Available online at www.sciencedirect.com

SCIENCE DIRECT®

Advances in Water Resources 28 (2005) 135–147

Advances in Water Resources

www.elsevier.com/locate/advwatres

Dual state–parameter estimation of hydrological models using ensemble Kalman filter

Hamid Moradkhani [a,*], Soroosh Sorooshian [a], Hoshin V. Gupta [b], Paul R. Houser [c]

[a] *Department of Civil and Environmental Engineering, University of California, Irvine, CA 92697-2175, USA*
[b] *Department of Hydrology and Water Resources, University of Arizona, Tucson, AZ 85721, USA*
[c] *Hydrological Sciences Branch, NASA-GSFC Code 974, Greenbelt, MD 20771, USA*

Received 25 March 2004; received in revised form 20 September 2004; accepted 21 September 2004

Abstract

Hydrologic models are twofold: models for understanding physical processes and models for prediction. This study addresses the latter, which modelers use to predict, for example, streamflow at some future time given knowledge of the current state of the system and model parameters. In this respect, good estimates of the parameters and state variables are needed to enable the model to generate accurate forecasts. In this paper, a dual state–parameter estimation approach is presented based on the Ensemble Kalman Filter (EnKF) for sequential estimation of both parameters and state variables of a hydrologic model. A systematic approach for identification of the perturbation factors used for ensemble generation and for selection of ensemble size is discussed. The dual EnKF methodology introduces a number of novel features: (1) both model states and parameters can be estimated simultaneously; (2) the algorithm is recursive and therefore does not require storage of all past information, as is the case in the batch calibration procedures; and (3) the various sources of uncertainties can be properly addressed, including input, output, and parameter uncertainties. The applicability and usefulness of the dual EnKF approach for ensemble streamflow forecasting is demonstrated using a conceptual rainfall-runoff model.

Keywords: Streamflow forecasting; Stochastic processes; Data assimilation; Ensemble Kalman filter; Dual estimation; Kernel smoothing

1. Introduction and scope

Hydrologic models are defined largely by parameters and states, parameters being physical and generally time-invariant descriptions of surface and subsurface characteristics, and states being fluxes and storages of water and energy that are propagated in time by the model physics. In practice, in addition to model simulation, reliable operation of a watershed system requires a continuous correction of the forecast as observational data become available. This entails the critical need to extend the applicability of data assimilation in hydrology as emphasized by Troch et al. [39]. However, the successful use of data assimilation relies on unbiased model state prediction, which is largely dependent on accurate parameter estimation. During the past two decades, much effort has been directed toward the estimation of hydrologic model parameters (calibration) to improve the forecast accuracy [7,8,11,32]. Conceptual hydrologic models are usually deterministic representations, which typically do not contain descriptions of the various sources of uncertainties. Although it has been common to translate the inability of a model to generate accurate streamflow forecasts into parameter uncertainty, other sources of uncertainties, such as model structural error, input, and output measurement errors, also need to be accounted for [16,17]. Several authors have studied the uncertainties associated with parameter estimation, and procedures have been developed for the statistical analyses of parameter uncertainties [18–20,33,34,42,43].

The aforementioned calibration procedures generally minimize long-term prediction error using a historical batch of data assuming time-invariant parameters, and

* Corresponding author. Tel.: +1 949 824 8821; fax: +1 949 884 8831.
E-mail address: moradkha@uci.edu (H. Moradkhani).

0309-1708/$ - see front matter © 2004 Elsevier Ltd. All rights reserved.
doi:10.1016/j.advwatres.2004.09.002

thus make no attempt to include information from new observations. Batch calibration requires a set of historical data to be kept in storage and processed enmasse while lacking the flexibility to investigate possible temporal evolution of the model parameters. Thiemann et al. [35] emphasized another limitation of batch calibration in hydrological prediction of an ungauged watershed where the lack of sufficient historical data makes the batch method infeasible. These limitations, as well as an interest in inferring the uncertainty in the estimated parameters, motivated Thiemann et al. [35] and Misirli et al. [25] to develop a recursive scheme for model prediction and parameter estimation in the online mode. From another perspective, Kitanidis and Bras [15] stated that adaptive estimation might be suitable when the forecast lead-time is short in comparison to the response time of the watershed. They explained that it would be the case when the error in input is large while the error in output measurement is small.

