^i$

    end if

end for

In this step, we run the EPFM filter with the initial proposal states $x_0^{i,p}$ and ensemble members θ^i to create state and parameter posterior distributions in the window interval $[t_0, t_K]$ and assimilation cycle τ. Then, we calculate the mean of state posterior distribution at $k = K$ in order to initialize the 4DVAR approach for the next assimilation cycle $\tau + 1$. This means that the prior knowledge for 4DVAR cost function at cycle τ is derived from the information of EPFM posterior at the last time step of cycle $\tau - 1$. This indicates how 4DVAR and EPFM feed each other from cycle to cycle. The parameter x_k^{i+} is the state posterior distribution obtained by EPFM at time k ($1 \le k \le K$).

$$x_{0,b} = \frac{1}{N} \sum_{i=1}^N x_{k=K}^{i+} \tag{35}$$

Step 4: In this paper, EPFM as an ensemble-based sequential data assimilation method is combined with a 4DVAR approach to capture the model structural uncertainty and dynamically update the prior error covariance matrix B. B incorporates two components, B_d and B_s. The dynamic error covariance matrix B_d is calculated using the ensemble estimates of model error within the assimilation cycle τ, while static error covariance matrix B_s is obtained from the previous assimilation cycle $\tau - 1$. The best estimates of the state variables and parameters are obtained as the expected values from their posterior distributions at each time within the assimilation window. θ_k^{i+} is the parameter posterior distribution obtained by the EPFM at time k ($1 \le k \le K$).

$$\overline{x}_k^+ = \frac{1}{N} \sum_{i=1}^N x_k^{i+} \text{and} \ \overline{\theta}_k^+ = \frac{1}{N} \sum_{i=1}^N \theta_k^{i+} \quad \forall k = 1, \ \ldots \ , K \tag{36}$$

$$\eta_k = \overline{x}_k^+ - M_{k-1 \to k}\left(\overline{x}_{k-1}^+, \overline{\theta}_k^+, u_k\right) \tag{37}$$

η_k is the estimate of model error at each time within the assimilation window.

$$q = \frac{1}{K} \sum_{k=1}^{K} \eta_k \qquad (38)$$

q is the model error bias within the assimilation window.

$$B_d = \frac{1}{K-1} \sum_{k=1}^{K} [\eta_k - q][\eta_k - q]^T \qquad (39)$$

B_d dynamic prior error covariance matrix in the assimilation cycle τ.

$$B_s = B \qquad (40)$$

B is the prior error covariance matrix from the previous assimilation cycle $\tau - 1$. B_s is the static prior error covariance matrix at assimilation cycle τ.

Step 5: In this step, we update prior error covariance matrix B using the following formula. The parameter γ is a tuning factor.

$$B = (\gamma \times B_s) + (1-\gamma) \times B_d \qquad 0 \leq \gamma \leq 1 \qquad (41)$$

In this formalism, if γ is zero, the prior error covariance B is updated only using the model error calculated within the EPFM framework at assimilation cycle τ. If $\gamma = 1$, the B matrix is the one used in the previous assimilation cycle $\tau - 1$ within the 4DVAR cost function, meaning that no update is made from cycle $\tau - 1$ to τ. The updated matrix B in the assimilation cycle τ will be used as prior error covariance for the 4DVAR cost function in the next assimilation cycle $\tau + 1$. This completes one assimilation cycle of the HEAVEN (Figure 1). As shown in Algorithm 1, it is reminded that the model evolution errors (η_k) are used to update only the B matrix, while the Q matrix remains unchanged throughout the assimilation cycles. It is also important to note that in higher dimensional problems, regularization of B_d, such as localization, might be necessary (see Bannister, 2017).

Algorithm 1: HEAVEN

$\{x_{0,\,b}, \Theta, B, Q, R_k\}$ ▷ initialization (equations (28)–(31)).

for $\tau = 1 : \frac{T}{K}$ **do**   ▷ assimilation cycle

  $t_{k=0} = t_{(\tau \times K)} - K \qquad t_{k=K} = t_{(\tau \times K)}.$

  $[t_{k=0}, t_{k=K}] \qquad \forall k = 0, 1, 2, \dots, K$ ▷ assimilation interval

  Step 1

  if Synthetic case **then**

    $x_a = \text{Minimize } (J^b + J^o)$ ▷ equation (33)

  end if

  if Real case **then**

    $x_a = \text{Minimize } (J^b + J^o + J^q)$ ▷ equation (32)

  end if

  Step 2

  x_0^i ▷ equation (34)

  Step 3

  $\text{x}_0^{i,\text{p}}$

  Step 4

  Run EPFM filter $\left\{ x_0^{i,p}, \theta^i \right\}$

  $\{x_k^{i+}, \theta_k^{i+}\}$ ▷ posterior distributions

  $\{x_{0,\,b}\}$ ▷ equations (35) and (36)

  Step 5

  $\{B_d\}$ and $B_s = B$ ▷ equations (37)–(39)

Step 6
Update B matrix   ▷ equation (41)
Create θ^i using Latin hypercube sampling
$\Theta = \frac{1}{N}\sum_{i=1}^{N}\theta^i$
$\{x_{0,\,b}, \Theta, B, Q, R_k\}$ ▷ available for the next cycle
end for

3. Experimental Design

In this study, a synthetic and several real data experiments are performed to evaluate the effectiveness, usefulness, and robustness of HEAVEN in estimating the states and parameters with better characterization of their posteriors. All the experiments are conducted with the HyMOD and Sacramento Soil Moisture Accounting (SAC-SMA) models.

3.1. HyMOD Model

The HyMOD model (Boyle, 2001) is a conceptual and parsimonious-lumped hydrologic model containing five parameters and five state variables. Using these parameters, the model moves the water through a series of quick-flow and slow-flow tanks, to route the runoff to the outlet. For more information about this model, we refer the readers to Moradkhani, Sorooshian, et al. (2005).

3.2. SAC-SMA Model

The SAC-SMA model, initially introduced by Burnash et al. (1973), is a spatially lumped continuous soil moisture model used operationally at the National Weather Service River Forecast System to generate daily streamflow from mean areal precipitation and daily potential evapotranspiration data. In this model, each basin is represented vertically by two soil zones: an upper zone, which models the short-term moisture storage, and a lower zone, which accounts for long-term groundwater storage. In SAC-SMA model, the precipitation over the impervious area contributes directly to generating direct runoff. The model has 14 parameters and 9 state variables. The SAC-SMA is the model used by the NWS for operational river and flash flood forecasting throughout the United States.

3.3. Case Studies

3.3.1. Synthetic Case Study

In this paper, we use Leaf River basin, located in southern Mississippi (MS), for a synthetic analysis. Leaf River basin with an area of 1,944 km^2 is the main tributary of the Pascagoula River, which drains into the Gulf of Mexico. The prevailing climate of this region is humid subtropical with mild winters and dry summers and well distributed precipitation throughout the year. Data for the synthetic case study was collected from the NWS Hydrology Laboratory. This basin has been extensively used as a synthetic case study in several hydrologic data assimilation studies (e.g., Bulygina & Gupta, 2011; DeChant & Moradkhani, 2012; Parrish et al., 2012; Vrugt et al., 2006).

3.3.2. Real Case Studies

For the real data experiments, we use seven basins located in different climate regimes and geographical conditions in the western United States (Figure 2) to better examine the robustness and performance of the proposed method. These river basins are among those test basins included in the Model Parameter Estimation Experiment (MOPEX) project (Duan et al., 2006) where the water management effects can be ignored. In the following, we briefly describe these watersheds and their hydroclimate characteristics.

