Journal of Hydrology 529 (2015) 1147–1160

Contents lists available at ScienceDirect

Journal of Hydrology

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

On the difficulty to optimally implement the Ensemble Kalman filter: An experiment based on many hydrological models and catchments

A. Thiboult *, F. Anctil

Dept. of Civil and Water Engineering, Laval University, 1065 avenue de la Medecine, Quebec G1V 0A6, Canada

ARTICLE INFO

Article history:
Received 28 April 2015
Received in revised form 10 September 2015
Accepted 14 September 2015
Available online 21 September 2015
This manuscript was handled by Konstantine P. Georgakakos, Editor-in-Chief, with the assistance of Hamid Moradkhani, Associate Editor

Keywords:
Data assimilation
Uncertainty estimation
Ensemble Kalman filter

SUMMARY

Forecast reliability and accuracy is a prerequisite for successful hydrological applications. This aim may be attained by using data assimilation techniques such as the popular Ensemble Kalman filter (EnKF). Despite its recognized capacity to enhance forecasting by creating a new set of initial conditions, implementation tests have been mostly carried out with a single model and few catchments leading to case specific conclusions. This paper performs an extensive testing to assess ensemble bias and reliability on 20 conceptual lumped models and 38 catchments in the Province of Québec with perfect meteorological forecast forcing. The study confirms that EnKF is a powerful tool for short range forecasting but also that it requires a more subtle setting than it is frequently recommended. The success of the updating procedure depends to a great extent on the specification of the hyper-parameters. In the implementation of the EnKF, the identification of the hyper-parameters is very unintuitive if the model error is not explicitly accounted for and best estimates of forcing and observation error lead to overconfident forecasts. It is shown that performance are also related to the choice of updated state variables and that all states variables should not systematically be updated. Additionally, the improvement over the open loop scheme depends on the watershed and hydrological model structure, as some models exhibit a poor compatibility with EnKF updating. Thus, it is not possible to conclude in detail on a single ideal manner to identify an optimal implementation; conclusions drawn from a unique event, catchment, or model are likely to be misleading since transferring hyper-parameters from a case to another may be hazardous. Finally, achieving reliability and bias jointly is a daunting challenge as the optimization of one score is done at the cost of the other.

1. Introduction

Despite the modelling advances in representing hydrological processes and providing more accurate streamflow forecasts, there is still a need for reducing and quantifying uncertainty. Most hydrological prediction systems are still deterministic and provide only the most likely outcome without addressing estimates of their uncertainty. The sources of uncertainty stem from multiple places in the hydrometeorological chain such as in inputs, initial conditions, parameter estimation, model structure, and outputs (e.g. Ajami et al., 2007; Salamon and Feyen, 2010; Liu and Gupta, 2007; Liu et al., 2012) and these uncertainties should be deciphered to enhance model predictive abilities and reliability for efficient decision making (Ramos et al., 2010).

A broad range of techniques has been developed to control uncertainty at different levels such as the Generalized Likelihood Uncertainty Estimation (GLUE), Shuffle Complex Evolution Metropolis algorithm (SCEM) for parameter uncertainty (Beven and Binley, 1992; Vrugt et al., 2003) and BMA combination technique for structural uncertainty (Jeremiah et al., 2011; Duan et al., 2007; Parrish et al., 2012; Ajami et al., 2007). Proper initial conditions are frequently identified as one of the main factors that contributes to an accurate forecast (DeChant and Moradkhani, 2011; Lee et al., 2011). Among others, data assimilation (DA) is commonly used in hydrometeorology to reduce initial condition uncertainty and proved to be a useful tool for modelling. DA incorporates observations into the numerical model to issue an analysis, which is an estimation of the best current state of the system. This has not only been largely applied to remote sensing for snow (Kuchment et al., 2010), soil moisture estimates (Forman et al., 2012; Meier et al., 2011; Renzullo et al., 2014; Alvarez-Garreton et al., 2014) or hydraulic information (Bailey and Bau, 2012), but also to update radar forcing (Harader et al., 2012; Kim and Yoo, 2014). Many applications also use in situ observations such as catchment discharge, snowpack measurements, or soil moisture to update models (e.g., Seo et al., 2009; Clark et al., 2008; Thirel et al., 2010; DeChant and Moradkhani, 2011; Franz et al., 2014).

* Corresponding author.
E-mail address: antoine.thiboult.1@ulaval.ca (A. Thiboult).

http://dx.doi.org/10.1016/j.jhydrol.2015.09.036
0022-1694/

In addition, DA may be coupled with parameter optimization (Vrugt et al., 2005; Moradkhani et al., 2005; Nie et al., 2011).

Sequential DA techniques such as particle filter and the Kalman filter family are frequently used for recursive updating of the states of a system, each time an observation is made available. Among them, the Ensemble Kalman filter (EnKF, Evensen, 1994) proved to be a powerful tool for hydrological forecasting (DeChant and Moradkhani, 2012; Rakovec et al., 2012; Vrugt and Robinson, 2007; Weerts and El Serafy, 2006; Abaza et al., 2014) that is effective and reliable enough for operational use (Andreadis and Lettenmaier, 2006). Several studies claim that they developed techniques that improved upon traditional EnKF (e.g., Clark et al., 2008; Whitaker and Hamill, 2002) by focusing on the relaxation of constrains of traditional EnKF implementation, or by explicitly including time lag between the soil moisture and the discharge in the updating process (Li et al., 2013, 2014; McMillan et al., 2013).

