Advances in Water Resources 53 (2013) 231–241

Contents lists available at SciVerse ScienceDirect

Advances in Water Resources

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

Evolutionary assimilation of streamflow in distributed hydrologic modeling using in-situ soil moisture data

Gift Dumedah [a,b,*], Paulin Coulibaly [a]

[a] School of Geography and Earth Sciences, and Department of Civil Engineering, McMaster University, 1280 Main Street West, Hamilton, Ontario, Canada L8S4L8
[b] Department of Civil Engineering, Monash University, Melbourne, Victoria 3800, Australia

ARTICLE INFO

Article history:
Received 13 May 2011
Received in revised form 4 December 2011
Accepted 6 July 2012
Available online 15 July 2012

Keywords:
Data assimilation
Evolutionary algorithms
Soil moisture
Streamflow
SWAT

ABSTRACT

This study has applied evolutionary algorithm to address the data assimilation problem in a distributed hydrological model. The evolutionary data assimilation (EDA) method uses multi-objective evolutionary strategy to continuously evolve ensemble of model states and parameter sets where it adaptively determines the model error and the penalty function for different assimilation time steps. The assimilation was determined by applying the penalty function to merge background information (i.e., model forecast) with perturbed observation data. The assimilation was based on updated estimates of the model state and its parameterizations, and was complemented by a continuous evolution of competitive solutions.

The EDA was illustrated in an integrated assimilation approach to estimate model state using soil moisture, which in turn was incorporated into the soil and water assessment tool (SWAT) to assimilate streamflow. Soil moisture was independently assimilated to allow estimation of its model error, where the estimated model state was integrated into SWAT to determine background streamflow information before they are merged with perturbed observation data. Application of the EDA in Spencer Creek watershed in southern Ontario, Canada generates a time series of soil moisture and streamflow. Evaluation of soil moisture and streamflow assimilation results demonstrates the capability of the EDA to simultaneously estimate model state and parameterizations for real-time forecasting operations. The results show improvement in both streamflow and soil moisture estimates when compared to open-loop simulation, and a close matching between the background and the assimilation illustrates the forecasting performance of the EDA approach.

1. Introduction

An important state variable which is used in hydrological streamflow forecasts is soil moisture. Soil moisture partitions available rainfall into runoff and infiltration, and influences the exchange of mass and energy between the land surface and the atmosphere. Numerous studies [38,8,52,48,7] have identified soil moisture as a crucial input to facilitate streamflow estimation, and others [14,13,26,5] have incorporated soil moisture data into hydrological models to estimate streamflow using data assimilation (DA). DA provides an innovative approach to quantify uncertainties in both observation and hydrological model, with the aim to merge inaccurate model output with imperfect observation data [52,22,39].

But challenges remain in integrating in situ soil moisture into hydrological models for streamflow forecasts. In particular, accurate estimation of the model state, the dynamics between simulated output and observation, the model error, and accurate model parameterizations are key drivers for efficient DA methods. A persistent question is a framework which could simultaneously propagate model state and provide accurate model parameterizations while finding a merger between simulated outputs and the observation dataset.

Studies have applied several DA methods including the Kalman filter [5,25], Extended Kalman filter (EKF) [51], the ensemble Kalman filter (EnKF) [66,15,22], the variational DA [18,17,16,47,49], and particle filtering [54,53,43] to assimilate data into hydrological models. Analytical weakness of the Kalman filter including its incompatibility with nonlinear models, and the problems of instability in strong nonlinear models are well documented in the hydrological data assimilation literature [2,22,51,50].

In the EnKF, ensemble of model states are usually generated by randomly changing/perturbing model input values. These model states are propagated to future time and are updated using the Kalman gain (denoted, K) matrix. A key limitation in the EnKF is that its probability distributions are assumed to be Gaussian as in the EKF and variational methods [22,43,20]. The variational method minimizes a cost (or a penalty) function (denoted, J) by finding the least squares estimate based on the observation dataset, the model estimate, and their associated uncertainties. Particle filters (PFs) are based on recursive Bayesian filter which use Monte Carlo

* Corresponding author at: Department of Civil Engineering, Monash University, Melbourne, Victoria 3800, Australia.

E-mail address: gift.dumedah@monash.edu (G. Dumedah).

0309-1708/$ - see front matter © 2012 Published by Elsevier Ltd.
http://dx.doi.org/10.1016/j.advwatres.2012.07.012

simulation. PFs require posterior probability distribution function which is usually generated from a random sample (or particles) with associated weights where updates are performed on particle weights. Model errors in the PF method do not assume normality but particle weights can quickly descend to zero after few iterations [20].

Typically DA studies [54–56,66,42,22,62,45] have used observed streamflow data to update model state where there is lack of observed state data (e.g. soil moisture). Few studies including [59,60,5] have actually used observed state data (i.e. soil moisture) in model state update process to improve streamflow estimation. The above studies usually lump different error sources into a single model [63] by estimating the K or by minimizing the J in the model state update process. These studies assimilate soil moisture into a hydrological model where model parameterizations are assumed to be constant. But it is important to incorporate other error sources in the DA procedure including model structure, model parameterizations, and inaccuracies in input data (e.g. precipitation) and observation datasets [39,62,64].