1. The Chehalis River basin (area = 2,318 km^2) is one of the largest basins in Washington dominated mostly by oceanic climate. This region is generally wet throughout the year, except the summer, which is relatively dry. With more than 1,400-mm precipitation annually, this watershed is the wettest area in the state. This basin is mainly fed by rainfall precipitation, although a small portion of mountainous regions accumulates snow during the winter. (2) The Indian Creek basin (area = 1,914 km^2) lies in the Klamath National Forest in California, which drains into the Klamath River. This basin, with mean annual precipitation of 762 mm, is characterized by semiarid to subhumid climate conditions. (3) The Carson River basin (area = 3,372 km^2) is an endorheic basin, which starts at the Alpine County in California and ends at the Nevada state. This region is mostly dominated by arid and hot climate with varying precipitation throughout the year. The annual precipitation in this basin
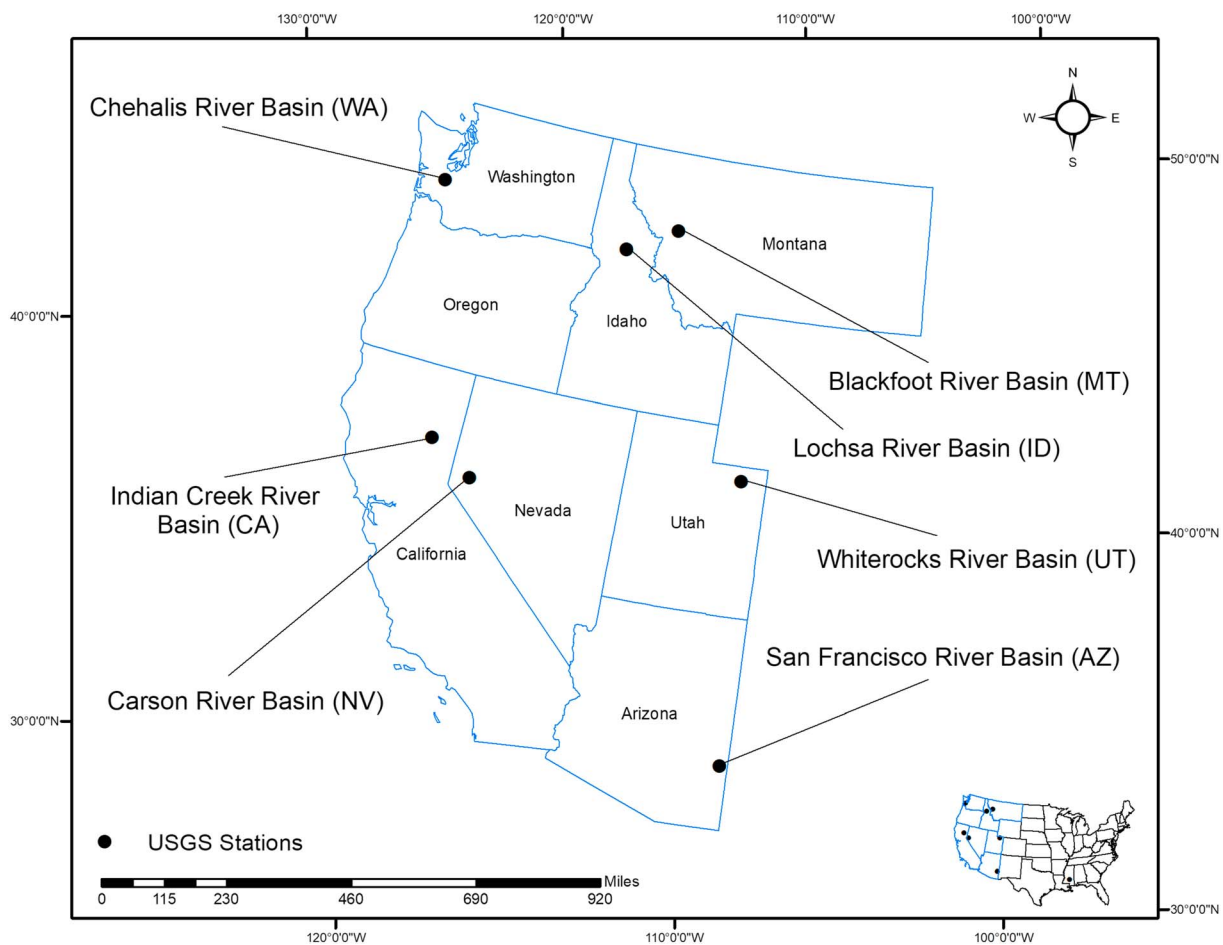
Figure 2. Seven study basins located in the western United States.

reaches to 127 mm, while annual evaporation exceeds 1,524 mm. Therefore, ground water becomes the primary source for municipal and industrial water use in this watershed. (4) The Blackfoot River basin (area = 5,923 km^2), the snow-fed and spring-fed river, lies in west central Montana. This basin encompasses a diverse range of ecosystems from high-elevation glaciated peaks, montane forests, and foothills to semiarid prairie pothole regions. A dry and cold climate dominates the watershed, with an average annual rainfall of 430 mm and an average annual snowfall of 2,000 mm. (5) The Lochsa River basin (area = 3,051 km^2) is located in the mountains of north central Idaho. This basin is in the temperate climate zone where the precipitation generally occurs as rain in the summer and snow in the winter. (6) The Whiterocks River basin (area = 282 km^2) is located in the semiarid region of northeastern Utah. In the mountainous regions of the watershed, more often precipitation falls as rain from April through September and as snow in winter months. This basin typically experiences hot, dry summers and cold winters with average annual precipitation of around 250 mm. (7) The San Francisco River basin (area = 7,163 km^2) is the largest tributary to the upper Gila River, which is located in southeastern Arizona and southwestern New Mexico. This basin is characterized by semiarid to humid climatic conditions with extreme seasonal precipitation events. The location of these watersheds are depicted in Figure 2.

3.4. Performance Measures

In order to provide a robust analysis of each assimilation run, it is necessary to assess the model performance through multiple deterministic and probabilistic measures. These are described below.

$$\text{Nash-Sutcliffe efficiency (NSE)} = 1 - \frac{\sum_{t=1}^{T}\left(y_t - y_t'\right)^2}{\sum_{t=1}^{T}\left(y_t - \overline{y_t}\right)^2} \tag{42}$$

$$\text{Reliability} = 1 - \frac{2}{T}\sum_{t=1}^{T}\left|\frac{z_t}{T} - U_t\right| \tag{43}$$

$$95\% \text{ Exceedance Ratio (ER95)} = \frac{1}{T}\sum_{t=1}^{T}\left(y'_{97.5\%,t} < y_t \text{ or } y'_{2.5\%,t} > y_t\right) \times 100\% \tag{44}$$

$$\text{Kling-Gupta Efficiency (KGE)} = 1 - \sqrt{\left(\left(\frac{\text{Cov}_{y_t,y_t'}}{\sigma\sigma'}\right) - 1\right)^2 + \left(\left(\frac{\sigma'}{\sigma}\right) - 1\right)^2 + \left(\left(\frac{\mu'}{\mu}\right) - 1\right)^2} \tag{45}$$

$$\text{Mean Absolute Bias (MAB)} = \frac{1}{T}\sum_{t=1}^{T}\left|y_t - y_t'\right| \tag{46}$$

$$\text{Correlation Coefficient} = \frac{\text{Cov}_{y_t,y_t'}}{\sigma\sigma'} \tag{47}$$

$$\text{centered Root Mean Square Difference (RMSD)} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(\left(y_t' - \overline{y}_t'\right) - \left(y_t - \overline{y}_t\right)\right)^2} \tag{48}$$

Nash-Sutcliffe efficiency (NSE; Nash & Sutcliffe, 1970) is widely used to assess the goodness of fit of hydrologic models. NSE varies between 0 (worse prediction skill) to 1 (perfect prediction skill). Parameters y_t and y_t' are the observed and simulated streamflow values, respectively. Reliability (Renard et al., 2010) is a measure of the fit of the Q-Q quantile plot to uniform. Similar to NSE, this measure also varies between 0 (farthest possibility from uniform) to 1 (exactly uniform). For the description of the z_t and U_t calculation, we refer the readers to Renard et al. (2010). The 95% Exceedance Ratio (ER95; Moradkhani et al., 2006) is an indicator to evaluate the spread of the ensemble. A perfect ensemble has a 5% exceedance of the 95% predictive bounds. A value greater than 5% (too narrow predictive distribution) indicates an overconfident forecast, while a value less than 5% (too wide predictive distribution) indicates an underconfident forecast. Similar to NSE, the Kling-Gupta Efficiency (KGE; Gupta et al., 2009) varies from $-\infty$ to 1, such that a value of 1 indicates a perfect fit between observed and simulated streamflow values. The pairs of (μ, σ) and (μ', σ') represent the first two statistical moments (means and standard deviations) of y_t and y_t', respectively. Mean absolute bias (MAB) is a metric that indicates the magnitude of the bias for a given estimate. A perfect prediction has a MAB value of 0 indicating no bias between the observed and simulated streamflow data. The correlation coefficient indicates the strength and direction of a linear relationship between two variables. The centered root mean square difference (RMSD) is similar to the RMSD, but both observed and simulated data are centered before the differences are calculated. This performance measure has a range between 0 (perfect analogs) and ∞ (total dissimilarity). RMSD along with correlation coefficient provide complementary statistical information of pattern similarity between two series (Taylor, 2001).

3.5. Data Assimilation Setting

This section describes the data assimilation setting for both synthetic and real data experiments. In this study, for all data assimilation runs, we assume a lognormal and a normal error distribution with a relative error of 25% for precipitation and potential evapotranspiration, respectively. Therefore, using these values, we assume that all meteorological observation errors, including sensor errors and those associated with spatial heterogeneity are taken into account. In addition, streamflow observation errors are assumed to be normally distributed with a 15% relative error. For real case studies, the model error is assumed to follow a normal distribution with a relative error of 25%.

4. Results and Discussions

In this section, first we examine the applicability and usefulness of the proposed method on the synthetic case (section 4.1), and later several real case studies (section 4.2) are discussed to indicate the effectiveness of this method in streamflow prediction and uncertainty quantification.

4.1. Streamflow Prediction Results (I): Synthetic Case Study

The first task in synthetic case study is to define two parameters of K and γ at which the proposed data assimilation approach operates. K is the assimilation window size, and γ specifies the portion of dynamic (B_d) and static (B_s) prior error covariances that participate in evolving prior error covariance (B) in each cycle throughout the assimilation period. The performance of HEAVEN can be optimized by tuning both K and γ parameters. By running the HEAVEN model for different K and γ, we obtained the most appropriate values of these two parameters. Here, we considered that K changes between 1 and 30 days and knowing that γ varies from 0 to 1, we considered a step size of $\Delta\gamma = 0.01$ for fine tuning of this parameter. For $\gamma = 0$, only dynamic error covariance is propagated throughout the assimilation process, while for $\gamma = 1$, the initial guess for B matrix ($B = B_s$) is linearly augmented along the assimilation cycles.

In addition to the SAC-SMA model, we used the HyMOD model, a more parsimonious hydrologic model, to further investigate the influence of these two parameters on the model performance. The tuning factor γ versus model performance for both SAC-SMA and HyMOD hydrologic models showed that the covariance matrix corresponding to $\gamma = 0.75$ and $k = 7$ days provides better results in all performance measures.