A key feature of EnKF is the proper specification of hyper-parameters (perturbations of inputs and outputs) and model states to be updated (Moradkhani et al., 2005). In most studies, EnKF implementation is based on an a priori selection of the hyper-parameters and updated states combination, which is then scarcely justified. Noteworthy exceptions are Moradkhani et al. (2005) and Chen et al. (2013), but these studies are very specific as they are performed on a single model and one or two catchments. Accurate perturbations representing error estimates are crucial since the EnKF updating scheme is based on the weighting of the model and observation relative error. However this specification is complex in practice as the different sources of uncertainty experience strong interactions (Moradkhani et al., 2006; Hong et al., 2006; Kuczera et al., 2006). Several attempts to account explicitly for structural error have been reported, for example by directly adding perturbations to the state variables (Reichle et al., 2002; Vrugt et al., 2006; Clark et al., 2008), or by updating model parameters (Moradkhani et al., 2005; Vrugt et al., 2005; Naevdal et al., 2003).

Moreover, despite encouraging results, DeChant and Moradkhani (2012) point that little research has been done to examine the effectiveness and robustness of EnKF and that “studies need to provide a more rigorous testing of these techniques than has previously been presented”. Another issue that needs consideration is that EnKF performance is mostly discussed as ‘standalone’, regardless of the influence of the coupling with the hydrological model. This is mainly due to the fact that EnKF is often tested on a single model. Thus, the question of adequacy between the DA technique and the model is rarely assessed.

The present study aims at identifying EnKF parametrization to reduce and quantify optimally the uncertainty related to initial conditions in a forecast mode. A second scope addresses the question of EnKF and hydrological model adequacy. In order to achieve this, the analysis is conducted on 20 structurally dissimilar lumped conceptual models, 38 catchments, 12 hyper-parameter sets, and all possible combinations of the state variables to strive for general results. Finally, the effectiveness of identifying the best EnKF parametrization without exploring all combinations is discussed.

Section 2 presents EnKF’s basics, models, basins and scores. Section 3 presents the results of the DA techniques followed by a discussion and the conclusion statements are provided in Section 4.

2. Material and methods

2.1. Hydrological models, snowmelt modules, and PET

The EnKF is tested individually on 20 lumped conceptual models, which differ by their structure. The selection was initially carried out by Perrin (2000) and revised by Seiller et al. (2012) for hydrological projection purposes. Because they are based on diverse hydrological concepts and present different degrees of complexity (4–10 calibrated parameters and 2–7 reservoirs to represent perceptual and conceptual hydrologic processes), they allow to test the EnKF in a comprehensive manner according to structure diversity (see Table 1). The models have been modified to match a common frame and they should not be directly compared to their original version. In the case where the original models included a module to compute evapotranspiration or snow accumulation and melting, the module has been omitted as these processes are computed externally beforehand.

The models exploit various conceptualizations and thus their parameters and state variables perform particular roles in simulating rainfall–runoff processes. Their reservoirs may describe systems ranging from precipitation interception to routing (or more conceptual functions). The role of state variables is not detailed in the article for concision purpose. For the same reason, the state variable values before and after the analysis step will not be discussed here but only the outputs of the models, i.e., simulated streamflow will be considered. For further details on state variable meaning, refer to Perrin (2000).

The lumped models are driven by potential evapotranspiration and precipitation. The potential evapotranspiration is computed from the formula proposed by Oudin et al. (2005), which relies on mean air temperature and the calculated extraterrestrial radiation. To partition snow accumulation, snowmelt, and liquid precipitation, the snow module (Cemaneige, Valery et al., 2014) is executed before hydrological models.

Table 1
Main characteristics of the 20 lumped models (Seiller et al., 2012).

Model acronym	Number of optimized parameters	Number of reservoirs	Derived from
M01	6	3	BUCKET (Thornthwaite and Mather, 1955)
M02	9	2	CEQUEAU (Girard et al., 1972)
M03	6	3	CREC (Cormary and Guilbot, 1973)
M04	6	3	GARDENIA (Thiery, 1982)
M05	4	2	GR4J (Perrin et al., 2003)
M06	9	3	HBV (Bergström and Forsman, 1973)
M07	6	5	HYMOD (Wagener et al., 2001)
M08	7	3	IHACRES (Jakeman et al., 1990)
M09	7	4	MARTINE (Mazenc et al., 1984)
M10	7	2	MOHYSE (Fortin and Turcotte, 2007)
M11	6	4	MORDOR (Garçon, 1999)
M12	10	7	NAM (Nielsen and Hansen, 1973)
M13	8	4	PDM (Moore and Clarke, 1981)
M14	9	5	SACRAMENTO (Burnash et al., 1973)
M15	8	3	SIMHYD (Chiew et al., 2002)
M16	8	3	SMAR (O'Connell et al., 1970)
M17	7	4	TANK (Sugawara, 1979)
M18	7	3	TOPMODEL (Beven et al., 1984)
M19	8	3	WAGENINGEN (Warmerdam et al., 1997)
M20	8	4	XINANJIANG (Zhao et al., 1980)