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
0022-1694/

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.

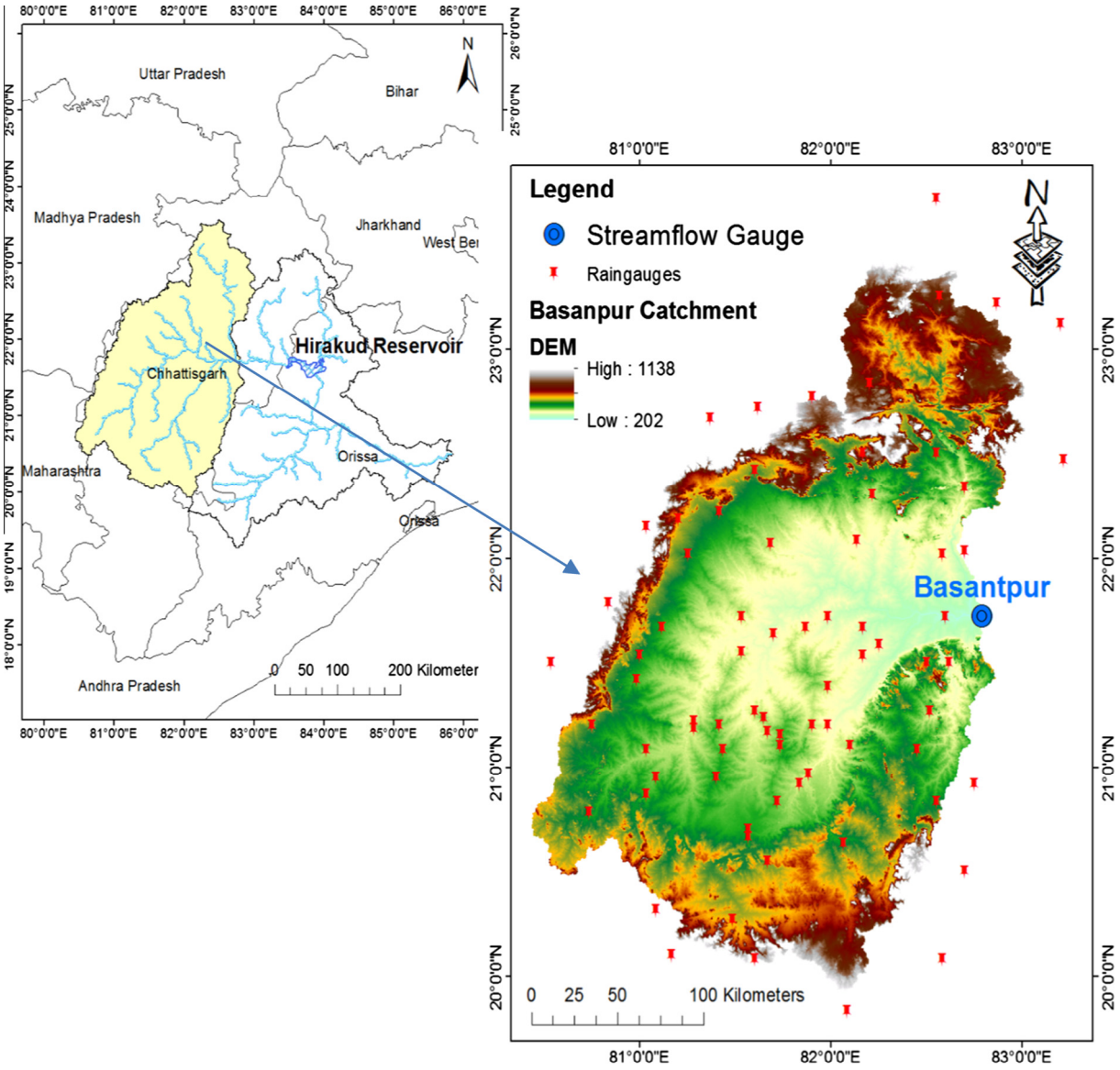

Fig. 1. Index map of the Mahanadi River basin at Basantpur showing location of raingauges and streamflow gauge.

Daily rainfall data of 75 raingauge stations in the Basantpur catchment for the period 2000–2010 are collected from the India Meteorological Department (IMD). Moreover, with the initiative of different satellite-based rainfall information, especially, the TMPA, real-time flood prediction could become possible in this flood prone area. The TMPA rainfall products are obtained at 0.25° × 0.25° (latitude × longitude) spatial and 3-h temporal resolutions. The TMPA products are available in two versions: post real-time research product, 3B42 from <ftp://disc2.nascom.nasa.gov/data/s4pa/TRMM_L3/TRMM_3B42> (henceforth, called as TRMM) and near real-time, 3B42RT from <ftp://trmmopen.gsfc.nasa.gov/pub/merged/3B42RT> (henceforth, called as TRMM-RT). In this study, the TMPA V7 of both 3B42 and 3B42RT rainfall products (Huffman and Bolvin, 2013) are used. The 6-h maximum (T_{max}) and minimum (T_{min}) temperature data are obtained from the coupled National Centers for Environmental Prediction (NCEP) and CFSR products from <http://rda.ucar.edu/datasets/ds093.1/> in grid format with resolution of 0.3125° × 0.3125°. Thereafter, the rainfall and temperature data are aggregated temporally to obtain these products at 1-day temporal resolution for their use in the developed models. Since all the data-driven models developed herein use the mean areal rainfall as one of the inputs, the mean catchment rainfall at daily-scale is computed by using the dynamic Thiessen polygon method. By this method, depending on the available station-specific rainfall data on a particular day, the Thiessen weights are reconstructed on a daily-scale to account for the missing rainfalls at a specific station. As the available IMD gauge-based rainfall data are having a temporal resolution of 1-day, the 3-h satellite rainfalls are aggregated to daily-scales. Daily discharges at the Basantpur gauging site in the Mahanadi River observed by the Central Water Commission (CWC), New Delhi is collected from <www.indiawris.com>. Although daily discharge and IMD rainfall data are available for longer durations, these data for the period from 2000 to 2010 are chosen since the TMPA products are only available after the year 2000. Table 1 shows the statistical characteristics of ranges, means and standard deviations of the data records used for training (2000–2007) and testing (2008–2010) the models. Only the monsoon climate–streamflow data (from 15 June to 15 October) are used, as the study aims to develop an efficient flood forecasting model which is operated during the high flow periods. Moreover, as compared to high flows, the low flows are mostly contributed

Table 1
Statistical parameters of data sets used for training and testing of models.