Further analysis revealed that the model performance is more dependent on the assimilation window size than the parameter γ. Thus, the assimilation window size should be appropriately chosen, as a small window size (~2 days) may unnecessarily increase the computational time and decrease the model performance, while a large window size (>20 days), irrespective of the potential for improving the model performance, may lead to local optima of the cost function. Here, a range of k values under different model scenarios, that is, using different ensemble sizes (50, 100, 200, and 500) and γ values were examined and concluded that for both hydrologic models a time interval of 7 days results in more reliable and accurate streamflow prediction. Also, we investigated the evolution of prior error covariance matrix B throughout the entire assimilation period, and detected that it directly correlates with the streamflow observation which enables the HEAVEN to properly capture the high and low flows over 4 years of analysis in the Leaf River basin. In general, when streamflow rises, the covariance matrix inflates, and when streamflow declines, the covariance matrix shrinks. For clarity, we illustrate this evidence for 1 year in Figure 3. Both deterministic and probabilistic measures shown in this figure indicates the superiority of HEAVEN in streamflow prediction. Note that this figure is reported based on an ensemble size of 50, although improved results are achievable by using larger ensembles.

In the synthetic experiment, we initialize B ($B=B_s$) with a zero matrix. Since this is a singular matrix at time zero, we initialize the B matrix with a very small value. This is apparent from the top panel of Figure 3, which demonstrates almost zero variance for all state variables in the first assimilation window (from $t = 0$ to $t = 7$). For different ensemble sizes, the HEAVEN outperforms the EPFM according to both deterministic and probabilistic measures. For instance, for the ensemble size of 50, the reliability increases from 0.71 for the EPFM to 0.83 for the HEAVEN. To analyze the ensemble spread, the ER95 is calculated. The ER95 of HEAVEN at 3.6% is closer to the optimal value (5%), indicating an accurate uncertainty characterization. Deterministic measures also indicated that the proposed HEAVEN (KGE = 0.91 and MAB = 8.18 m^3/s) produces an ensemble mean very close to the observation than the EPFM (KGE = 0.78 and MAB = 11.09 m^3/s). Although the proposed method requires additional steps as compared to the EPFM, both filters show almost the same computational demand for a given ensemble size. It is worth noting that the performance of HEAVEN even using much smaller ensemble sizes (e.g., 50) is similar to that of EPFM with the ensemble size of 200, meaning that HEAVEN with less computational demand is able to provide the same level of accuracy and precision as EPFM. This can be attributed to the derivative-free optimization algorithm used in the HEAVEN, which operates fast and does not increase the computing complexity. The results also reveal that the proposed filter does not suffer from the particle degeneracy and sample impoverishment problems even when a small ensemble size is used. In fact, the GA combined with MCMC in the EPFM (Abbaszadeh et al., 2018) considerably alleviates the weight degeneracy and sample impoverishment problems. Also, the formulation used in regenerating the replicates of state variables and parameters from one cycle to another throughout the assimilation period helps reduce the ensemble size without compromising the accuracy during the assimilation.
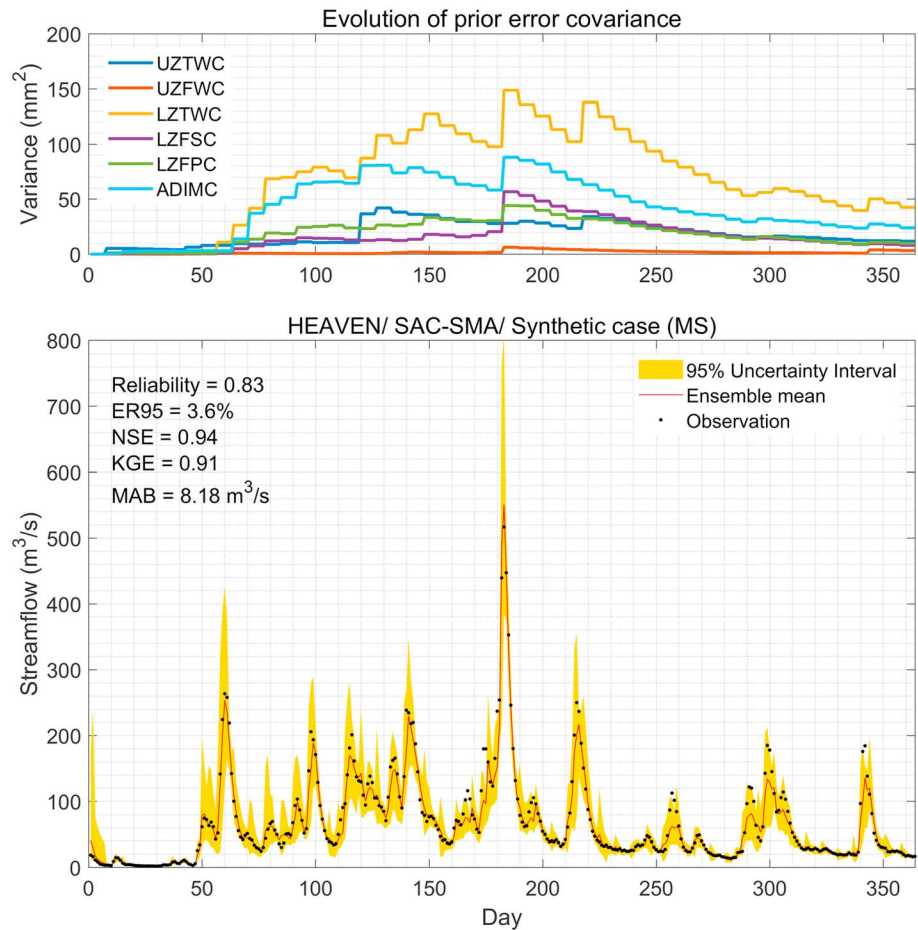
Figure 3. (top) Evolution of prior error covariance matrix (*B*) for six state variables in SAC-SMA model. (bottom) Synthetic streamflow prediction by the HEAVEN for SAC-SMA experiment over 1 year based on an ensemble size of 50. SAC-SMA = Sacramento Soil Moisture Accounting model; HEAVEN = Hybrid Ensemble And Variational data assimilation method for ENvironmental systems; ER95 = 95% Exceedance Ratio; NSE = Nash-Sutcliffe efficiency; KGE = Kling-Gupta Efficiency; MAB = Mean Absolute Bias.

4.2. Streamflow Prediction Results (II): Real Case Studies

In addition to the synthetic study, several real data experiments are conducted in the basins located in different environmental settings (i.e., climate zones and hydrological properties) to fully examine the effectiveness and usefulness of the proposed method. One of these experiments is based on the real data for the same basin used in the synthetic case. Figure 4 demonstrates streamflow prediction along with its uncertainty estimate for the Leaf River basin using the EPFM and HEAVEN over 1 year. This result is based on an ensemble size of 100. Both deterministic (KGE, NSE, and MAB) and probabilistic (Reliability and ER95) measures are summarized in this figure, and they all indicate that the HEAVEN provides more accurate and reliable predictions than the EPFM approach. For example, the HEAVEN improves the EPFM performance by 12% and 36% in terms of Reliability and MAB, respectively. Further analysis revealed that although the EPFM performs well, it remains slightly biased for low flows as compared to the HEAVEN unless a large ensemble size is employed. As seen in Figure 4, the EPFM shows a relatively low performance in predicting low flows during dry periods. These results are further supported in the right panel of Figure 4, where the *S*-shape predictive QQ plots show that the HEAVEN is better able to characterize high and low flows compared to the EPFM. Also, we noticed a large contrast between these two filters when data assimilation is performed on the real data set. This explains the importance of accounting for model error using the HEAVEN method. Unlike the synthetic case where *B* and *R* are the only error covariance matrices, here in the real case data assimilation scenario, the model error covariance matrix *Q* plays a crucial role. Conventionally, in the
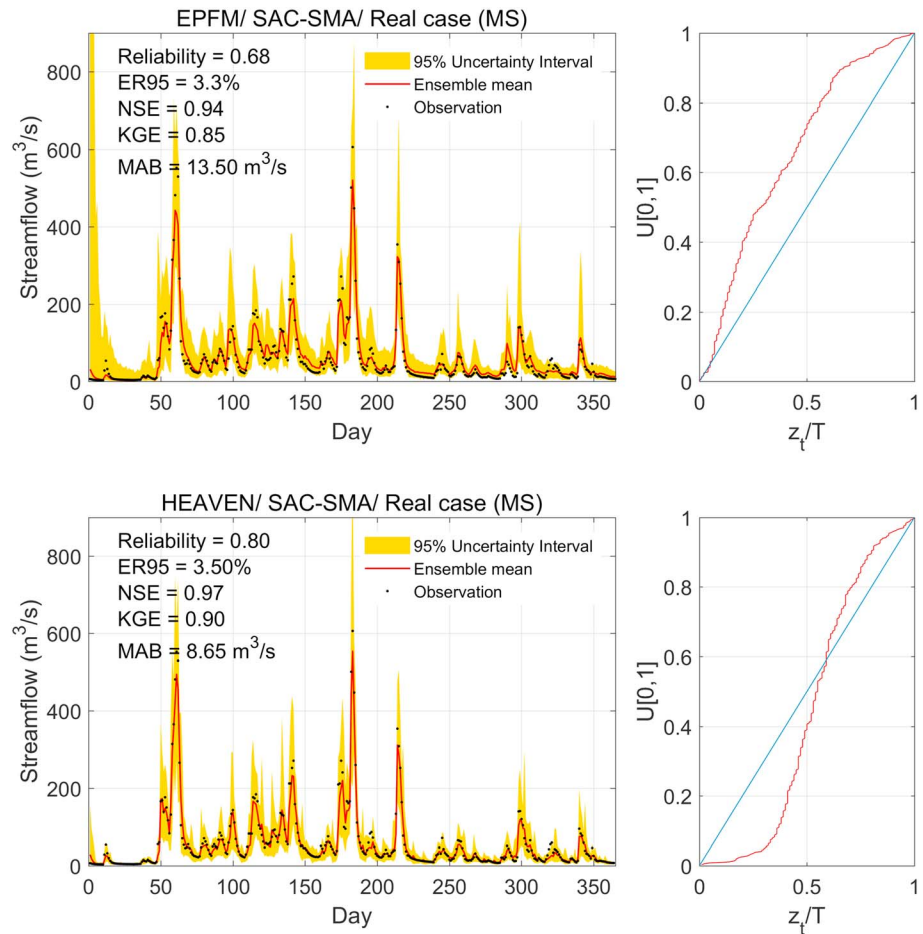
Figure 4. Streamflow prediction using the EPFM and HEAVEN for the Leaf River basin over 1 year (left panels). Predictive *QQ* plots (right panels). EPFM = evolutionary particle filter with MCMC; SAC-SMA = Sacramento Soil Moisture Accounting model; HEAVEN = Hybrid Ensemble And Variational data assimilation method for ENvironmental systems; ER95 = 95% Exceedance Ratio; NSE = Nash-Sutcliffe efficiency; KGE = Kling-Gupta Efficiency; MAB = Mean Absolute Bias.

