Journal of Hydrology 548 (2017) 484–497

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

# Towards robust quantification and reduction of uncertainty in hydrologic predictions: Integration of particle Markov chain Monte Carlo and factorial polynomial chaos expansion

S. Wang [a,*], G.H. Huang [b], B.W. Baetz [c], B.C. Ancell [a]

[a] Department of Geosciences, Texas Tech University, Lubbock, TX, USA
[b] Institute for Energy, Environment and Sustainable Communities, University of Regina, Regina, Saskatchewan, Canada
[c] Department of Civil Engineering, McMaster University, Hamilton, Ontario, Canada

A R T I C L E I N F O

*Article history:*
Received 13 December 2016
Received in revised form 28 February 2017
Accepted 13 March 2017
Available online 15 March 2017
This manuscript was handled by Corrado Corradini, Editor-in-Chief, with the assistance of Renato Morbidelli, Associate Editor

*Keywords:*
Hydrologic prediction
Uncertainty
Data assimilation
Particle filter
Markov chain Monte Carlo
Polynomial chaos expansion

A B S T R A C T

The particle filtering techniques have been receiving increasing attention from the hydrologic community due to its ability to properly estimate model parameters and states of nonlinear and non-Gaussian systems. To facilitate a robust quantification of uncertainty in hydrologic predictions, it is necessary to explicitly examine the forward propagation and evolution of parameter uncertainties and their interactions that affect the predictive performance. This paper presents a unified probabilistic framework that merges the strengths of particle Markov chain Monte Carlo (PMCMC) and factorial polynomial chaos expansion (FPCE) algorithms to robustly quantify and reduce uncertainties in hydrologic predictions. A Gaussian anamorphosis technique is used to establish a seamless bridge between the data assimilation using the PMCMC and the uncertainty propagation using the FPCE through a straightforward transformation of posterior distributions of model parameters. The unified probabilistic framework is applied to the Xiangxi River watershed of the Three Gorges Reservoir (TGR) region in China to demonstrate its validity and applicability. Results reveal that the degree of spatial variability of soil moisture capacity is the most identifiable model parameter with the fastest convergence through the streamflow assimilation process. The potential interaction between the spatial variability in soil moisture conditions and the maximum soil moisture capacity has the most significant effect on the performance of streamflow predictions. In addition, parameter sensitivities and interactions vary in magnitude and direction over time due to temporal and spatial dynamics of hydrologic processes.



## 1. Introduction

Hydrologic systems involve dynamic interactions between water, climate, vegetation, and soil processes. Hydrologic models typically conceptualize the complex behavior of hydrologic systems by using mathematical equations, in which model parameters represent temporal and spatial variability in watershed characteristics that cannot be observed or measured explicitly. As a result, the predictive performance of hydrologic models is inevitably affected by various sources of uncertainty, including the descriptions of boundary and initial conditions, the errors in model forcing data and output observations, inaccurate parameter estimates, and model structural deficiencies (Ajami et al., 2007; DeChant and Moradkhani, 2014). Therefore, robust quantification and reduction of uncertainties are necessary to improve the reliability of hydrologic predictions.

Over the past few decades, tremendous efforts have been made in the development of data assimilation techniques for addressing various sources of uncertainty in hydrologic modeling (Liu and Gupta, 2007; Ryu et al., 2009; Gharamti et al., 2013; Abaza et al., 2014; Panzeri et al., 2014; Randrianasolo et al., 2014; Sun et al., 2015; Xu and Gómez-Hernández, 2015). Data assimilation techniques are recognized as a powerful tool for probabilistic hydrologic predictions through recursively updating model states and parameters when new observations become available. Previously, the ensemble Kalman filter (EnKF) introduced by Evensen (1994) was the most commonly used technique for uncertainty assessment of hydrologic model parameters and state variables due to its attractive features of real-time adjustment and efficient implementation (Reichle et al., 2002; Moradkhani et al., 2005b; Xie and Zhang, 2010; Cammalleri and Ciraolo, 2012; Rafieeinasab et al.,

* Corresponding author.
*E-mail address:* shuo.wang@ttu.edu (S. Wang).

http://dx.doi.org/10.1016/j.jhydrol.2017.03.027
0022-1694/

2014; Samuel et al., 2014; Gharamti et al., 2015; Liu et al., 2016; Pathiraja et al., 2016). However, the EnKF relies on a Gaussian assumption that cannot be met in practical applications, resulting in unrealistic model simulations and unreliable hydrologic predictions.

In recent years, the particle filter (PF) technique has been introduced as an attractive alternative to remove the unrealistic Gaussian assumption of the EnKF, improving the reliability of hydrologic predictions. The PF is able to fully represent the posterior distributions of model parameters and state variables through a number of independent random samples called particles, and the particles are weighted and propagated sequentially by assimilating available observations. Over the past few years, the PF and its variants have been receiving increasing attention from the hydrologic community due to its ability to properly estimate the state of nonlinear and non-Gaussian systems (Weerts and El Serafy, 2006; Smith et al., 2008; Salamon and Feyen, 2009; DeChant and Moradkhani, 2012; Dumedah and Coulibaly, 2013; Bi et al., 2015). Moradkhani et al. (2005a) introduced a PF with sampling importance resampling (PF-SIR) algorithm for uncertainty assessment of hydrologic model states and parameters. To further improve the PF-SIR algorithm, Moradkhani et al. (2012) proposed a PF with a Markov chain Monte Carlo (PF-MCMC) algorithm to increase parameter diversity within the posterior distribution, reducing the risk of sample impoverishment and leading to a more accurate streamflow forecast. Plaza Guingla et al. (2013) proposed two alternatives that included a resample-move step in the standard PF and an optimal importance density function in the Gaussian PF to improve the effectiveness of the PF. Vrugt et al. (2013) combined the strengths of sequential Monte Carlo sampling and the MCMC simulation with the DiffeRential Evolution Adaptive Metropolis (DREAM) algorithm for the joint estimation of hydrologic model parameters and state variables. Yan et al. (2015) used the PF-SIR and PF-MCMC algorithms to estimate soil moisture states and soil hydraulic parameters through assimilating remotely sensed near-surface soil moisture measurements.

