Journal of Hydrology 519 (2014) 2676–2691

Contents lists available at ScienceDirect

Journal of Hydrology

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

Impact of streamflow data assimilation and length of the verification period on the quality of short-term ensemble hydrologic forecasts

A. Randrianasolo [a], G. Thirel [a,*], M.H. Ramos [a], E. Martin [b]

[a] Irstea, UR HBAN, 1, rue Pierre-Gilles de Gennes, CS 10030, 92761 Antony, France
[b] CNRM-GAME, Météo-France, CNRS, Toulouse, France

ARTICLE INFO

Article history:
Available online 3 October 2014

Keywords:
Data assimilation
Forecast verification
Forecast quality
Hydrologic ensemble prediction

SUMMARY

Data assimilation has gained wide recognition in hydrologic forecasting due mainly to its capacity to improve the quality of short-term forecasts. In this study, a comparative analysis is conducted to assess the impact of discharge data assimilation on the quality of streamflow forecasts issued by two different modeling conceptualizations of catchment response. The sensitivity of the performance metrics to the length of the verification period is also investigated. The hydrological modeling approaches are: the coupled physically-based hydro-meteorological model SAFRAN–ISBA–MODCOU, a distributed model with a data assimilation procedure that uses streamflow measurements to assess the initial state of soil water content that optimizes discharge simulations, and the lumped soil moisture-accounting type rainfall–runoff model GRP, which assimilates directly the last observed discharge to update the state of the routing store. The models are driven by the weather ensemble prediction system PEARP of Météo-France, which is based on the global spectral ARPEGE model zoomed over France. It runs 11 perturbed members for a forecast range of 60 h. Forecast and observed data are available for 86 catchments over a 17-month period (March 2005–July 2006) for both models and for 82 catchments over a 52-month period (April 2005–July 2009) for the GRP model. The first dataset is used to investigate the impact of streamflow data assimilation on forecast quality, while the second is used to evaluate the impact of the length of the verification period on the assessment of forecast quality. Forecasts are compared to daily observed discharges and scores are computed for lead times 24 h and 48 h. Results indicate an overall good performance of both hydrological models forced by the PEARP ensemble predictions when the models are run with their data assimilation procedures. In general, when data assimilation is performed, the quality of the forecasts increases: median differences between performance with and without data assimilation are of the order of 16%, varying with catchment, lead time and according to the metric used to evaluate the forecasts. We also show that, for the configuration studied here, forecast verification carried out over a period greater than 24 months of daily discharges provides average score values already close to the final values that would be obtained with a longer verification period.

1. Introduction

Improving flood forecasting is a key issue of increasing importance in hydrology. It is motivated mainly by the need to lower the human and financial losses caused by severe and/or sudden overflows of watercourses. Liu and Gupta (2007) indicate three ways to explore how to improve the performance of hydrologic forecasting systems: improving hydrological modeling, strengthening instrumentation for better quality of hydrologic information and assimilating data for improving the initial states of the model. In this study, we focus on the last point. In general, data assimilation (DA) in hydrologic forecasting refers to any technique that aims to assimilate observations into a modeling system in order to improve the initial states of real-time forecasting systems and, consequently, their predictions. In the assimilation process, modifications are brought to the model with changes to its input or output variables, parameters or states (Refsgaard, 1997). In the review of DA techniques used for operational forecasting proposed by Liu et al. (2012), the authors show that despite the numerous studies carried out on the topic there is still a gap between theory and operational practice. There is a clear need for new data sources in hydrologic DA and for automated procedures to move forward in improving the predictive skill of hydrologic forecasts. Complementarily, the authors also point out the need for robust forecast

* Corresponding author.
E-mail address: guillaume.thirel@irstea.fr (G. Thirel).

http://dx.doi.org/10.1016/j.jhydrol.2014.09.032
0022-1694/

verification that, among others, will "help quantify the value of DA for operational hydrologic forecasting".

A robust verification of the impacts of DA on the quality of hydrologic predictions for operational forecasting is not an easy task. It requires the analysis of different combinations of DA techniques and modeling approaches, implemented on a variety of hydro-climatic conditions and running over a long time period. Usually, the impact of DA on the quality of forecasts is assessed at specific sites and for specific configurations, when new techniques are developed, although not necessarily operationally implemented. Intercomparison studies providing generalized conclusions on the impacts of DA on hydrologic verification are rare and may be more complex than they appear at first sight. Studies on the literature have however shown the potential positive impacts of automated DA techniques over different applications on operational forecasting. Moradkhani et al. (2006), for instance, highlight the improvements obtained in the quality of streamflow forecasts after using a sequential data assimilation technique based on particle filtering to propagate input errors through a conceptual hydrologic model. Seo et al. (2009) describe the improvements made through the use of a state updating procedure based on variational assimilation of streamflow, precipitation and potential evaporation data into lumped and routing models in operational river stage forecasting of fast-responding headwater catchments in the USA. The authors also compared automated DA performance with manual updating done by forecasters, highlighting the importance of automated DA when human expertise is not available as well as the complementarities between automated and manual DA. DeChant and Moradkhani (2011) explore the use of ensemble data assimilation to better estimate snow quantities and create a framework to account for initial condition uncertainty in addition to forcing uncertainty. Their results suggest that data assimilation may also improve ensemble streamflow volume forecasts.