hydrologic data assimilation, one may account for model structural error by simple perturbation of the model predictions y_t. This is because the model ($\mathcal{M}$) and observation (h) operators are not explicitly distinguishable in hydrologic systems. Therefore, matrix Q is not basically formed and fully taken into account within the EPFM filter.

The proposed HEAVEN filter is able to fully account for model structural error, as it uses weak-constraint 4DVAR approach to explicitly characterize the Q matrix, and consequently provides a more accurate and reliable posterior distribution. It should be noted that in this study the model error covariance is neither inflated nor deflated ($\lambda = 1$). However, depending on the case study, tuning of this parameter may improve the assimilation results. Unlike the synthetic case in which the B matrix was initialized with zero matrix, in the real case study, this matrix is initialized according to equation (29). As seen in the top panel of Figure 3, although the prior error covariance matrix B automatically inflates or deflates over time, it is always greater than the initialized B_s matrix. The Q matrix is initialized similar to the B_s matrix (see equations (29) and (30)) and used in a static form throughout the assimilation period. This means that the magnitude of prior error covariance matrix B is always greater than the Q matrix, unless the Q matrix is inflated. The B and Q matrices collectively capture the model structural error and contribute to generating more reliable and accurate posterior distributions.

In summary, the proposed algorithm, regardless of the ensemble size, outperforms the original EPFM according to both deterministic and probabilistic measures. This finding indicates that integrating the
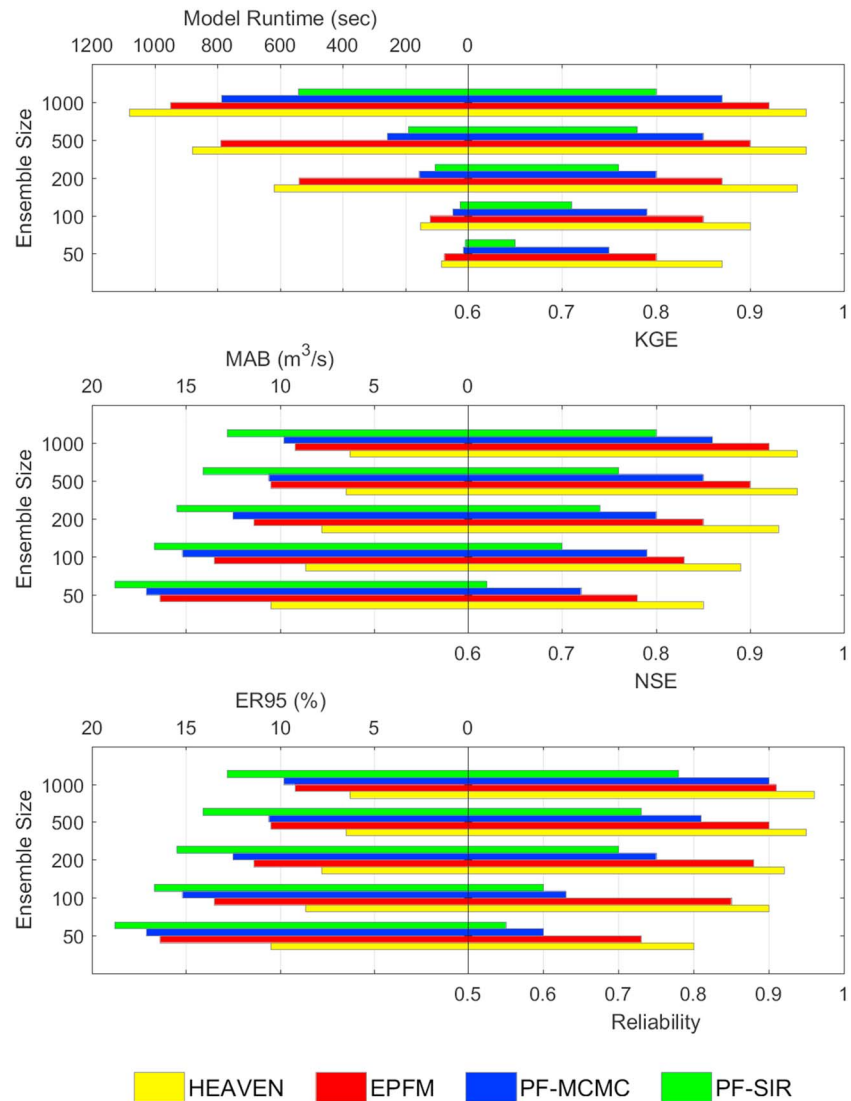
Figure 5. The comparison of accuracy, distribution spread, and computational demand for four assimilation strategies in the real case study over 4 years. EPFM = evolutionary particle filter with MCMC; HEAVEN = Hybrid Ensemble And Variational data assimilation method for ENvironmental systems; PF = particle filter; MCMC = Markov chain Monte Carlo; SIR = sampling importance resampling; ER95 = 95% Exceedance Ratio; NSE = Nash-Sutcliffe efficiency; KGE = Kling-Gupta Efficiency; MAB = Mean Absolute Bias.

sequential and variational data assimilation approaches (EPFM and 4DVAR) provides the possibility to fully account for all sources of uncertainties involved in the hydrologic predictions. It is important to mention that by applying the proposed framework on both hydrologic models, we were able to achieve better performance in forecasting. This was validated by comparing the current approach with the variants of the particle filter. Figure S1 (in supporting information) shows the PF-MCMC and HEAVEN skills for 3-day streamflow forecast in Leaf River basin in Mississippi, during the flood season. To reduce the possibility of sample impoverishment in the particle filter, Moradkhani et al. (2012) combined the PF with MCMC algorithm. In addition, in Figure 5, we summarized multiple deterministic (i.e., KGE, NSE, and MAB) and probabilistic (i.e., ER95 and Reliability) measures along with model runtime for four different assimilation strategies, that is, HEAVEN, EPFM, PF-MCMC, and PF-SIR. This comparison was made for the real case study, Leaf River basin, over 4 years. This figure indicates that even using small ensemble size (e.g., 50), HEAVEN outperforms the benchmark approach, PF-SIR, with much larger ensemble size (e.g., 1,000). This means that the HEAVEN approach with a smaller ensemble size has less model runtime and better model
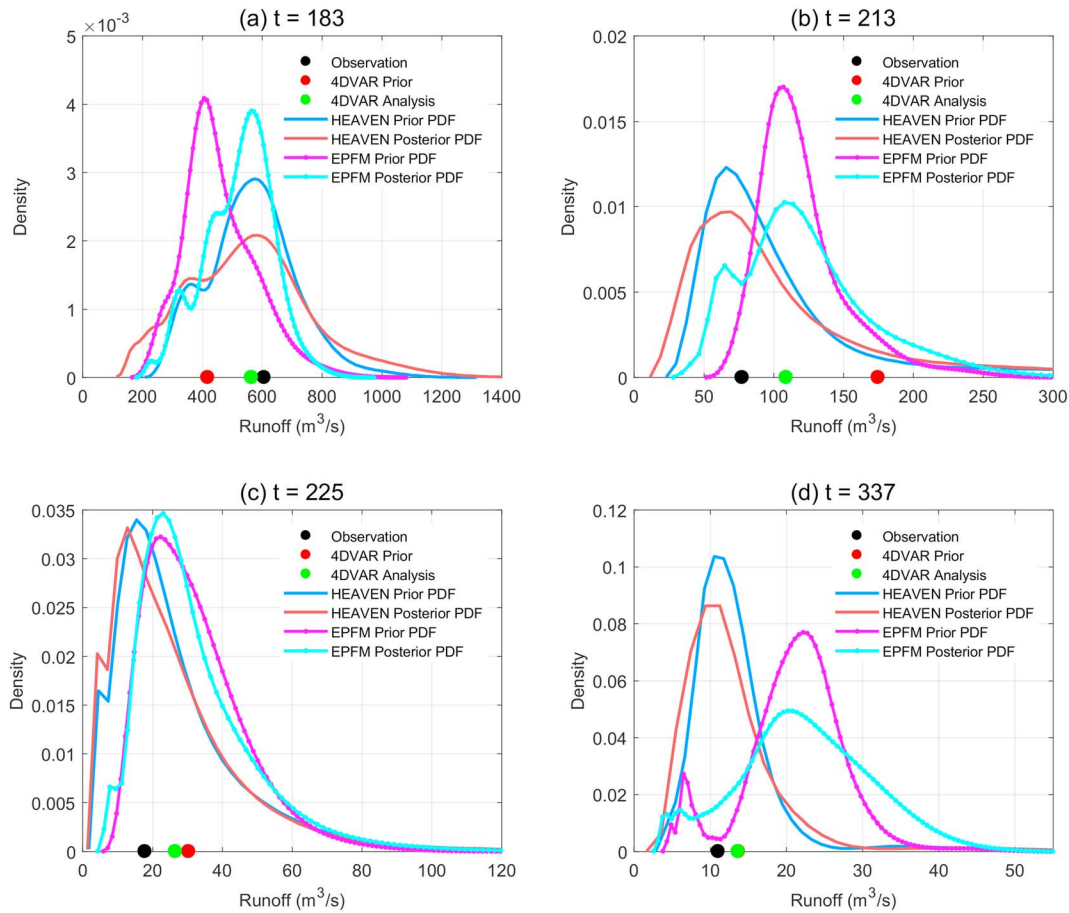
Figure 6. The prior and posterior distributions obtained by the EPFM and HEAVEN for 4 days (a. 183, b. 213, c. 225, and d. 337), where each represents the initial time of an assimilation window.