Dataset		Range	Mean	Stdev
Rainfall (mm/day)	Training			
	IMD	0–96.65	8.14	10.99
	TRMM	0–104.00	8.86	12.25
	TRMM-RT	0–100.29	7.76	10.57
	Testing			
	IMD	0–69.12	7.49	10.05
	TRMM	0–92.09	9.12	12.68
	TRMM-RT	0–67.16	8.58	11.33
Discharge (m^3/s)	Training	9–33087.95	1736.46	2394.97
	Testing	3.61–16826.73	1362.83	2022.38

by base flow, rather by the rainfall occurring on previous days or in real-time. At the same time, it is common to develop separate NNs over distinct hydrological seasons to improve the forecasts (Singh and Deo, 2007); and thus, to show the proof-of-concept, the monsoonal streamflow events are modeled to reduce the influence of low flow events.

3. Methodology

3.1. Input selection

The data-driven models studied herein are trained and verified for daily discharge (Q) prediction using the areal estimate of daily rainfall (R) and daily average temperature (T), calculated as the average of daily T_{max} and T_{min}. The performance of these rainfall–runoff models is dependent on the selection of appropriate input vector *a priori*. The time-lag for the independent inputs of R and T is decided based on the maximum cross-correlation function (CCF) between the inputs and the observed discharge (Q); and for the dependent variable (Q), it is decided based on the auto-correlation function (ACF) and partial auto-correlation function (PACF) of the Q timeseries (Tiwari and Chatterjee, 2011; Sehgal et al., 2014a). Fig. 2a–d shows the CCFs between the input variables (IMD, TRMM and TRMM-RT rainfalls, and temperature) and discharge; and Fig. 2e and f shows the ACF and PACF for the discharge variable, respectively, indicating the 95% confidence limits with 'bound 1' and 'bound 2'. The cross correlation analysis illustrated in Fig. 2a–d reveals the temporal climate (rainfall, temperature)-flow dependency of R and T up to a delay time of 4-days and 3-days, respectively. Similarly, it is observed from Fig. 2e–f that the magnitudes of ACF and PACF at lag times of 1- to 4-days are high; hence, discharge variables up to 4-days lag is considered to be significant. Therefore, to simulate daily discharge, the developed NN-based models are provided with these significant time-lagged inputs of R and T and significant time-lagged discharge feedback to the network.

However, it is challenging to identify the pre-event precipitation that triggers flood peaks. It is also difficult to determine the average time-lag between the most intense rainfall rate and the discharge peak for a particular catchment for all the flood events. Therefore, cumulative rainfall over a particular period before the flood day could be used as the antecedent rainfall (AR). As Figs. 2a–c indicate that rainfall in the catchment still correlates discharge well up to about 8-days lag, the sum of rainfalls before 4–8 days of the flood event is considered as AR in the input scenario. The efficacy of the models is tested for four different sets of input options given by: Option [1] with only time-lagged rainfall variables (only R), Option [2] with time-lagged rainfall variables and antecedent rainfall (R and AR), Option [3] with time-lagged rainfall and temperature variables (R and T), and Option [4] with time-lagged rainfall and temperature variables, and antecedent rainfall (R, AR and T). Before performing the NN-based modeling, the input and output variables are normalized and scaled between [0 1] in order to ensure that all the variables receive equal attention during the training of a model.

3.2. Neural network (NN)-based models

Since data-driven models are based on learning and pattern recognition, it is very common in NN-based rainfall–runoff models to use time delay inputs in addition to the current information. A brief detail of the developed static and dynamic NN-based models are presented in the following sections.

3.2.1. Multilayer perception (MLP) ANN model

Multilayer perception (MLP), a class of NN, is a multilayer series of nodes or neurons of input, hidden, and output layers. Of the varieties of NN structures available, the commonly used feed-forward NN used herein is given by

$$\widehat{Q} = f_o\left[\sum_h w_{oh}\cdot f_h\left(\sum_i w_{hi}x_i + b_h\right) + b_o\right] \quad (1)$$

where x_i is the independent input variables of R and T; w_{hi} is the weight from input to hidden layer; w_{ho} is the weight from hidden to output layer; f_h and f_o are the transfer functions for hidden and output layers, respectively; and b_h and b_o are the bias for the hidden and output layers, respectively. The network uses a logistic sigmoid transfer function in the hidden layer and a linear transfer function is used for the output layer.

3.2.2. NARX model

For flood forecasting, the NARX model predicts the future values of streamflow timeseries, $Q(t)$ by regressing from the past prediction values of Q and the past values of exogenous (external) input timeseries of rainfall and temperature, $x(t)$. The form of the NARX model is expressed as (Lin et al., 1996)

$$\widehat{Q}(t) = f\left(x(t-1),\ldots,x(t-d),\widehat{Q}(t-1),\ldots,\widehat{Q}(t-d)\right) \quad (2)$$

where t is the current time-step, d = time lag; the terms $x(t)$, …, $x(t-d)$ are the time lagged exogenous inputs of R and T. Similarly, $\widehat{Q}(t-1)$, …, $\widehat{Q}(t-d)$ are the endogenous inputs that are being produced with any time delay d, and $f(\cdot)$ is a nonlinear function that estimates the next forecast value.