The PF has been introduced as a powerful tool to reduce uncertainty in the joint estimation of hydrologic model parameters and state variables through sequential assimilation of observations. When the posterior distributions of model parameters are derived from the PF, uncertainty is propagated through the hydrologic models, leading to probabilistic hydrologic predictions. In addition, model parameters are correlated with each other in the uncertainty propagation process, and their interactions have significant impacts on the predictive performance. Therefore, it is necessary to reveal the forward propagation and evolution of parameter uncertainties and their interactions after the data assimilation operation, leading to a robust quantification and reduction of uncertainty in hydrologic predictions.

Polynomial chaos expansion (PCE) techniques have been extensively used to represent stochastic processes through the propagation of random uncertainties in dynamic systems (Xiu and Karniadakis, 2002; Marzouk and Najm, 2009; Najm, 2009; Konda et al., 2010; Oladyshkin et al., 2011). In recent years, the PCE has attracted increased attention in hydrologic studies (Lin and Tartakovsky, 2009; Müller et al., 2011; Ciriello et al., 2013; Sochala and Le Maître, 2013; Dai et al., 2016). Fajraoui et al. (2011) used global sensitivity analysis in conjunction with the PCE methodology to provide valuable information for the interpretation of transport experiments in laboratory-scale heterogeneous porous media. Rajabi et al. (2015) used the non-intrusive PCE for efficient uncertainty propagation and moment independent sensitivity analysis of seawater intrusion simulations. Wang et al. (2015b) developed a polynomial chaos ensemble hydrologic prediction system for efficiently quantifying uncertainties in hydrologic predictions. Consequently, the PCE is recognized as a promising technique for explicitly revealing the propagation of uncertainty through hydrologic models and for efficiently quantifying uncertainty in hydrologic predictions.

In this work, we propose a unified probabilistic framework to robustly quantify and reduce uncertainties in hydrologic predictions. In the unified probabilistic framework, a particle Markov chain Monte Carlo (PMCMC) algorithm will be first used to infer the posterior distributions of hydrologic model parameters through sequential assimilation of available observations. After the data assimilation operation, a factorial polynomial chaos expansion (FPCE) technique will then be introduced to examine the forward propagation and evolution of uncertainty by revealing complex parameter interactions and their impacts on the predictive performance of hydrologic models. In addition, a Gaussian anamorphosis technique will be used to establish a seamless bridge between the data assimilation using the PMCMC and the uncertainty propagation using the FPCE through a straightforward transformation of posterior distributions of model parameters. The proposed unified probabilistic framework will be applied to predict daily streamflow in the Xiangxi River watershed which is located in the Three Gorges Reservoir (TGR) region, China.

This paper is organized as follows. Section 2 introduces the proposed unified probabilistic framework for robustly quantifying uncertainties in hydrologic predictions. Section 3 provides details on the study area and the experimental setup. Section 4 presents a systematic analysis of sequential streamflow assimilation and uncertainty quantification along with a thorough discussion on the recursive inference of model parameters and the explicit characterization of parameter interactions. Finally, conclusions are drawn in Section 5.

## 2. Development of a unified probabilistic framework

The unified probabilistic framework merges the strengths of the PMCMC and the FPCE algorithms for robustly quantifying and reducing uncertainties in hydrologic predictions. A general overview of the steps involved within the unified probabilistic framework is provided as follows: (1) inference of model parameters and state variables through sequential data assimilation using the PMCMC; (2) transformation of irregular posterior distributions of model parameters into standard normal distributions using the Gaussian anamorphosis technique; (3) characterization of uncertainty propagation and evolution using the FPCE technique; (4) quantification of uncertainties in hydrologic predictions through the propagation of parameter uncertainties; and (5) examination of sensitivities of hydrologic model parameters and their interactions that affect the predictive performance using factorial analysis.

### 2.1. Particle Markov chain Monte Carlo

The PF technique is recognized as an effective means to remove the unrealistic assumption of Gaussian errors involved in hydrologic modeling, improving the robustness of hydrologic predictions. The PF is a sequential Monte Carlo approach that can be used to implement a recursive Bayesian filter through Monte Carlo simulations, by which the posterior distributions of state variables are represented by a set of random particles with corresponding importance weights (Bi et al., 2015). Thus, the most important concept in the particle filtering is the sequential importance sampling (SIS) used for estimating the particle weights in a recursive form.

To better understand the SIS algorithm, the hydrologic model can be formulated as follows:

$$x_{i,t+1}^{-} = f(x_{i,t}^{+}, u_{i,t+1}^{-}, \theta_{i,t+1}^{-}) + \omega_{i,t+1}, \omega_{i,t+1} \sim N(0, Var_{t+1}) \quad (1)$$