performance than the PF-SIR at a larger ensemble size. This further highlights the superiority of the proposed HEAVEN approach to its simpler versions (PF-SIR, PF-MCMC, and EPFM).

Up to this point, we compared the efficacy of HEAVEN and EPFM approaches over the Leaf River basin for both synthetic and real case scenarios. To further interpret how 4DVAR cost function provides a more reliable and accurate initial condition for particle filtering in each assimilation cycle, we examine the prior and posterior distributions at four daily time steps ($t = 183, 213, 225,$ and 337 days) for the Leaf River basin using the HEAVEN method. These 4 days represent the initial time of four different assimilation cycles, which were chosen according to different streamflow regimes. For example, Figure 6a ($t = 183$) displays a peak flow (with streamflow observation of 605.9 m^3/s), while Figure 6c ($t = 225$) indicates a low flow (with streamflow observation of 17.8 m^3/s). The black points in Figure 6 illustrate such extreme events. We recall that the results being discussed here for these 4 days were actually obtained from the real case data assimilation performed by HEAVEN, which was previously reported in Figure 4 for 1-year assimilation period. To interpret the results presented in Figure 6, we first explain the keywords included in the figure legend. In Figure 6a ($t = 183$), 4DVAR prior (shown with a red point) is a deterministic streamflow value based on a priori model state ($x_{0,b}$). The parameter $x_{0,b}$ is available from the mean of model state ensemble at time $t = 182$ (or the last time step of the previous assimilation cycle as seen in step 4 of Figure 1). The 4DVAR analysis (green point) is a deterministic streamflow value based on an optimal initial state (x_a). The x_a is obtained by minimization of the 4DVAR cost function (step 2 in Figure 1). HEAVEN prior PDF (shown with a blue curve) is a kernel probability distribution fitted to the streamflow values based on the simulations obtained from the prior ensemble model states ($x_0^{i,P}$, step 3 in Figure 1). HEAVEN posterior PDF (brown curve) is a kernel probability distribution fitted to the streamflow values based on the simulations obtained from the posterior ensemble
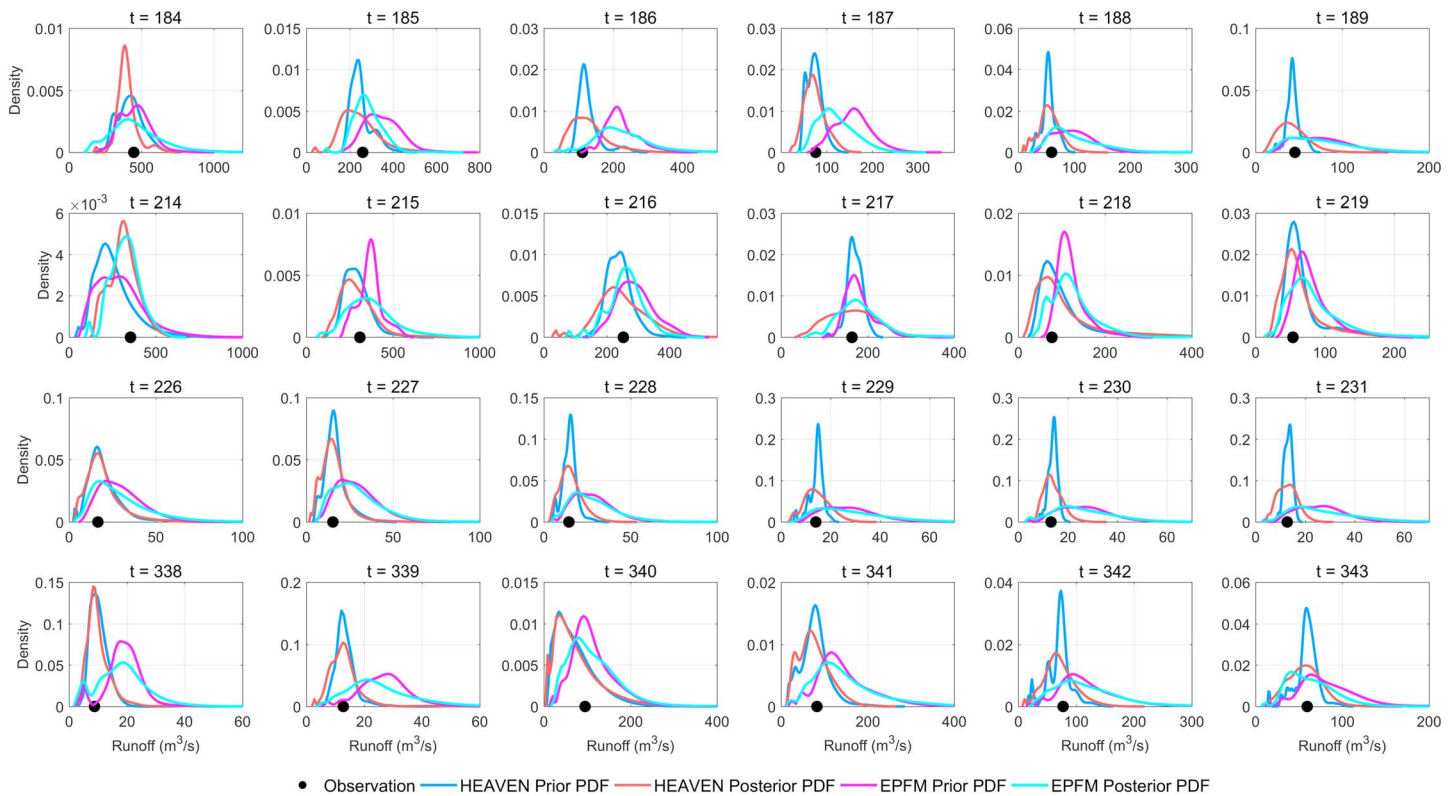
Figure 7. (first to fourth rows, from top to bottom) The prior and posterior distributions obtained by the EPFM and HEAVEN for the next 6 days in each assimilation window corresponding to the initial conditions reported in Figure 6. HEAVEN = Hybrid Ensemble And Variational data assimilation method for ENvironmental systems; EPFM = evolutionary particle filter with MCMC.

model states ($x_{k=0}^{i+}$). EPFM prior and posterior PDFs are also shown in magenta and cyan colors, respectively. To avoid confusion, it should be noted that all 4DVAR products in this study (such as those shown in Figure 6) are obtained from the variational portion of the HEAVEN model and not related to any single 4DVAR data assimilation run.