3.2.3. WANN model

Considering the analogy of input signals with the hydro-meteorological timeseries data, the Wavelet Transform (WT) method (Morlet et al., 1982) is used to express the asymmetric and irregular input data signals or waves as sum of the sub-signals or wavelets. This transform is based on a mother wavelet function that constructs a family of wavelets of finite interval, mainly characterized as the continuous and discrete wavelet transforms. Considering the discrete nature of observed hydro-meteorological timeseries data, the discrete wavelet transformation (DWT) is preferred in most hydrological forecasting problems. The DWT operates on two sets of functions viewed as high-pass and low-pass filters to produce discrete wavelet coefficients (DWC). For an input signal x, the first step produces two sets of DWCs: high pass approximation coefficients, A_1 (low frequency) and low pass detail coefficients, D_1 (high frequency). The next step splits the approximation coefficient A_1 into two parts using the same scheme, replacing x by A_1, and producing A_2 and D_2, and so on. The wavelet decomposition of the input x analyzed at level n has the structure of [A_n, D_n, D_{n-1}, …, D_2, D_1]. Recent studies on

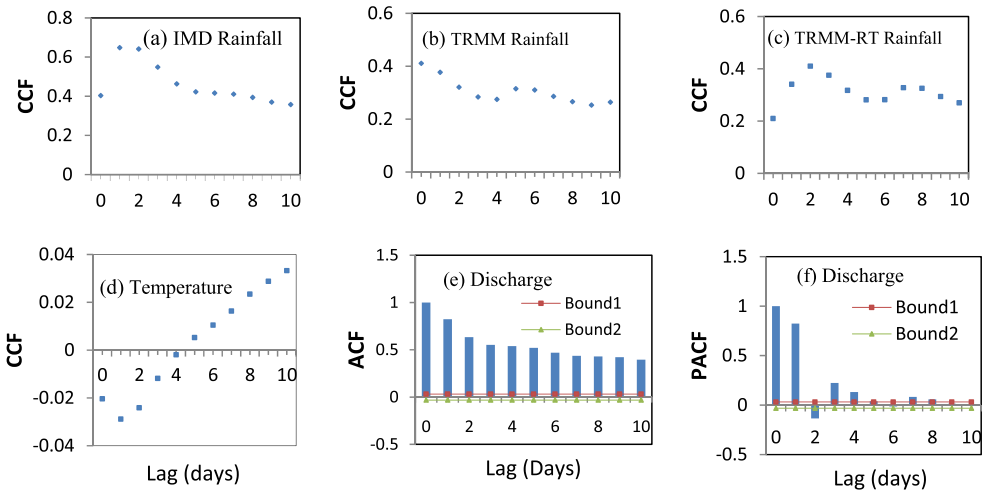

Fig. 2. Lag correlation analysis of the input variables used in the study.

coupling of any data-driven technique with the wavelet transform have been proved to improve the rainfall–runoff modeling performance (Tiwari and Chatterjee, 2011; Badrzadeh et al., 2015; Sehgal et al., 2014a,b; Shoaib et al., 2014). Hence, the simple ANN is coupled with the wavelet decomposition method to develop the wavelet-based ANN.

3.2.4. Development of the proposed WNARX model

Since the TRMM-RT rainfall product has very high biases, generally this product is not found to be very much useful for real-time flood forecasting with the existing rainfall–runoff models. To make this rainfall product useful for operational flood forecasting, there is a need to develop an ANN model framework which has the capability to minimize the random error component associated with the input data with the dynamism to incorporate the catchment behavior involving the catchment response in real-time due to the exogenous inputs of rainfall and temperature. Generally, the DWT has the capability to minimize the random error associated with the input data using the high pass and low pass filtering approaches. However, for incorporating the catchment dynamism in real-time due to the meteorological forcings, the streamflow data is not available in real-time. Consequently, the model simulated streamflow in real-time could be used as the feedback input to the model to improve forecast efficiency of the proposed model. Using this philosophy, the WNARX model is developed herein that couples the concepts of ANN, wavelet analysis and NARX model. This hybrid WNARX model integrates the NARX model with the DWT of the input variables (see Fig. 3). The framework of DWT is given by

$$T(a,b) = \frac{1}{\sqrt{a}} \sum_{t=0}^{N-1} f(t) g^* \left(\frac{t-b}{a} \right) \quad (3)$$

where $T(a,b)$ are the wavelet coefficients; g^* corresponds to the complex conjugate; g is the wavelet function or mother wavelet; and the variables a and b are the scale factor and time factor, respectively. For DWT, the available wavelet functions, varying from smaller vanishing moment 'Haar' to higher vanishing moment 'Daubechies' (db) functions (Sehgal et al., 2014b), are tested to achieve the best wavelet function. For any input timeseries x_N of length N, the DWT consists of $\log_2 N$ stages at most, but it is valid for fully autoregressive data without considering the seasonal variation of hydrological processes. If the level of decomposition is increased, more temporal and spectral variation in input data could be achieved (Shoaib et al., 2014); and the input timeseries is decomposed into a large-scale approximation and small-scale detail coefficients. In order to obtain the optimum level of decomposition, the DWT is tested by varying the decomposition level. For the current input datasets of rainfall and temperature, db25 wavelet function with nine-level of decomposition is found to be the best. At the first step, data pre-processing procedure is applied, in which the rainfall and temperature timeseries is decomposed to produce DWCs. After decomposition, the significant DWCs or wavelets are selected by correlation analysis between the DWCs and the observed discharge; and the sum of these selected wavelets forms the new timeseries (e.g., Tiwari and Chatterjee, 2010b, 2011). Finally, the new timeseries of rainfall and temperature variables are used as input to the NARX model (Eq. (2)). The model output is the undecomposed discharge timeseries which produces the feedback input to the network. The model is developed using the MATLAB toolbox. The difficulties arising from the timeseries of river discharge with different frequency components could be overcome by the proposed WNARX model as this model accounts for both the long-term seasonality of the meteorological inputs through DWT as well as the memory of the catchment through the autoregressive components of the output discharge variable.