Much of the efforts in simulation-based methods of hydrologic system analyses have been focused on (1) improved methods for parameter estimation wherein state variable uncertainties were not explicitly taken into account or (2) improved procedures for estimating time-varying state variables wherein the parameters were assumed to be known in advance. The commonly used batch calibration techniques only address parameter uncertainty while uncertainties in input, output and model structure are ignored. The main weakness of such approaches is that they attribute all errors from input, output and model structure to model parameter uncertainty. Sequential data assimilation procedures have the potential to overcome this drawback in simulation-based methods by explicitly taking into account all the sources of uncertainty. The Kalman filter [14], a recursive data-processing algorithm, is the most commonly used sequential data assimilation technique, which results in optimal estimation for linear dynamic models with Gaussian uncertainties.

Although filtering techniques can address the various sources of uncertainties in modeling, the typical presumption of these procedures is that the parameters are to be specified in advance and sequential estimation is applied only to the state variables. Because there is no guarantee that model behavior does not change over time, the model adjustment through the time variation of parameters together with state variables is incisive. Therefore a procedure that can provide the simultaneous estimate of states and parameters is required. The development of interactive (dual) state–parameter estimation using standard Kalman filter, in the context of hydrology, is traced back to [36,37] and later to the joint state–parameter by state augmentation technique [3,4] (see Section 3 for detail). Those techniques, however, were limited to linear dynamic systems. For nonlinear dynamics, the extended Kalman filter (EKF), which relies on linearization of model using first order approximation of Taylor series, can be used. As reported by Refs. [9,29,30] the EKF can lead to unstable results when the nonlinearity in the system is strong. To cope with the drawbacks of the EKF, a Monte Carlo-based Kalman filter called ensemble Kalman filter (EnKF) was introduced by Evensen [9]. One of the advantages of the EnKF comparing to the standard KF is that the estimation of priori model covariance (see Section 2.1) is not needed for the updating (analysis) step although its calculation using the model ensemble is straightforward.

The EnKF was originally developed for dynamic state estimation while in this paper its applicability to static state (parameter) estimation by dual state–parameter estimation strategy is extended and its usefulness on streamflow forecasting is examined.

The organization of the paper is as follow. In Section 2, the general framework for sequential data assimilation is explained, where the mathematical formulation of the EnKF as a special type of Monte Carlo procedure for state estimation is elaborated. A systematic approach for identifying the perturbation factor, as a key feature in the EnKF, and for tackling the uncertainties in forcing data (input) and observation (output) is suggested. In Section 3, the dual EnKF algorithm that deals simultaneously with both model parameters and state variables is explained and kernel smoothing of parameters is employed for parameter sampling to avoid the over-dispersion of parameters through random walk. In Section 4, the applicability of dual EnKF on a conceptual rainfall-runoff model and the power of this algorithm in streamflow forecasting is demonstrated.

5. Summary and conclusion

Hydrologic models are still far from perfect, and hydrologists need to put the models in better compliance with observations prior to use in forecasting. Batch calibration procedures as the most commonly used techniques in hydrology, and even the recursive calibration schemes concern primarily the estimation of parameters and the identification of uncertainties associated with them. However, more general algorithms that account for the simultaneous interactions of model states and parameters are encouraging while different sources of errors are considered. In this study, an integrated and algorithmic framework for dual state–parameter estimation using EnKF was presented, which leads to the ensemble streamflow forecasting. Perturbation of input and output to generate and modify the ensemble of model variables and to determine the ensemble size are key features of the EnKF, and identification of the magnitude of perturbation in a systematic framework is desired and elaborated in this study.