The HEAVEN approach seeks to find an optimal prior distribution at the beginning of each assimilation window by minimization of the 4DVAR cost function that measures the distance between prior state estimates and observations over the time interval $[t_{k=0}, t_{k=K}]$. For instance, in Figure 6a, $t_{k=0} = 183$ and $t_{k=K}=189$ (here, the assimilation interval K=7 days, as discussed before in section 4.1). As Figure 6 illustrates, the 4DVAR approach moves a deterministic background initial condition (shown with a red point) toward an optimal location (known as 4DVAR analysis and shown with a green point) close to the observation (shown with a black point). In Figure 6a, this improvement was made by 88% (from 412 to 583 m^3/s), showing a great success for this approach to capture the high flows. A similar interpretation can be drawn from Figure 6c, showing the applicability of this method in predicting low flows. Once the optimal deterministic initial state (x_a) is obtained, we perturb it by ε, where ε is assumed to follow normal distribution with mean zero and variance $B^{1/2}$. B has already been propagated through both static (B_s) and dynamic (B_d) prior error covariance matrices. The result of this step, x^i, undergoes an acceptance/rejection process in step 3 to generate an appropriate prior distribution ($x_0^{i,p}$). This results in prior streamflow distribution, shown with a blue curve in Figure 6. In step 4, we initialize the particle filtering (here EPFM) component of the HEAVEN with $x_0^{i,p}$ and obtain the posterior distributions for the time interval $[t_{k=0}, t_{k=K}]$. As an example, the posterior distribution at $t_{k=0} = 183$ is depicted with a brown curve in Figure 6a. The premise of this methodology is that the improved estimation of posterior distribution while accounting for model error within the HEAVEN results in more accurate and reliable streamflow predictions within the assimilation window. This is reported for all four time intervals in Figure 7 while compared with those obtained by the EPFM model when no account of model error exists. For instance, for the assimilation window initialized at
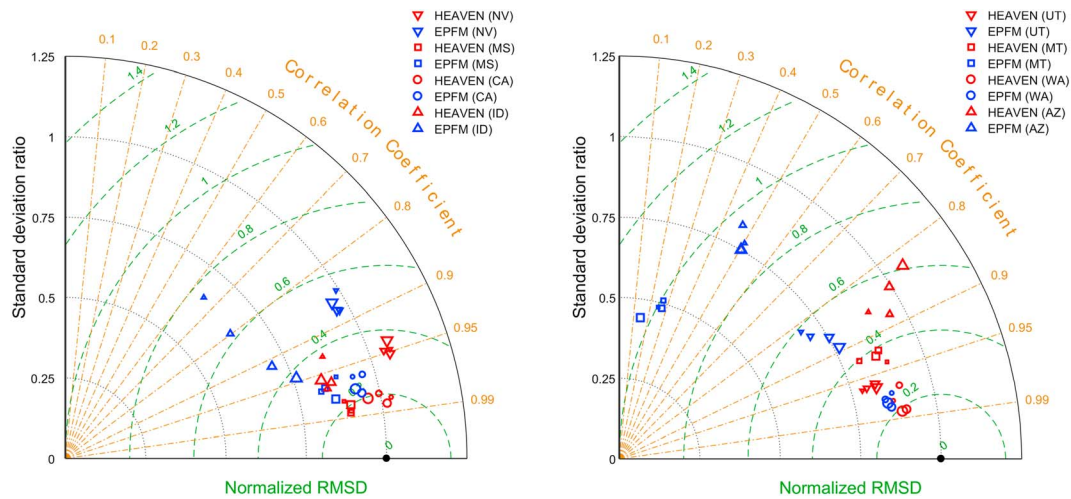
Figure 8. Taylor diagrams displaying the effectiveness of the two assimilation methods (i.e., EPFM and HEAVEN) for eight real case studies over 1 year of analysis. The symbols indicate different case studies in each panel. The size of each symbol represents the ensemble size (50, 100, 200, and 500) schematically. Normalized RMSD is represented by green dashed line, while correlation coefficient is displayed by brown dotted line. Normalized standard deviation and correlation coefficient are on the radial axis and angular axis, respectively. The observation is shown with a black point on the horizontal axis. HEAVEN = Hybrid Ensemble And Variational data assimilation method for ENvironmental systems; EPFM = evolutionary particle filter with MCMC; RMSD = root mean square difference.

$t_{k = 0} = 183$, the next 6 days show a flood recession period for streamflow, as it declines from 447.4 to 45.30 m^3/s. In this case, it is observed that the HEAVEN provides more accurate and reliable streamflow predictions due to better estimation of both prior and posterior distributions. As we discussed in our recent publication (Abbaszadeh et al., 2018), the EPFM is a robust evolutionary ensemble-based data assimilation approach to predict the streamflow owing to its unique features. However, the added feature in the HEAVEN that is mainly the characterization of model error make this new approach more suitable when model error is a serious concern. Further analysis reveals that the SAC-SMA model with improved initial condition from the HEAVEN at $t = 189$ is able to forecast streamflow up to 7 days with a better degree of accuracy and reliability. The second row in Figure 7 indicates that streamflow increased from 75.4 to 353.9 m^3/s in 1 day (between $t = 213$ and $t = 214$) followed by a slow decay for the next 5 days. The results demonstrate how both prior and posterior distributions characterized by the HEAVEN accurately and precisely track streamflow changes as compared to the EPFM. The third row in Figure 7 illustrates the capability of both filters in predicting low flows. It indicates that a small shift for both prior and posterior distributions at the beginning of the assimilation window (see Figure 6c) results in more accurate and reliable streamflow predictions in low flow events. Likewise, in the fourth row for Figure 7, the HEAVEN outperforms the EPFM, specifically on days $t = 338$ and $t = 339$, where the EPFM particles are located far from the observations.

Up to this point, using synthetic and real case experiments, we examined how HEAVEN is capable of properly characterizing the posterior distribution with higher accuracy and reliability for different streamflow regimes. The ensuing task is to further investigate the robustness and scalability of the proposed HEAVEN method compared to its original EPFM version. To accomplish this, we applied both filters on seven real case experiments introduced in section 3.3.2 and analyzed their performances over four different ensemble sizes of 50, 100, 200, and 500. In order to summarize the obtained results, we used the Taylor diagram that displays the comparative assessment of different techniques by using three performance measures, that is, normalized standard deviation, correlation coefficient, and normalized centered RMSD. Figure 8 presents two Taylor diagrams summarizing the statistics of the comparison between the ensemble mean, obtained by both HEAVEN and EPFM filters, and observations for seven real case experiments plus Leaf River basin used in this study.

The symbol below/above the standard deviation of 1 represents the ensemble mean wherein the variability is smaller/larger than that of the streamflow observation. The black point (observation) shown on the horizontal axis represents the normalized RMSD equal to 0, and both the ratio of standard deviations and the correlation coefficient are equal to 1. In all cases, it is seen that the ratios of standard deviation for the

HEAVEN ensemble mean are closer to 1 than the EPFM ensemble means. These diagrams also show that there is a good range of correlation obtained by the HEAVEN, with most values above 0.90. In addition to these two measures, the normalized centered RMSD indicates that the HEAVEN outperforms the EPFM regardless of the ensemble size. For all eight real case experiments, it is seen that the HEAVEN with a small ensemble size (e.g., 50) yields more accurate streamflow predictions as compared to the EPFM even with an ensemble size of 500. As shown in Figure 8, for two watersheds located in Arizona and Montana, there is a considerable difference between the performances of the two filters. A more plausible reason for this may be attributed to the hydroclimate conditions of these two regions. The Blackfoot River basin in Montana is a snow-dominated basin due to the strong atmospheric river in winter, and floods often happen due to peak spring melt. However, the San Francisco River basin in Arizona is a rain-fed basin with extreme rainfall events that are followed by dry periods. The results from the EPFM imply that this model does not properly capture the streamflow fluctuations resulting from either evapotranspiration (AZ) or snowmelt (MT). However, the HEAVEN better represents the streamflow variations and detects extreme events. Furthermore, the HEAVEN robustness can be explained by how it characterizes the model structural error by taking into account both B and Q error covariance matrices in the 4DVAR cost function. However, EPFM does not account for the model structural error. Therefore, the ensemble mean obtained by the HEAVEN shows higher skill than the EPFM for all real case experiments. We also compared the probabilistic performance measures (i.e., ER95 and Reliability) between these two filters and found that the HEAVEN always outperforms the EPFM regardless of the ensemble size. For example, for the Chehalis River basin (WA) with an ensemble size of 50, although both filters deterministically show similar performance in predicting streamflow, the probabilistic measures show that the HEAVEN provides more reliable posterior distribution than the EPFM. The ER95 of the HEAVEN (3.9%) is closer to the ideal value of 5% than the EPFM with ER95 = 10.21%. The higher ER95 value of the EPFM is because the ensemble distribution is too narrow, which indicates an overconfident streamflow prediction. The Reliability increases from 0.72 for the EPFM to 0.85 for the HEAVEN, indicating a more reliable ensemble prediction by the HEAVEN. Therefore, we conclude that the proposed HEAVEN approach, by characterizing the model structural uncertainty, along with the model parameter and input data uncertainties, provides a more accurate and reliable posterior distribution as compared with the EPFM method. The superiority of the HEAVEN is further confirmed when we assess the capability of both approaches in predicting high flow events. This corroborates with the findings of others (Pathiraja, Moradkhani, et al., 2018; Pathiraja, Anghileri, et al., 2018a, 2018b; Shoaib et al., 2016), who demonstrated that accounting for the model structural uncertainty is particularly important in predicting high flow events.

It should be noted that, although techniques to combine the hybrid particle filter and 4DVAR have been previously explored (Atkins & Morzfeld, 2013; Chorin et al., 2010; Chorin & Tu, 2009; Morzfeld et al., 2012; Slivinski et al., 2015; van Leeuwen, 2015; Zhu et al., 2016), the objective, formulation, and implementation of the proposed approach is fundamentally different from other methods. HEAVEN is a framework in which any variant of particle filtering can be used to account for and quantify the hydrologic model structural uncertainty, along with other sources of uncertainties involved in model predictions, without a need to calculate the adjoint and tangent linear of forecast model, which is most often not practical in hydrologic studies.

5. Concluding Remarks

This paper proposes a new hybrid ensemble and variational data assimilation method that effectively combines both sequential (EPFM) and variational (4DVAR) assimilation approaches to account for all sources of uncertainties involved in hydrologic predictions and thus leads to more accurate and reliable posterior distributions for both state variables and parameters in data assimilation applications. The effectiveness and usefulness of this technique was evaluated by both deterministic and probabilistic measures, and the robustness and superiority of this filter was examined through eight real case studies located in different geographical and climate zones across the United States. This study suggests using the HEAVEN approach for the following features:

1. It operates simultaneously on both batch processing and sequential manners, leading to a more complete estimation of posteriors for any streamflow regimes, including low and high flows.

2. It characterizes model structural uncertainty by incorporating an explicit form of model error covariance matrix (Q) in the 4DVAR cost function.

3. It propagates the prior error covariance matrix (B), which consists of a linear combination of static (B_s) and dynamic (B_d) error covariance matrices, from one cycle to another cycle over the entire assimilation period to fully account for a wide range of uncertainties in model predictions, and thus lead to more accurate and reliable posterior distributions.