3.3. ARMAX model

The dynamic timeseries-based linear ARMAX model is a commonly used method which is used for comparison purposes in this study. This model is capable of incorporating an exogenous input variable unlike the static ARMA model. The form of the ARMAX model is given by

$$A(q)Q(t) = B(q)x(t-d) + C(q)\varepsilon(t) \quad (4)$$

where $x(t-d)$ = time delay input variables of R and T; $Q(t)$ = output discharge variable; ε = white noise; d = time delay between input and output; and $A(q)$, $B(q)$, and $C(q)$ are the polynomials of the regression equation. The time lagged inputs of R and T are used herein as exogenous input variables to predict the future discharge. The degree of individual polynomials is decided by trial and error approach.

3.4. Multi-time-step-ahead discharge forecasting

For discharge simulation at several time-steps ahead, the selected time-lagged input variables (see Section 3.1) are used in the developed models for generating medium-range forecasts with lead times of 1-, 2- and 3-days to predict $Q(t+1)$, $Q(t+2)$, $Q(t+3)$,

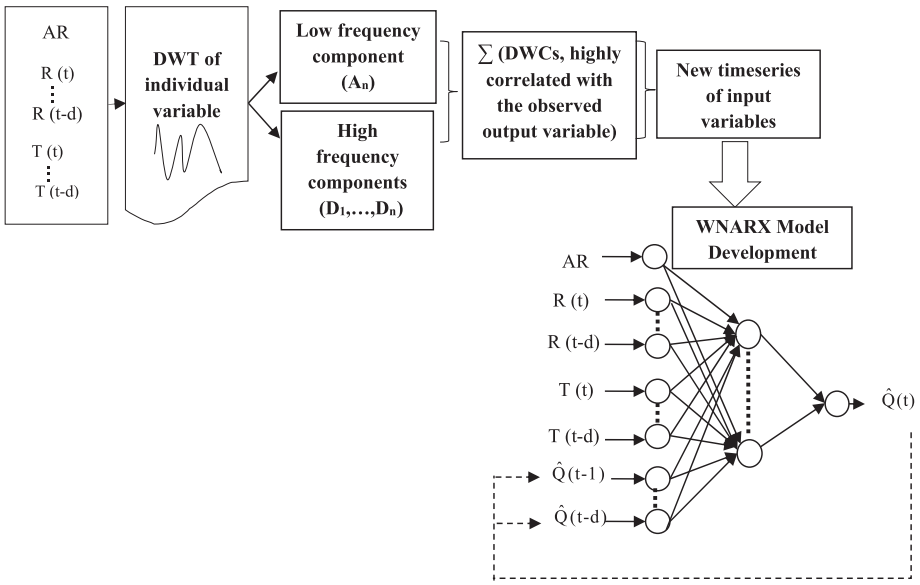

Fig. 3. Architecture of the WNARX model.

respectively. The models are applied and tested with the observed (IMD) and satellite-based (TRMM and TRMM-RT) rainfall products as forcings to forecast the discharge up to 3-days lead times.

3.5. Criteria for performance evaluation

First, the accuracy of the satellite rainfall products of TRMM and TRMM-RT are evaluated with respect to the benchmark IMD rainfalls by using the performance evaluation measures of correlation coefficient (r), percent bias ($PBias$), probability of detection (POD) and false alarm ratio (FAR) expressed as

$$r = \frac{\sum_{i=1}^{n}\left[(O_i - \overline{O})(P_i - \overline{P})\right]}{\sqrt{\left[\left(\sum_{i=1}^{n}(O_i - \overline{O})^2\right)\left(\sum_{i=1}^{n}(P_i - \overline{P})^2\right)\right]}} \tag{5}$$

$$PBias = \frac{\sum_{i=1}^{n}(P_i - O_i)}{\sum_{i=1}^{n}O_i} \times 100 \tag{6}$$

$$POD = \frac{No\ of\ hits}{No\ of\ hits + misses} \tag{7}$$

$$FAR = \frac{No\ of\ false\ alarms}{No\ of\ hits + false\ alarms} \tag{8}$$

The POD is the ratio of number of extreme rainfall events that are correctly forecast to the total number of extreme rainfall events at a given threshold, which varies from 0 to 1, 1 being when 100% of the storms are detected. FAR is the measure of failure of the forecaster to exclude the non-extreme events that varies from 0 to 1, 0 being the best case in which no false alarms are issued.

Second, the efficacy of the developed models with different rainfall products are compared using seven performance evaluation measures of: (i) Correlation coefficient (r), (ii) Nash–Sutcliffe efficiency (NSE) (Nash and Sutcliffe, 1970), (iii) ratio of root mean square error ($RMSE$) to standard deviation of measured data (RSR), (iv) mean absolute error (MAE), and (v) error in volume ($Evol$), (vi) error in peak flow ($Epeak$) and (vii) error in time to peak flow (Et). These error measures are expressed as

$$NSE = \left[1 - \frac{\sum_{i=1}^{n}(O_i - P_i)^2}{\sum_{i=1}^{n}(O_i - \overline{O})^2}\right] \times 100 \tag{9}$$

$$RSR = \frac{RMSE}{STDEV_{OBS}} = \frac{\sqrt{\sum_{i=1}^{n}(O_i - P_i)^2}}{\sqrt{\sum_{i=1}^{n}(O_i - \overline{O})^2}} \tag{10}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|O_i - P_i| \tag{11}$$

$$Evol = \left(\sum_{i=1}^{n}P_i \Big/ \sum_{i=1}^{n}O_i - 1\right) \times 100 \tag{12}$$

$$Epeak_i = (Qp_i/Qo_i - 1) \times 100 \tag{13}$$

$$Et_i = (Tp_i/To_i - 1) \times 100 \tag{14}$$

where O_i and P_i are the observed and predicted discharge values for ith day, respectively; $\overline{O}$ is the mean of the observed discharge values; n is the number of data points; Qo_i and Qp_i are the observed and predicted peak flood values, respectively; and To_i and Tp_i are observed and predicted time to peak flood values, respectively.

