Journal of Hydrology 519 (2014) 2692–2706

Contents lists available at ScienceDirect

Journal of Hydrology

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

Sequential streamflow assimilation for short-term hydrological ensemble forecasting

Mabrouk Abaza [a,*], François Anctil [a], Vincent Fortin [b], Richard Turcotte [c]

[a] Department of Civil and Water Engineering, Pavillon Adrien-Pouliot, 1065, avenue de la Médecine, Université Laval, Québec, Canada
[b] Recherche en prévision numérique environnementale, Environnement Canada, Montréal, 2121, route Transcanadienne Dorval, Québec H9P 1J3, Canada
[c] Centre d'Expertise Hydrique du Québec, 675, boul. René-Levesque Est Aile Louis-Alexandre-Taschereau, 4[e] étage, Québec G1R5V7, Canada

ARTICLE INFO

Article history:
Available online 23 August 2014

Keywords:
Ensemble streamflow forecasting
Streamflow assimilation
Ensemble Kalman Filter
Performance
Reliability
Economic value

SUMMARY

This paper evaluates the application of the Ensemble Kalman Filter (EnKF) for streamflow assimilation within an ensemble prediction system designed for short-term hydrological forecasting at the outlet of the au Saumon watershed. The EnKF updates three state variables of a distributed hydrological model (soil moisture in the intermediate layer, soil moisture in the deep layer, and land routing) to improve the initial conditions of the forecasts. A systematic method for the identification of the perturbation factors (ensemble generation) and for the selection of the ensemble size is discussed. EnKF results show a substantial improvement in performance and reliability over the open-loop estimates. Manual assimilation was also assessed and led to a performance similar to the EnKF; however, the EnKF forecasts are substantially more reliable. While an ensemble size of 1000 members was required to fully sample the hydrological and meteorological uncertainty, similar results are obtained in terms of skill when limiting the ensemble size to 50.

1. Introduction

Schaake et al. (2007) suggested that significant advances in operational hydrological forecasting could be made by developing forecasting systems designed to process and disseminate probabilistic information about the uncertainty of hydrological forecasts. The deterministic approach to streamflow forecasting provides no information on the uncertainty of future events, justifying the need for a probabilistic one (Ehrendorfer, 1997). Meteorological ensemble prediction systems (M-EPS) have proven capable of increasing the forecast horizon and of providing an estimate of the uncertainty associated with each forecast. The objective is simple: to make available a set of forecasts at each time step, so that this ensemble allows the users to assess the uncertainty of the issued forecast. Hydrological ensemble prediction systems (H-EPS) share the same purpose but for streamflow (Palmer, 2002); they dynamically allow the spread of the ensemble to change with the information content of the forecast.

The skill of short-range forecasts from a numerical model depends largely on how the state variables of the model are initialized (DeChant and Moradkhani, 2011a,b). Data assimilation is an invaluable tool as it allows obtaining improved predictions after combining data to a numerical model (Liu and Gupta, 2007). Data assimilation techniques have been used for many years in meteorology (Rabier et al., 2000; Gauthier et al., 2007; Rawlins et al., 2007; Fillion et al., 2010; Tanguay et al., 2012) but similar developments are more recent in hydrology (Liu and Gupta, 2007; Liu et al., 2012; Dechant and Moradkhani, 2011a; Thirel et al., 2013; McMillan et al., 2013). There are nonetheless many available data assimilation techniques suitable for hydrological applications, such as sequential or variational methods, as extensively reviewed by Liu and Gupta (2007), Reichle (2008) and Liu et al. (2012). They express the need for assimilation techniques that improve forecasts and reduce uncertainty. Most of data assimilation systems developed for hydrological models are concerned with streamflow assimilation (Seo et al., 2003; Clark et al., 2008; Warrach-Sagi and Wulfmeyer, 2010; Lee et al., 2011, 2012) while others explored soil moisture (Pauwels et al., 2001; Reichle et al., 2002, 2008; De Lannoy et al., 2007; Moradkhani and Sorooshian, 2008; Crow and Ryu, 2009; Kumar et al., 2009; Brocca et al., 2010; Peters-Lidard et al., 2011; Montzka et al., 2011; Liu et al., 2012) and snow (Rodell and Houser, 2004; Lee et al., 2005; Andreadis and Lettenmaier, 2006; Liston and Hiemstra, 2008; Zaitchik et al., 2008; Durand et al., 2009; Kuchment et al., 2010;

* Corresponding author.
E-mail addresses: mabrouk.abaza.1@ulaval.ca (M. Abaza), Francois.Anctil@gci.ulaval.ca (F. Anctil), Vincent.Fortin@ec.gc.ca (V. Fortin), richard.turcotte2@mddep.gouv.qc.ca (R. Turcotte).

http://dx.doi.org/10.1016/j.jhydrol.2014.08.038
0022-1694/

Kolberg and Gottschalk, 2010; DeChant and Moradkhani, 2011a; De Lannoy et al., 2012).

Manual state updating, in which an operational hydrologist adjusts the model's inputs or state variables over a predetermined number of previous time steps in order for the model to better simulate the observed discharge, is also another viable but human resources intensive option. It has been used successfully with distributed hydrological modeling and ensemble forecasting (Boucher et al., 2012; Abaza et al., 2013). The need for sound assimilation techniques for operational distributed hydrological models is stressed by Weerts et al. (2014) and Liu et al. (2012), among others.

There exists a wide range of data assimilation techniques suitable for hydrological applications, namely variational (Seo et al., 2003, 2009) and sequential methods (e.g. the Ensemble Kalman Filter (EnKF) developed by Evensen (2003) and particle filters as in Vrugt et al. (2013) and Moradkhani et al. (2012)) and methods based on evolutionary algorithms (Dumedah, 2012; Dumedah and Coulibaly, 2013a,b). A certain consensus has emerged in favor of Ensemble Kalman and particle filters, especially in the presence of strong nonlinearities (Pham, 2001).