This paper aims to assess the impact of streamflow data assimilation on the quality of ensemble short-term hydrologic forecasts issued by two different configurations of combined DA technique and hydrological model, both developed in France for operational flood forecasting. In a previous study (Randrianasolo et al., 2010), we investigated the performance of two hydrologic models and showed that both provided a good overall performance for hydrological ensemble prediction. Preliminary tests were carried out to evaluate the importance of data assimilation, but the study pointed out the need of further analyses intercomparing both modeling systems with and without data assimilation and over a longer verification period to achieve more robust conclusions. These analyses are presented in this paper. Additionally, the diagnostic verification carried out here also examines the quality of streamflow forecasts in relation to catchment characteristics and flow regimes.

Since most metrics used to assess the quality of ensemble forecasts are based on average values over time series of pairs of observations and forecasts (see, for instance, Brown et al., 2010), this study also evaluates the impact of the length of the verification period on forecast quality assessment. To the best knowledge of the authors, almost nothing has been written in the literature on this subject so far, whereas many hydrologists have studied the impact of the number of data available for the calibration of hydrological models on parameter estimation and model robustness during validation (Gupta and Sorooshian, 1985). For instance, Yapo et al. (1996) compared calibration results using different lengths of data, taken from different sections of historical archives, and concluded that eight years of data are required to get insensitive calibration. Studying the HBV model, Harlin (1991) highlighted a rule which consisted in having diverse climatic and flow conditions in a data set to get the right representation of their natural variability during the calibration procedure. Brath et al. (2004) tested different scenarios of historical records to calibrate a distributed rainfall–runoff model in order to get an efficient parameterization. They obtained acceptable results with only three months of data and an adequately extended number of rain gauges. In general, in forecast verification, as in hydrologic model calibration, the conventional view is to consider that 'the longer, the better' (Perrin et al., 2007). In this study, we explore a 52-month archive of reforecast data to evaluate if this is long enough to assess the quality of daily ensemble streamflow forecasts. The aim is to investigate the questions: which length of data is needed to properly assess the quality of ensemble forecasts? What is the influence of the length of data on verification scores and, consequently, on the conclusions that can be drawn from the quality assessment?

In the following, Section 2 presents the hydrologic models used and their data assimilation techniques. Section 3 describes the data used (weather forecasts, observed data and catchments), while Section 4 presents the methods applied for the evaluation of forecast quality. Section 5 presents the results obtained when models are used with and without DA, as well as the impact of the length of the data on the assessment of forecast quality. Finally, Section 6 presents the discussions and conclusions.

2. Forecasting models and their data assimilation techniques

2.1. A lumped rainfall–runoff model (GRP)

GRP is a lumped soil-moisture-accounting type rainfall–runoff model designed specifically for flood forecasting (see Perrin et al., 2003 and Berthet et al., 2009 for details). Input data are precipitation and mean evapotranspiration (Oudin et al., 2005). The model links a production function, which computes the effective rainfall over the catchment, to a routing function, which transfers the flow toward the catchment outlet. The production function is a storage that reflects the memory of the moisture in the catchment and loses water for the routing function through percolation. The routing function includes a unit hydrograph and a non-linear routing store, which delays the release of the effective precipitation over the next time step. In this study, the model is calibrated and run at the daily time-step. Three parameters have to be calibrated against observed data: a volume-adjustment factor that controls the volume of the effective rainfall, the maximum capacity of the routing store and the base time of the unit hydrograph. The model was calibrated with observed data available prior to the period chosen for forecast and evaluation (see Section 4) in order to guarantee independence between calibration and evaluation. It was calibrated with RMSE as objective function, i.e., automated model calibration was performed to identify parameter sets that minimize the Root Mean Square Error.

2.2. A distributed hydro-meteorological coupled model (SIM)

SIM (SAFRAN–ISBA–MODCOU) is a distributed hydrometeorological model developed at Météo-France, validated over 881 French catchments and able to reproduce water and energy budgets. It is composed of three models: SAFRAN, ISBA and MODCOU (Habets et al., 2008). SAFRAN (Quintana-Seguí et al., 2008) is a system which produces an analysis on an 8 km $\times$ 8 km grid of observed meteorological parameters: 10-m wind speed, 2-m relative humidity, 2-m air temperature, total cloud cover, incoming solar radiation, atmospheric/terrestrial radiations, snowfall and rainfall. It generates mean values of pressure, radiation, wind and humidity for the ISBA model, which is a land-surface model that simulates the water and energy fluxes between the soil and the atmosphere (Noilhan and Planton, 1989). ISBA is a soil–vegetation–atmosphere-transfer model and includes the use of an exponential profile of hydraulic conductivity (Quintana-Seguí et al., 2009). Through the simulation