Analysis indicates that, unlike $RMSE$, MAE is the most natural and unambiguous measure of average error magnitude (Willmott and Matsuura, 2005). However, RSR incorporates the benefits of error statistics ($RMSE$) and includes a scaling/normalization factor (Akhtar et al., 2009) so that the resulting statistics can be applied to different variable types. $Evol$ is a very important index, especially in flow prediction, since it indicates the conservation of flow volume; but it is seldom used by researchers in ANN-based rainfall–runoff models. Moreover, for any hydrological model, the volume conservation ability of the model is very much needed to test its suitability for field application. Similarly, since the simulated output may not always fall on the 1:1 plot of the observed-simulated values, even with high r values, only using the correlation coefficient as the single most error measure does not guarantee the best performance of the model.

4. Results and discussion

4.1. Comparison of TRMM and TRMM-RT rainfall estimates with respect to the observed data

First, the satellite rainfall products of TRMM and TRMM-RT are evaluated with respect to the observed IMD rainfall, by inter-comparing the magnitudes. Fig. 4a shows the scatter plots of the areal rainfall timeseries of the TRMM and TRMM-RT against the IMD gauge-rainfall during the monsoon seasons (15 June to 15 October) from 2000 to 2010. The TRMM rainfall product has a correlation of 0.83 with the IMD-gauge rainfall; whereas, the TRMM-RT product is found to be moderately correlated ($r = 0.74$) with the IMD rainfall. The TRMM performs better than the TRMM-RT in which the former explains about 70% rainfall variability as compared to 55% variability by the latter.

The discrepancy or bias in rainfall magnitudes are separately computed for low ($R <$ mean), medium (mean $< R <$ mean + standard deviation) and high ($R >$ mean + standard deviation) rainfall regimes in the box and whisker plots as illustrated in Fig. 4b. For low rainfall events, the TRMM is relatively less biased than TRMM-RT, although both the products over-predict the rainfalls (*PBias* > 0). For the medium rainfall events, the TRMM shows slightly more positive bias indicating over-prediction than the TRMM-RT. The high rainfall events are better estimated by TRMM with *PBias* < 5%, whereas, TRMM-RT is negatively biased implying under-prediction. In essence, it is clear that the TRMM performs better than the TRMM-RT for all kind of rainfall regimes considered herein. This is because, the TRMM-RT product is a multi-satellite rainfall product disseminated in real-time and it is not bias corrected. On the other hand, the bias corrected TMPA product is combination of multiple satellites as well as gauge analyses (Huffman et al., 2007). The TRMM product is bias corrected by the TMPA using the Global Precipitation Climatology Project (GPCP)'s monthly rain gauge analysis and the Climate Assessment and Monitoring System (CAMS)'s monthly rain gauge analysis (Huffman et al., 2007; Yong et al., 2010); and the same cannot be carried out for the TRMM-RT during real-time flood forecasting. However, for real-time flood forecasting, the TRMM-RT rainfall is only useful since this product is available at near real-time for field application, whereas TRMM is available at a time lag of 2–3 months making it non-usable.

Fig. 5 illustrates the *POD* and *FAR* values of daily-scale TRMM and TRMM-RT rainfall products for the specified threshold levels of 0–100 mm/day. The results reveal that the *POD* values diminish with increase in threshold showing that the high intensity rainfall events are deviated from the gauge-based counterparts. However, the high rainfall events with 30.11 mm/day (95 percentile) are better detected by the TRMM with a high *POD* of 80% as compared to that of 45% by the TRMM-RT. For extremely high rainfall (>80 mm/day), neither the TRMM nor the TRMM-RT captures the observed rainfall magnitudes. Similarly, as envisaged from Fig. 5, *FAR* increases with increase in rainfall thresholds for both the TRMM and TRMM-RT rainfalls. At high precipitation (30.11 mm/day), the *FAR* is about 50% for both TRMM and TRMM-RT. In general, there is no significant difference between the *FARs* for TRMM and TRMM-RT, except for 60–70 mm/day rainfall threshold ranges (common in monsoon period), when TRMM-RT creates 80% false alarm as compared to 60% by TRMM. The patterns of *POD* and *FAR* indicate poor reflection of IMD gauge-based areal rainfall by the TRMM and TRMM-RT products. Hence, the effect of these satellite-based rainfall products has to be further evaluated for flood forecasting, being used as the basic forcing input into the NN-based models.

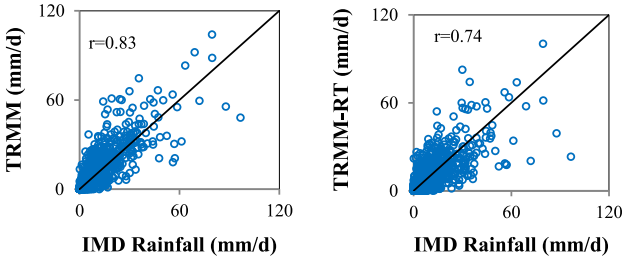

Fig. 4a. Scatter plots for the satellite-based rainfall products against IMD rainfall.

4.2. Hydrological evaluation of the developed models with different rainfall products

4.2.1. Modeling catchment runoff at daily time-scale

Since the discharge estimation by the rainfall–runoff models developed herein are subjected to uncertainties due to error in rainfall inputs, the satellite-based rainfall estimates are evaluated based on their performances in estimating the corresponding discharges at catchment-scale by all the developed models. For each input scenario (Options [1]–[4]), as detailed in Section 3.1, the ARMAX and the NN-based models are trained and tested for the three sets of rainfall products of IMD, TRMM and TRMM-RT.

