Journal of Hydrology 539 (2016) 57–73

Contents lists available at ScienceDirect

Journal of Hydrology

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

A wavelet-based non-linear autoregressive with exogenous inputs (WNARX) dynamic neural network model for real-time flood forecasting using satellite-based rainfall products

Trushnamayee Nanda [a,1], Bhabagrahi Sahoo [b,*], Harsh Beria [a], Chandranath Chatterjee [a,2]

[a] Agricultural and Food Engineering Department, Indian Institute of Technology Kharagpur, India
[b] School of Water Resources, Indian Institute of Technology Kharagpur, India

ARTICLE INFO

Article history:
Received 30 December 2015
Received in revised form 3 May 2016
Accepted 5 May 2016
Available online 10 May 2016
This manuscript was handled by Andras Bardossy, Editor-in-Chief, with the assistance of Fi-John Chang, Associate Editor

Keywords:
ANN
WANN
NARX
WNARX
Flood forecasting
TRMM

SUMMARY

Although flood forecasting and warning system is a very important non-structural measure in flood-prone river basins, poor raingauge network as well as unavailability of rainfall data in real-time could hinder its accuracy at different lead times. Conversely, since the real-time satellite-based rainfall products are now becoming available for the data-scarce regions, their integration with the data-driven models could be effectively used for real-time flood forecasting. To address these issues in operational streamflow forecasting, a new data-driven model, namely, the wavelet-based non-linear autoregressive with exogenous inputs (WNARX) is proposed and evaluated in comparison with four other data-driven models, viz., the linear autoregressive moving average with exogenous inputs (ARMAX), static artificial neural network (ANN), wavelet-based ANN (WANN), and dynamic nonlinear autoregressive with exogenous inputs (NARX) models. First, the quality of input rainfall products of Tropical Rainfall Measuring Mission Multi-satellite Precipitation Analysis (TMPA), viz., TRMM and TRMM-real-time (RT) rainfall products is assessed through statistical evaluation. The results reveal that the satellite rainfall products moderately correlate with the observed rainfall, with the gauge-adjusted TRMM product outperforming the real-time TRMM-RT product. The TRMM rainfall product better captures the ground observations up to 95 percentile range (30.11 mm/day), although the hit rate decreases for high rainfall intensity. The effect of antecedent rainfall (*AR*) and climate forecast system reanalysis (CFSR) temperature product on the catchment response is tested in all the developed models. The results reveal that, during real-time flow simulation, the satellite-based rainfall products generally perform worse than the gauge-based rainfall. Moreover, as compared to the existing models, the flow forecasting by the WNARX model is way better than the other four models studied herein with the TRMM and TRMM-RT rainfalls at 1–3 days lead times. The results confirm the robustness of the WNARX model with only the satellite-based (TRMM-RT) rainfall (without use of gauge data) to provide reasonably good real-time flood forecasts. The utility of the TRMM-RT solves the real-time flood forecasting issues, since this is the only rainfall product disseminated in real-time. Hence, the WNARX model with the TMPA rainfall products can offer an exciting new horizon to provide flood forecasting and early warning in the flood prone catchments.

1. Introduction

In a changing global climate, there is an increase in frequency of flood extremes worldwide (Trenberth, 1999; Petra and Naef, 2010; Jena et al., 2014; Ezer and Atkinson, 2014) in both urban and rural watersheds, increasing the severity of flood hazards manifolds. Hence, streamflow forecasting in real-time always becomes important for water resources management and flood risk analysis (Perumal and Sahoo, 2007; Perumal et al., 2011). Real-time flood monitoring requires dense gauge-based observed precipitation data, the dominant forcing component, which may not be available in real-time at the desired spatial and temporal resolutions. Particularly, in India and other developing countries, the sparse raingauge networks and data unavailability in remote areas are the prime obstacles.

* Corresponding author. Tel.: +91 3222 281884.
E-mail addresses: nanda.trushnamayee@yahoo.com (T. Nanda), bsahoo2003@yahoo.com (B. Sahoo), harsh.beria93@gmail.com (H. Beria), cchat@agfe.iitkgp.ernet.in (C. Chatterjee).
[1] Tel.: +91 9475083931.
[2] Tel.: +91 3222 283158.