4. It precludes the particle degeneracy and sample impoverishment.

In this study, we used a lumped hydrologic model as a proof of concept for the proposed joint Bayesian and variational data assimilation approach, although implementation and analysis with a distributed hydrologic model would be the next step to further analyze the improved performance and investigate the HEAVEN scalability. These aspects will be included in our upcoming research paper. One attractive feature of HEAVEN is that it needs neither tangent linear nor adjoint versions of forecast model, making it more suitable in hydrologic applications. However, for those models in which such linearization of model and observation operators are accessible, it is expected that the results and the computational efficiency of the implementation may be further improved. This conjecture will be investigated in future research.

References

Abbaszadeh, P., Moradkhani, H., & Yan, H. (2018). Enhancing hydrologic data assimilation by evolutionary particle filter and Markov chain Monte Carlo. *Advances in Water Resources*, *111*, 192–204. https://doi.org/10.1016/j.advwatres.2017.11.011

Abdolghafoorian, A., & Farhadi, L. (2016). Uncertainty quantification in land surface hydrologic modeling: Toward an integrated variational data assimilation framework. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *9*(6), 2628–2637. https://doi.org/10.1109/JSTARS.2016.2553444

Anderson, J. L. (2001). An ensemble adjustment Kalman filter for data assimilation. *Monthly Weather Review*, *129*(12), 2884–2903. https://doi.org/10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2

Anderson, J. L., & Anderson, S. L. (1999). A Monte Carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Monthly Weather Review*, *127*(12), 2741–2758. https://doi.org/10.1175/1520-0493(1999)127<2741:AMCIOT>2.0.CO;2

Andrieu, C., Doucet, A., & Holenstein, R. (2010). Particle Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society Series B-Statistical Methodology*, *72*(3), 269–342. https://doi.org/10.1111/j.1467-9868.2009.00736.x

Atkins, E., Morzfeld, M., & Chorin, A. J. (2013). Implicit particle methods and their connection with variational data assimilation. *Monthly Weather Review*, *141*(6), 1786–1803. https://doi.org/10.1175/MWR-D-12-00145.1

Bannister, R. N. (2017). A review of operational methods of variational and ensemble-variational data assimilation. *Quarterly Journal of the Royal Meteorological Society*, *143*(703), 607–633. https://doi.org/10.1002/qj.2982. https://doi.org/10.1002/qj.2982

Boyle, D. (2001). *Multicriteria calibration of hydrological models*, (PhD dissertation). Tucson: University of Arizona.

Bulygina, N., & Gupta, H. (2011). Correcting the mathematical structure of a hydrological model via Bayesian data assimilation. *Water Resources Research*, *47*, W05514. https://doi.org/10.1029/2010WR009614

Burnash, R. J., Ferral, R. L., & McGuire, R. A. (1973). A generalized streamflow simulation system: Conceptual modeling for digital computers, report, U. S. Department of Commerce, National Weather Service, and State of California, Department of Water Resources.

Chorin, A. J., Morzfeld, M., & Tu, X. (2010). Implicit filters for data assimilation. *Communications in Applied Mathematics and Computational Science*, *5*(2), 221–240. https://doi.org/10.2140/camcos.2010.5.221

Chorin, A. J., & Tu, X. (2009). Implicit sampling for particle filters. *Proceedings of the National Academy of Sciences*, *106*(41), 17,249–17,254. https://doi.org/10.1073/pnas.0909196106

Clark, M. P., Rupp, D. E., Woods, R. A., Zheng, X., Ibbitt, R. P., Slater, A. G., et al. (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*(10), 1309–1324. https://doi.org/10.1016/j.advwatres.2008.06.005

Crow, W. T., & Wood, E. F. (2003). The assimilation of remotely sensed soil brightness temperature imagery into a land surface model using ensemble Kalman filtering: A case study based on ESTAR measurements during SGP97. *Advances in Water Resources*, *26*(2), 137–149. https://doi.org/10.1016/S0309-1708(02)00088-X

De Lannoy, G. J. M., Reichle, R. H., Houser, P. R., Pauwels, V. R. N., & Verhoest, N. E. C. (2007). Correcting for forecast bias in soil moisture assimilation with the ensemble Kalman filter. *Water Resources Research*, *43*, W09410. https://doi.org/10.1029/2006WR005449

DeChant, C. M., & Moradkhani, H. (2011). Improving the characterization of initial condition for ensemble streamflow prediction using data assimilation. *Hydrology and Earth System Sciences*, *15*(11), 3399–3410. https://doi.org/10.5194/hess-15-3399-2011

DeChant, C. M., & Moradkhani, H. (2012). Examining the effectiveness and robustness of sequential data assimilation methods for quantification of uncertainty in hydrologic forecasting. *Water Resources Research*, *48*, W04518. https://doi.org/10.1029/2011WR011011

Dong, J., Steele-Dunne, S. C., Judge, J., & Van De Giesen, N. (2015). A particle batch smoother for soil moisture estimation using soil temperature observations. *Advances in Water Resources*, *83*, 111–122. https://doi.org/10.1016/j.advwatres.2015.05.017. https://doi.org/10.1016/j.advwatres.2015.05.017

Doucet, A., & Johansen, A. M.. (2009). A tutorial on particle filtering and smoothing: Fifteen years later. Handbook nonlinear Filter.12, 656–704.

Duan, Q., Schaake, J., Andréassian, V., Franks, S., Goteti, G., Gupta, H. V., et al. (2006). Model parameter estimation experiment (MOPEX): An overview of science strategy and major results from the second and third workshops. *Journal of Hydrology*, *320*, 3–17. https://doi.org/10.1016/j.jhydrol.2005.07.031

Efstratiadis, A., & Koutsoyiannis, D. (2010). Une décennie d'approches de calage multi-objectifs en modélisation hydrologique: Une revue. *Hydrological Sciences Journal*, *55*(1), 58–78. https://doi.org/10.1080/02626660903526292

Gao, F., & Han, L. (2012). Implementing the Nelder-Mead simplex algorithm with adaptive parameters. *Computational Optimization and Applications*, *51*(1), 259–277. https://doi.org/10.1007/s10589-010-9329-3

Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, *377*(1–2), 80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003

Han, H., Ding, Y. S., Hao, K. R., & Liang, X. (2011). An evolutionary particle filter with the immune genetic algorithm for intelligent video target tracking. *Computers and Mathematics with Applications*, *62*(7), 2685–2695. https://doi.org/10.1016/j.camwa.2011.06.050

Hernández, F., & Liang, X. (2018). Hybridizing Bayesian and variational data assimilation for robust high-resolution hydrologic forecasting. *Hydrology and Earth System Sciences*, *22*, 5759–5779. https://doi.org/10.5194/hess-22-5759-2018

Jones, A. S., Vukićević, T., & Vonder Haar, T. H. (2004). A microwave satellite observational operator for variational data assimilation of soil moisture. *Journal of Hydrometeorology*, *5*(1), 213–229. https://doi.org/10.1175/1525-7541(2004)005<0213:AMSOOF>2.0.CO;2

Kwok, N. M., Fang, G., & Zhou, W. (2005). Evolutionary particle filter: Re-sampling from the genetic algorithm perspective. 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, 1053–1058. https://doi.org/10.1109/IROS.2005.1545119

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*(12), 1597–1615. https://doi.org/10.1016/j.advwatres.2011.08.012

Leisenring, M., & Moradkhani, H. (2012). Analyzing the uncertainty of suspended sediment load prediction using sequential data assimilation. *Journal of Hydrology*, *468-469*, 268–282. https://doi.org/10.1016/j.jhydrol.2012.08.049

Liang, F., & Wong, W. H. (2001). Real-parameter evolutionary Monte Carlo with applications to Bayesian mixture models. *Journal of the American Statistical Association*, *96*(454), 653–666. https://doi.org/10.1198/016214501753168325

Liu, Y., & Gupta, H. V. (2007). Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. *Water Resources Research*, *43*, W07401. https://doi.org/10.1029/2006WR005756

Lorenc, A. C. (2017). Improving ensemble covariances in hybrid variational data assimilation without increasing ensemble size. *Quarterly Journal of the Royal Meteorological Society*, *143*(703), 1062–1072. https://doi.org/10.1002/qj.2990

Matgen, P., Montanari, M., Hostache, R., Pfister, L., Hoffmann, L., Plaza, D., et al. (2010). Towards the sequential assimilation of SAR-derived water stages into hydraulic models using the particle filter: Proof of concept. *Hydrology and Earth System Sciences*, *14*(9), 1773–1785. https://doi.org/10.5194/hess-14-1773-2010

Moradkhani, H., DeChant, C. M., & Sorooshian, S. (2012). Evolution of ensemble data assimilation for uncertainty quantification using the particle filter-Markov chain Monte Carlo method. *Water Resources Research*, *48*, L22702. https://doi.org/10.1029/2012WR012144

Moradkhani, H., Hsu, K., Hong, Y., & Sorooshian, S. (2006). Investigating the impact of remotely sensed precipitation and hydrologic model uncertainties on the ensemble streamflow forecasting. *Geophysical Research Letters*, *33*, L22702. https://doi.org/10.1029/2006GL026855

Moradkhani, H., Hsu, K.-L., Gupta, H., & Sorooshian, S. (2005). Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter. *Water Resources Research*, *41*, W01403. https://doi.org/10.1029/2004WR003604

Moradkhani, H., Nearing, G., Abbaszadeh, P., & Pathiraja, S. (2018). Fundamentals of data assimilation and theoretical advances. In Duan, et al. (Eds.), *Handbook of Hydrometeorological Ensemble Forecasting* (pp. 1–26). Heidelberg: Springer, Berlin. https://doi.org/10.1007/978-3-642-40457-3_30-1