Case studies built around a distributed hydrological model are less numerous than for lumped ones; we note EnKF and variational streamflow assimilation (Clark et al., 2008; McMillan et al., 2013; Lee et al., 2012; Thirel et al., 2010) and variational soil moisture and streamflow assimilation (Lee et al., 2011). Data assimilation was also developed for other applications such as groundwater (Valstar et al., 2004; Franssen et al., 2011), coupled surface–subsurface (Camporese et al., 2009), and sediment transport (Stroud et al., 2009).

Most authors select updating the states variables of their model, but some simultaneously update states and parameters (Moradkhani et al., 2005a,b; Vrugt et al., 2005; Franssen and Kinzelbach, 2008; Lu et al., 2011; Leisenring and Moradkhani, 2011; Nie et al., 2011).

Weerts and El Serafy (2006) compared EnKF, particle filter (PF), and residual resampling performance. They found that the EnKF was more efficient and, in general, more robust than the other methods. Dumedah and Coulibaly (2013a,b) arrived to a different conclusion comparing the EnKF, the particle filter, and an evolutionary-based assimilation. They concluded that the evolutionary-based assimilation, which they proposed, surpasses the others for lead times of 10 days or less, and that the particle filter performs best for longer lead times.

The EnKF was introduced by Evensen (1994) as an alternative to the extended Kalman filter (EKF), addressing difficulties arising from high-dimensional nonlinear filtering problems. According to Komma et al. (2008), the EnKF is considered as an obvious choice for flood forecasts updating (Komma et al., 2008). Its advantage resides in its computational efficiency and in its straightforward implementation for lumped and distributed hydrological models (Pauwels and De Lannoy, 2009). Note that there exist many variants of the EnKF, including the ensemble adjustment Kalman filter (Anderson, 2001), the ensemble square root filter (Tippett et al., 2003; Clark et al., 2008), and the Ensemble Kalman smoother (Li et al., 2013), which are not explored here.

The specific purpose of this study is to explore the capabilities of the EnKF for the sequential streamflow assimilation of an ensemble prediction system for short-term hydrological forecasting. More specifically, the EnKF, which is among the most popular methods in meteorology, is setup for the state updating of Hydrotel, a semi-physical distributed model operationally used in the Province of Québec (Canada) to issue forecasts for watersheds with quick hydrological responses, located upstream of dams (Turcotte et al., 2004). Even though the selected watershed is dominated by snow accumulation and melt, the study focuses on a rainfall-runoff implementation (summer and autumn) to produce ensemble hydrological forecasts. The freshet period will remain unexplored here since it may require a different implementation combining streamflow and snow information.

The watershed, database, model, and assimilation technique are defined in the Section 2, while the experimental procedure is detailed in Section 3. Results for hyper-parameters, ensemble streamflow forecasting, ensembles with a lower number of members, and cost-loss analysis are detailed in Section 4. Conclusions are presented in Section 5.

2. Material and methods

2.1. Watershed and data

All model simulations are performed on the au Saumon watershed of the upper Saint-Francois River in the Province of Québec (Canada): a snow-influenced basin subjected to contrasted climate conditions (Seiller et al., 2012).

The roughly 80-km au Saumon River drains 767 km^2 of mostly forested land at gauge 030282 located in the municipality of Lingwick. Elevations within the watershed vary between 277 and 1092 m. The average annual air temperature is 4.5 °C. Mean annual precipitation reaches about 1250 mm, of which one third is snow, which leads to a mean annual streamflow of roughly 750 mm. The hydrological regime of the au Saumon is dominated by a spring freshet and high fall flows. The main characteristics of the watershed (soil texture, land cover and location of the gauge station) are illustrated by Fig. 1.

Model simulations are performed at a 3 h time step. Precipitation and temperature observations extend from August 2010 to August 2011. All observations and model calibration were provided by the Centre d'Expertise Hydrique du Québec (CEHQ), including 3-h streamflow time series at the outlet of the watershed and 3-h rainfall and temperature time series krigged to a 0.1° resolution grid (about 10 by 15 km) encompassing the river system. Station data from the Québec «Programme de surveillance du climat du Ministère du Développement durable, de l'Environnement, de la Faune et des Parcs (MDDEFP)» were used as original data source. Simulation with assimilation is performed from August 2010 to August 2011, but the hydrological ensemble forecasts were only issued on the rainfall-runoff period (summer and autumn). The Nash coefficient over the calibration period (2000–2010) reached 0.59, but is 0.77 when only considering the period used here for the assessment of the EnKF.

The Canadian Operational Global Meteorological Ensemble Prediction System (M-GEPS) is used to produce hydrological ensemble forecasts. This meteorological forecasting system relies on the GEM model for issuing 20-member predictions with a 3-h time step at a 100-km resolution (at mid-latitudes) and a 15-day forecast horizon. GEM is a versatile operational model that can be implemented over a wide range of spatial scales and for a variety of meteorological applications (Toth et al., 2010). It is currently used operationally to produce the background fields in the global data assimilation cycle and in prediction mode to produce medium-range forecasts (Côté et al., 2003). All meteorological forecasts are issued at midnight and bi-linearly interpolated to the 0.1° resolution grid of the CEHQ observations. They are not accessible for the entire simulation period mentioned above; 256 days are available to the project. 160 days taken from the autumn of 2010 and the summer of 2011 are actually used.

No pre-processing of the meteorological forcing was performed, in contrast with other H-EPS implementations, because there is no long reforecast database available for the Canadian M-EPS, making it very challenging to assess and correct biases. Also, since the main objective of the paper is to establish whether or not the EnKF