This study uses an alternative DA method based on multi-objective evolutionary algorithms (MOEAs), which it is termed in this paper as evolutionary data assimilation (EDA). MOEAs are stochactic search tools which employ the concept of evolution and natural selection. Evolutionary algorithms possess stochastic [29,31,44,58] and adaptive [6,58] capabilities, and can provide information for estimating model state and observation-simulation dynamics which are critical DA methods. This study assimilates in situ soil moisture into a distributed hydrological model, and allows state parameterizations of the model where they are applied to simulate ensemble of streamflows. As will be demonstrated in this study, the updating procedure is aimed to better estimate the model state, and parameterizations of the hydrological model before merging imperfect simulated output to uncertain observation data. The updating process is supplemented by a continuous evolution of equally competitive solutions over multiple time steps which approximates the dynamics between simulated output and observation, and are applied to estimate model state at model forecasting stage.

The EDA is illustrated for the Spencer Creek watershed in southern Ontario using an advanced and widely used MOEA tool, the non-dominated sorting genetic algorithm-II (NSGA-II) which was developed by [28]. The watershed has an hourly in situ soil moisture data from the time domain reflectometer (TDR) probe measurements and stream gauge data for 4 different locations. Simulated output of soil moisture and streamflow are generated using the soil and water assessment tool (SWAT).

The remainder of the paper is organized as follows. The data and methods section describe the study area and data sources used, the watershed model (i.e., SWAT), and the NSGA-II method. A detailed description of the EDA framework and its implementation for soil moisture and streamflow assimilations into SWAT are also provided in this section. The results and discussion section presents outputs of the method to illustrate the effectiveness of the EDA approach. The paper concludes with findings on how joint updates of model state and parameterizations of SWAT have improved streamflow estimations, and its capability to estimate antecedent soil moisture for future streamflow forecasts.

2. Data and methods

2.1. Study area – Spencer Creek watershed, and data sources

The Spencer Creek watershed is a sub-catchment located at the western side of Lake Ontario in southern Ontario (Fig. 1). It runs from north to south and has a drainage area of about 280 km^2. The upstream section is mostly a plain physiographic region with shallow soils, isolated drumlins, gentle-sloping bedrocks, and few swamps and wetlands. The middle section has massive deposits of fine-grained sediments such as silt and clay, swamps and wetlands where the streams have low gradient and variable groundwater discharge. The Spencer Creek watershed covers parts of the Niagara Escarpment at its lower portion where topography is highly variable. As a result, streams at the lower section have steep gradient and increased groundwater discharge. The watershed also has two reservoirs: Valens reservoir located at the upstream section, and Christie reservoir located at the middle section.

Major land uses and their respective proportions in the watershed include agriculture (33%), forest (32%), wetlands and water bodies (18%), and urban (17%). The distribution of the various soil textures are: 15% for clay, 19% for fine sandy loam, 20% for gravel, 32% for organic, and 14% for impervious surfaces (including residential, industrial, and road networks). The watershed has a dendrite stream pattern with elevation ranging between 69 m to 340 m, and about 90% of the watershed has slope less than 10 degrees, and 9% has slope greater than 10 degrees.

The watershed has four streamflow gauging stations: Westover, Highway 5, Dundas, and Ancaster, but only data at the first three locations are used as data at Ancaster has large missing values. Streamflow at these locations from 2006 to 2010 vary widely from year to year where maximum and minimum discharges are 30.76 m^3/s at Dundas on 14 March 2010 and 0.015 m^3/s at Ancaster on 24 September 2007 respectively. The annual daily discharge is 1.34 m^3/s with standard deviation of 0.13 m^3/s. Extreme high discharges are occasionally observed in January due to sudden high temperatures which result in snowmelt floods, whereas low discharges occur usually in September. But overall maximum discharges usually occur in March and April as snowmelt floods which are due to the combined influence of accumulated snow and consistent high temperatures during the spring season.

Additionally, the Spencer Creek watershed has two in situ soil moisture sites: Orchard and Governor Road both at Dundas Valley (Fig. 1). The two soil moisture sites are part of the McMaster Mesonet – a network of continuous soil moisture monitoring sites and weather stations which are established by the McMaster University (School of Geography & Earth Sciences) in the Halton–Hamilton area. Each soil moisture site has nine soil moisture stations where each station has Campbell Scientific TDR soil probes (CS610) at six depths: 10 cm, 20 cm, 30 cm, 50 cm, 70 cm, and 100 cm. The soil moisture data were observed at hourly time intervals since November 2006 to present. The average and the hourly variance of the daily soil moisture data at Orchard and Governor Road from 2006 to 2010 were used in this study. The daily soil moisture error is approximated to be the hourly variance between the nine soil moisture stations at each site. Two weather stations are also located in the Spencer Creek watershed, and provide hourly meteorological data including precipitation, air temperature, solar radiation, relative humidity, wind speed, and evapotranspiration.

2.2. Watershed model: soil and water assessment tool

The soil and water assessment tool (SWAT) is a basin-scale continuous-time step model that was developed by the United States Department of Agriculture in the early 1990s and has been continuously updated. The model evaluates the impact of management and climate on water supplies in watersheds and large river basins [4,3]. SWAT is physically-based, spatially distributed model and requires extensive input data such as soil, land cover and terrain information, and can operate at sub-daily time intervals. SWAT divides the watershed into sub-basins which are further subdivided into hydrologic response units (HRUs) using digital elevation model. The HRUs are determined based on homogenous land areas within sub-basins which are categorized by unique land use, soil, and slope information.