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*(2), 135–147. https://doi.org/10.1016/j.advwatres.2004.09.002

Morzfeld, M., Tu, X., Atkins, E., & Chorin, A. J. (2012). A random map implementation of implicit filters. *Journal of Computational Physics*, *231*(4), 2049–2066. https://doi.org/10.1016/j.jcp.2011.11.022

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I—A discussion of principles*. *Journal of Hydrology*, *10*(3), 282–290. https://doi.org/10.1016/0022-1694(70)90255-6

Noh, S. J., Tachikawa, Y., Shiiba, M., & Kim, S. (2011). Applying sequential Monte Carlo methods into a distributed hydrologic model: Lagged particle filtering approach with regularization. *Hydrology and Earth System Sciences*, *15*(10), 3237–3251. https://doi.org/10.5194/hess-15-3237-2011

Parrish, M. A., Moradkhani, H., & Dechant, C. M. (2012). Toward reduction of model uncertainty: Integration of Bayesian model averaging and data assimilation. *Water Resources Research*, *48*, W03519. https://doi.org/10.1029/2011WR011116

Pathiraja, S., Anghileri, D., Burlando, P., Sharma, A., Marshall, L., & Moradkhani, H. (2018a). Insights on the impact of systematic model errors on data assimilation performance in changing catchments. *Advances in Water Resources*, *113*, 202–222. https://doi.org/10.1016/j.advwatres.2017.12.006

Pathiraja, S., Anghileri, D., Burlando, P., Sharma, A., Marshall, L., & Moradkhani, H. (2018b). Time-varying parameter models for catchments with land use change: The importance of model structure. *Hydrology and Earth System Sciences*, *22*(5), 2903–2919. https://doi.org/10.5194/hess-22-2903-2018

Pathiraja, S., Marshall, L., Sharma, A., & Moradkhani, H. (2016). Detecting non-stationary hydrologic model parameters in a paired catchment system using data assimilation. *Advances in Water Resources*, *94*, 103–119. https://doi.org/10.1016/j.advwatres.2016.04.021

Pathiraja, S., Moradkhani, H., Marshall, L., Sharma, A., & Geenens, G. (2018). Data-driven model uncertainty estimation in hydrologic data assimilation. *Water Resources Research*., *54*(2), 1252–1280. https://doi.org/10.1002/2018WR022627

Plaza, D. A., De Keyser, R., De Lannoy, G. J. M., Giustarini, L., Matgen, P., & Pauwels, V. R. N. (2012). The importance of parameter resampling for soil moisture data assimilation into hydrologic models using the particle filter. *Hydrology and Earth System Sciences*, *16*(2), 375–390. https://doi.org/10.5194/hess-16-375-2012

Reichle, R., McLaughlin, D. B., & Entekhabi, D. (2001). Variational data assimilation of microwave radio brightness observations for land surface hydrologic applications. *IEEE Transactions on Geoscience and Remote Sensing*, *39*(8), 1708–1718. https://doi.org/10.1109/36.942549

Reichle, R., McLaughlin, D. B., & Entekhabi, D. (2002). Hydrologic data assimilation with the ensemble Kalman filter. *Monthly Weather Review*, *130*(1), 103–114. https://doi.org/10.1175/1520-0493(2002)130<0103:HDAWTE>2.0.CO;2

Renard, B., Kavetski, D., Kuczera, G., Thyer, M., & Franks, S. W. (2010). Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors. *Water Resources Research*, *46*, W05521. https://doi.org/10.1029/2009WR008328

Rios, L. M., & Sahinidis, N. V. (2013). Derivative-free optimization: A review of algorithms and comparison of software implementations. *Journal of Global Optimization*, *56*(3), 1247–1293. https://doi.org/10.1007/s10898-012-9951-y

Salamon, P., & Feyen, L. (2009). Assessing parameter, precipitation, and predictive uncertainty in a distributed hydrological model using sequential data assimilation with the particle filter. *Journal of Hydrology*, *376*(3–4), 428–442. https://doi.org/10.1016/j.jhydrol.2009.07.051

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*(3–4), 255–275. https://doi.org/10.1016/j.jhydrol.2009.01.019

Shaw, J. A., & Daescu, D. N. (2016). An ensemble approach to weak-constraint four-dimensional variational data assimilation. *Procedia Computer Science*, *80*, 496–506. https://doi.org/10.1016/j.procs.2016.05.329

Shoaib, S. A., Marshall, L., & Sharma, A. (2016). A metric for attributing variability in modelled streamflows. *Journal of Hydrology*, *541*, 1475–1487. https://doi.org/10.1016/j.jhydrol.2016.08.050

Slivinski, L., Spiller, E., Apte, A., & Sandstede, B. (2015). A hybrid particle–ensemble Kalman filter for Lagrangian data assimilation. *Monthly Weather Review*, *143*(1), 195–211. https://doi.org/10.1175/MWR-D-14-00051.1

Talagrand, O., & Courtier, P. (1987). Variational assimilation of meteorological observations with the adjoint vorticity equation I: Theory. *Quarterly Journal of the Royal Meteorological Society*, *113*(478), 1311–1328. https://doi.org/10.1002/qj.49711347812

Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram. *Journal of Geophysical Research*, *106*(D7), 7183–7192. https://doi.org/10.1029/2000JD900719

Tian, X., Xie, Z., Dai, A., Shi, C., Jia, B., Chen, F., & Yang, K. (2009). A dual-pass variational data assimilation framework for estimating soil moisture profiles from AMSR-E microwave brightness temperature. *Journal of Geophysical Research*, *114*(16), 1–12. https://doi.org/10.1029/2008JD011600

Tremolet, Y. (2006). Accounting for an imperfect model in 4D-Var. *Quarterly Journal of the Royal Meteorological Society*, *132*(621), 2483–2504. https://doi.org/10.1256/qj.05.224

van Leeuwen, P. J. (2015). Nonlinear data assimilation for high-dimensional systems. In P. J. van Leeuwen, C. Yuan, & S. Reich (Eds.), *Nonlinear Data Assimilation* (Chap. 1, pp. 1–73). Basel, Switzerland: Springer.

Verkade, J. S., & Werner, M. G. F. (2011). Estimating the benefits of single value and probability forecasting for flood warning. *Hydrology and Earth System Sciences*, *15*(12), 3751–3765. https://doi.org/10.5194/hess-15-3751-2011

Vrugt, J. A., Gupta, H. V., Nualláin, B., & Bouten, W. (2006). Real-time data assimilation for operational ensemble streamflow forecasting. *Journal of Hydrometeorology*, *7*(3), 548–565. https://doi.org/10.1175/JHM504.1

Wang, X., Barker, D. M., Snyder, C., & Hamill, T. M. (2008). A hybrid ETKF–3DVAR data assimilation scheme for the WRF model. Part I: Observing system simulation experiment. *Monthly Weather Review*, *136*(12), 5116–5131. https://doi.org/10.1175/2008MWR2444.1

Wu, G., Zheng, X., Wang, L., Zhang, S., Liang, X., & Li, Y. (2013). A new structure for error covariance matrices and their adaptive estimation in EnKF assimilation. *Quarterly Journal of the Royal Meteorological Society*, *139*(672), 795–804. https://doi.org/10.1002/qj.2000

Yan, H., DeChant, C. M., & Moradkhani, H. (2015). Improving soil moisture profile prediction with the particle filter-Markov chain MonteCarlo method. *IEEE Transactions on Geoscience and Remote Sensing*, *53*, 6134–6147.

Yan, H., Moradkhani, H., & Zarekarizi, M. (2017). A probabilistic drought forecasting framework: A combined dynamical and statistical approach. *Journal of Hydrology*, *548*, 291–304. https://doi.org/10.1016/j.jhydrol.2017.03.004. https://doi.org/10.1016/j.jhydrol.2017.03.004

Yin, S., Zhu, X., & Member, S. (2015). Intelligent particle filter and its application to fault detection of nonlinear system. *IEEE Transactions on Industrial Electronics*, *62*(6), 3852–3861. https://doi.org/10.1109/TIE.2015.2399396

Zhu, G., Li, X., Ma, J., Wang, Y., Liu, S., Huang, C., et al. (2018). A new moving strategy for the sequential Monte Carlo approach in optimizing the hydrological model parameters. *Advances in Water Resources*, *114*, 164–179. https://doi.org/10.1016/j.advwatres.2018.02.007

Zhu, M., van Leeuwen, P. J., & Amezcua, J. (2016). Implicit equal-weights particle filter. *Quarterly Journal of the Royal Meteorological Society*, *142*(698), 1904–1919. https://doi.org/10.1002/qj.2784

Zhu, M., van Leeuwen, P. J., & Zhang, W. (2017). Estimating model error covariances using particle filters. *Quarterly Journal of the Royal Meteorological Society*, *144*(713), 1310–1320. https://doi.org/10.1002/qj.3132

Zhu, Y., Toth, Z., Wobus, R., Richardson, D., & Mylne, K. (2002). The economic value of ensemble-based weather forecasts. *Bulletin of the American Meteorological Society*, *83*(1), 73–83. https://doi.org/10.1175/1520-0477(2002)083<0073:TEVOEB>2.3.CO;2

Zupanski, D. (1997). A general weak constraint applicable to operational 4DVAR data assimilation systems. *Monthly Weather Review*, *125*(9), 2274–2292. https://doi.org/10.1175/1520-0493(1997)125<2274:AGWCAT>2.0.CO;2