4.2.1.1. Effect of different input scenarios on real-time discharge simulation. Fig. 6 shows the comparison of Nash–Sutcliffe efficiency (*NSE*) for discharge simulation by the models developed herein for all the four sets of inputs. Based on the *NSE* measures, it is noticed from Fig. 6 that, in general, the model efficiency increases in the order ARMAX < ANN < WANN < NARX < WNARX in simulation mode. The model performances may be considered satisfactory for streamflow prediction when *NSE* is >50% (Moriasi et al., 2007). The plots of *NSE* of the models show that the flow prediction performance increases from a minimum value with only *R* as input in Option [1], and is the best in Option [4] that accounts for *R*, *T* and *AR*. Further, Fig. 6 reveals that the linear ARMAX model performance improves substantially in Option [4] for the IMD and TRMM-driven simulations; however, with the TRMM-RT-driven simulations, the model does not show large improvement for the Option [4] over Option [1]. This indicates the limited efficacy of the ARMAX model, even while accounting for more input information to quantify the non-linear rainfall–runoff process. Furthermore, it is found that the static NN models are more input specific than the dynamic NNs. With this effect, the model efficiency increases by 8.79%, 14.86% and 13.68% by the ANN model; and 13.66%, 13.10% and 8.92% by the WANN model for the IMD, TRMM and TRMM-RT-driven simulations, respectively. Moreover, for the WNARX and NARX models, Option [4] has little improvement in real-time flow prediction as compared to that with Option [1]; hence, the model efficiencies are found to be acceptable with Option [1] also. Overall, the results demonstrate that the strong influence of additional inputs, such as, the previous antecedent rainfall and time-lagged temperature improves the model efficiencies of all the models. This finding can be attributed to the fact that the catchment produces floods when soil is saturated after meeting the losses (Berthet et al., 2009; Froidevaux et al., 2015). The inclusion of *AR* accounts for soil moisture, and temperature accounts for evapotranspiration loss. However, for large river basins like Ganges, studies have found that large number of input parameters are not always responsible for reliable flood forecasting, rather the most correlated inputs, if available in real-time, solves the problem (Akhtar et al., 2009). Moreover, from Fig. 6, it is found that for the

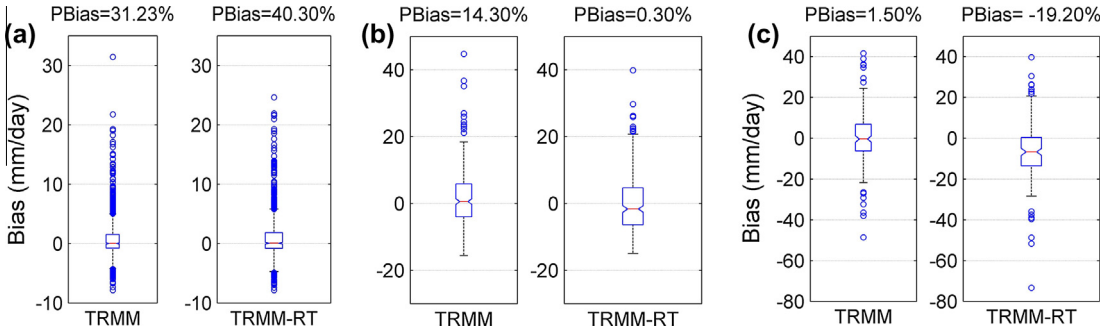

Fig. 4b. Comparison of TRMM and TRMM-RT products with IMD rainfall for (a) low (R < mean), (b) medium (mean < R < mean + standard deviation) and (c) high (R > mean + standard deviation) rainfall regimes.

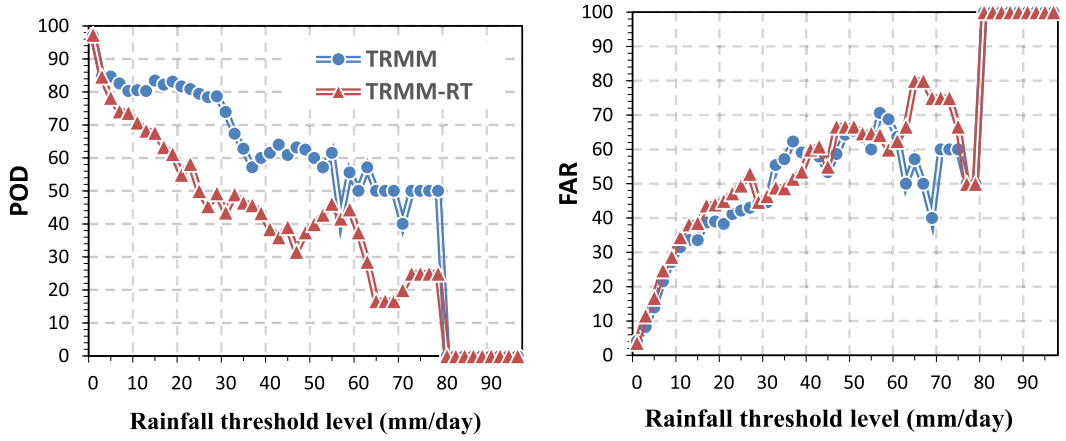

Fig. 5. Probability of detection (*POD*) and false alarm ratio (*FAR*) for TRMM and TRMM-RT rainfall products at different thresholds.