http://dx.doi.org/10.1016/j.jhydrol.2016.05.014

In the recent past, many devastating floods occurred in India, such as, the 2005 Mumbai flood in the Mithi River, 2008 Bihar flood in the Kosi River, the 2013 cloudburst in Uttarakhand, the 2014 flood havocs in Jammu and Kashmir and the 2015 Chennai flood. Hence, a reliable real-time flood forecasting method to transform rainfalls into runoff is needed; for which the satellite-based real-time availability of rainfalls could be more useful. The high resolution Tropical Rainfall Measuring Mission (TRMM) – Multisatellite Precipitation Analysis (TMPA) gridded rainfall products (Huffman et al., 2007) offer a promising aid for real-time flood forecasting and warning system in data-scarce regions worldwide with varied accuracy levels, when compared with observed rainfalls. However, since the uncertainties involved with the discharge data are much smaller than those with the precipitation data, the precipitation products need to be evaluated through their utility in hydrological streamflow modeling (Lammers et al., 2001). In hydrological modeling, the potential of satellite rainfall data in comparison with the observed rainfalls has been evaluated by many researchers (Artan et al., 2007; Asante et al., 2007; Hopson and Webster, 2010; Yong et al., 2010, 2012; Kneis et al., 2014; Tong et al., 2014). Kneis et al. (2014) found that the real-time rainfall product from the TMPA project is less preferable for hydrological analysis as compared to its gauge-adjusted estimate. Hence, to reduce the rainfall bias while making it reliable for operational flood forecasting in real-time, there is a scope to further evaluate the real-time satellite-based rainfall product with different data-driven techniques (e.g., ASCE, 2000a,b; Coulibaly et al., 2000; Khu et al., 2001; Kisi, 2009; Napolitano et al., 2010; Tiwari and Chatterjee, 2010a,b, 2011). Moreover, during flood period, one of the difficult tasks is inflow forecasting for real-time reservoir operation to protect the dam from failure and to reduce the magnitude of the downstream flood and to increase the time of translation (Coulibaly et al., 2000; Mohammadi et al., 2005; Valipour et al., 2013). Hence, in the situation where the real-time high resolution satellite precipitation products are available, the neural network (NN)-based flood forecasting at higher lead times can become more reliable (Akhtar et al., 2009). However, the NN-based models may not be able to cope up with the non-stationarity in input data, if pre-processing of input and/or output data is not performed (Cannas et al., 2006). In order to address this issue of non-stationarity, integration of discrete wavelet transform (DWT) techniques with the data-driven models have been advocated for rainfall–runoff modeling (Nourani et al., 2009, 2013; Tiwari and Chatterjee, 2011; Sehgal et al., 2014a,b; Seo et al., 2015). Furthermore, for larger basins, the past information on rainfall and discharge together contains more information on memory of the catchment response than rainfall only. Hence, the performance in discharge forecasting could be improved by either considering previous discharge as an input also or using the first level model output as the feedback input into the network for forecasting at the second level. For flow forecasting, the regression-based linear models, viz., the autoregressive moving average (ARMA), autoregressive integrated moving average (ARIMA), autoregressive moving average with exogenous inputs (ARMAX), and autoregressive integrated moving average with exogenous inputs (ARIMAX) models are found to be useful (Chang and Chen, 2001; Nourani et al., 2013; Valipour et al., 2013). Similarly, the nonlinear autoregressive with exogenous inputs (NARX) based recurrent NNs (Lin et al., 1996) are also used for flood forecasting (Kumar et al., 2004; Kisi, 2009; Besaw et al., 2010; Chang et al., 2014a,b).

Although most of the validation projects reveal that the precipitation radar of TMPA produce errors within the acceptable range after the gauge-adjustment, however, the real-time satellite rainfall product is not acceptable for flood forecasting due to the high bias involvement (Kneis et al., 2014; Tong et al., 2014). Although the TRMM-RT estimates proved to give promising flow forecasts by incorporating time-lagged observed streamflow in the input (Akhtar et al., 2009; Nourani et al., 2013), however, streamflow is seldom available on a real-time basis in many catchments worldwide. Hence, to address this shortfall in flow forecasting, there is a scope to use the time-lagged streamflows into the dynamic NARX model, where the forecast flow can also be used as a feedback input to the NN with the real-time satellite-based rainfall data. Moreover, the past studies have never verified the short- and medium-range flood forecasting accuracy with reanalysis temperature products available from the Climate Forecast System Reanalysis (CFSR). As far as the authors are aware, limited studies have been conducted to evaluate the utility of the satellite-based precipitation products in streamflow forecasting using NNs. Furthermore, the non-stationarity in the meteorological timeseries could also be reduced with multi-resolution analysis using wavelets that decomposes the input data into approximate and detail (noise) components, and finally the noise is eliminated through correlation analysis with the observed output data. However, the wavelet-based ANN (WANN) does not use the past streamflow information depicting the catchment and meteorological characteristics; due to which this model may not be always useful with the meteorological input data with high biases. Although, the NARX model has the advantage of using the past streamflow information by dynamic feedback inputs; however, non-stationarity in the rainfall timeseries could limit its performance.

In light of the above discussion, in this study, a novel hybrid NN model, namely, the wavelet-based non-linear auto regressive model with exogenous inputs (henceforth, called as WNARX) model is proposed, that utilizes the advantage of NN, wavelet-based error decomposition and memory of the catchment in terms of the dynamic feedback inputs. The multi-step ahead river flow forecast accuracy by the proposed WNARX model is compared with the ARMAX, static ANN, WANN and NARX models. The developed models are tested herein for real-time flood forecasting in the upper Mahanadi River basin, upstream to the Hirakud reservoir, in eastern India. According to the flood scenario of the Mahanadi River basin, very frequent floods occur at the downstream end of the river (Tiwari and Chatterjee, 2010a,b, 2011; Jena et al., 2014). Hence, prior forecast of high river discharges at the upper catchment outlets would contribute to the operation of the Hirakud reservoir, which eventually controls the downstream flood discharge.

This paper is organized as follows: succeeding to the Introduction section, Section 2 describes the details of the study site and data used; the methodology used is described in Section 3; the outcomes of the inter-comparison of the ARMAX, ANN, WANN, NARX and WNARX models with input of different rainfall products and assessment of their feasibility in flood forecasting at 1–3 days lead times are presented in Section 4; and Section 5 concludes the study with future scope of research.

2. Study area and data used

The models developed in this study are applied for streamflow forecasting at the Basantpur gauging station on the Mahanadi River in eastern India. This river basin is one of the largest river basins of India covering an area of 141,000 km^2 in the states of Chhattisgarh and Odisha. In the upstream of this flood prone river basin, the multi-purpose Hirakud dam provides some amount of flood relief by storing part of flood water. Of the tributaries in upstream to the Hirakud dam, the catchment of Basantpur gauging site contributes the maximum inflow into the dam. It encompasses a catchment area of 59200.10 km^2 which spans from 19.50° to 23.80°N latitudes and 80.00° to 83.00°E longitudes (Fig. 1). The study area is characterized by tropical monsoon receiving the maximum precipitation in the months of July–September.