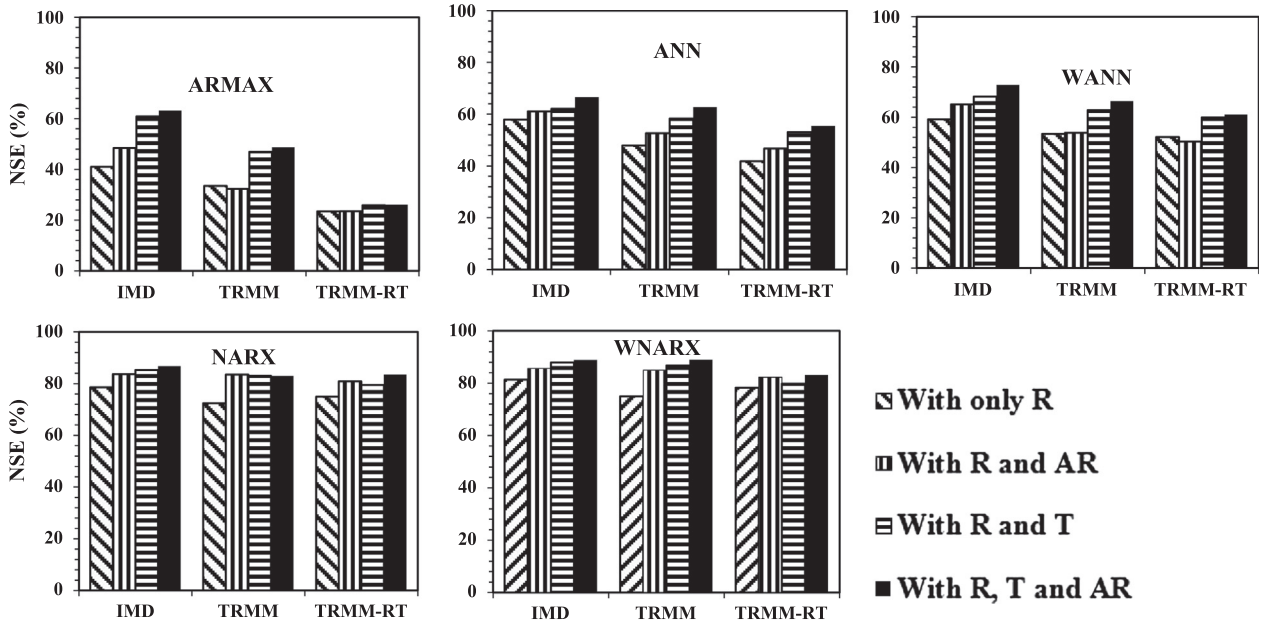

Fig. 6. Comparison of performance of the models for different input combinations to reproduce observed discharges during testing (2008–2010).

ARMAX and static NN models of ANN and WANN, Option [3] produces satisfactory results as compared to Option [2] for both the observed and satellite-based rainfall products. In case of Option [3], inclusion of temperature as a separate timeseries in the model indicates significant improvement in model results; however, the same is not significant in the WNARX and NARX models. This observation can be attributed to the fact that inclusion of temperature as a separate timeseries in the model includes four additional input variables of $T(t-3)$ up to $T(t)$; whereas, inclusion of *AR* adds only one variable. The analysis evidences relatively weak linkage of

the catchment to the rainfall forcing one week prior to flood. This corresponds with the previous finding that the correlation of triggering antecedent rainfall to produce flood is region-specific, and 3–4 days previous rainfall is significant to produce flood (Froidevaux et al., 2015). Comparison of Options [3] and [4] reveal that real-time discharge simulation efficiency increases by *NSE* of ~5% in case of Option [4] (with *R*, *AR* and *T* timeseries) over the Option [3], especially with the ANN and WANN models. From the overall observation of the plots, it is anticipated that for long-term flow forecasts, the antecedent catchment wetness long before the flood day may trigger flood. Hence, for assessment of the multi-step ahead streamflow forecasts, the Option [4] is considered to be the most suitable scenario for all the models.

4.2.1.2. Effect of observed and satellite-based rainfall data on real-time discharge simulation. The inter-comparison of rainfall products for real-time discharge simulation with Option [4] is illustrated in Fig. 7 for testing dataset, and the corresponding error measures are given in Table 2. It can be surmised from Table 2 that, for training datasets, the WNARX and NARX models always perform with *NSE* > 80%, *r* > 0.90 and |*Evol*| ⩽ 6%. Table 2 and Fig. 7 reveal that, for the testing datasets, the models perform similarly as in case of the training datasets with *NSE* > 80%, *r* > 0.90 and |*Evol*| ⩽ 9%. However, in general, during training phase, all the models are quite acceptable with the TRMM and TRMM-RT-driven simulations (Table 2). During testing with the ARMAX and static NN-based models, the high *RSR* values with the TRMM and TRMM-RT-

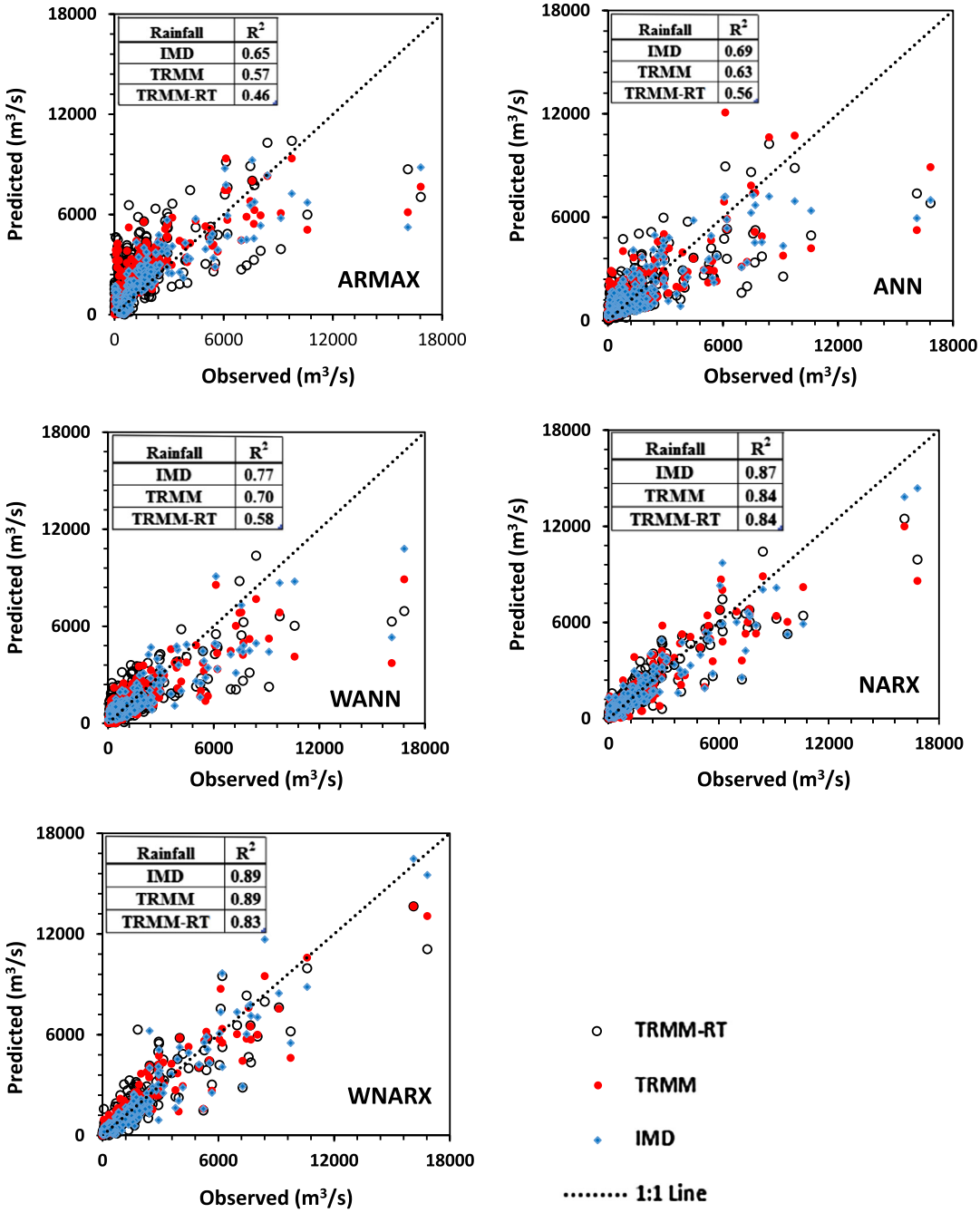

Fig. 7. Scatter plots of observed and simulated discharges using different rainfall products during testing (2008–2010).

driven flow predictions are quite noticeable, indicating higher *RMSE*. However, with the ANN models, the flow prediction error is within the acceptable limit of $RSR < 0.7$ (e.g., Moriasi et al., 2007). The lower accuracy of the conventional ARMAX model as compared to the NN-based models can be attributed to the linearity associated with the ARMAX model in modeling the nonlinear rainfall–runoff process. The satellite-based rainfall when passes through the NN optimization, the weighted magnitude of rainfall could reproduce the catchment flow timeseries after the nonlinear transformation. This factor can be attributed to the correction of rainfall bias by the NN schemes to some extent. In the WANN model, the pre-processing of input data moderately improves the performance over the ANN model with respect to some medium and high flow predictions due to correction of data for non-stationarity (e.g., Badrzadeh et al., 2015). While the hydrographs produced with the IMD (observed) rainfalls seem to outperform the satellite-based counterparts, the scenario is different for WNARX and NARX models. With the WNARX and NARX models, the satellite rainfalls reproduced the desired hydrograph with very small magnitude of *RSR* and *MAE* and high value of *r* (Table 2), and *NSE* in the range of 80–90%. Despite the slight under-estimation, the TRMM-RT rainfall conserves the catchment flow volume with *Evol* = −1.16% and −2.63% by the WNARX and NARX models, respectively. In the NARX and WNARX models, the consideration of the short-term autoregressive features ($\widehat{Q}(t-1)$, $\widehat{Q}(t-2)$, $\widehat{Q}(t-3)$, $\widehat{Q}(t-4)$) surpassed over the weakness of the NN models. More interestingly, it is found herein that the WNARX model, with the long-term seasonality being taken into account through wavelet transform, is efficient to transform the satellite-derived rainfall into runoff in simulation mode.

Furthermore, the impact of the rainfalls is assessed based on reproduction of peak flow events by different models. Fig. 8 shows the reproduction of the discharge hydrograph for two typical peak flood events that occurred in 2008 (event [1]) and 2009 (event [2]) driven by all the rainfalls. The event [1] has an observed peak of 16826.73 m^3/s precedent by heavy rainfall. Similarly, event [2] has a peak of 16117.48 m^3/s precedent by heavy rainfall with dual flood peaks. For the event [1], the magnitude of peak error is quite significant in the static NNs, but the simulations show no error in time to peak discharge (i.e. *Et* = 0%). It is evident that this peak event is quite closely reproduced by the WNARX model (*Epeak* = −7.83%) as compared to the NARX model (*Epeak* = −14.57%) while using the IMD rainfalls. In event [1], the simulated peak discharge resembles the observed one with the IMD gauge as well as TRMM and TRMM-RT-driven simulations by the WNARX and NARX models. For this event there exists a phase lag in the peak flow forecast by the ARMAX and the static NNs, which is a common problem in data-driven methods that is still to be addressed during NN optimization. However, it is not observed in case of event [2]. The peak flow during the event [2] is well predicted by the models without time offsetting. The dynamic networks of WNARX and NARX reproduced the peak flow closer to the observed one than the static models. The NARX model resulted in underestimation with *Epeak* of −14.19% with the IMD gauge rainfall. The error further increases with the TRMM and TRMM-RT, since their hit rate decreases for high rainfall intensity. The WNARX model produced the simulated peak with *Epeak* of −15.34% and −15.33% by the TRMM and TRMM-RT rainfalls, respectively, as compared to that of 2.26% by the IMD rainfall. The sharp observed peak in event [2] (Fig. 8) indicates the strong influence of heavy downstream rainfall accumulated within short duration to cause the high outflow. Although this event is more or less influenced by the antecedent rainfall, the inclusion of recurrent discharge into the network of WNARX model reproduced the peak very well with the TRMM and TRMM-RT rainfall products as compared to other models.

4.2.2. Performance of the developed models for multi-step-ahead discharge forecasting

Since the input scenario of Option [4] with all the input variables of *R*, *T* and *AR* results in the best performance of all the models in simulation mode with the gauged and satellite-based rainfall products, this scenario is only used for real-time flood forecasting. The NN-based models developed herein are extended for forecast at 1- to 3-days lead times for monsoonal climate-flow events (15 June to 15 October). The ARMAX model is not further considered for discharge forecasting because of its relatively poor performance.

The typical discharge reproductions by the models are evaluated for 1- to 3-days ahead forecasts, and the same are illustrated in Figs. 9 and 10. The performance measures of all the models in forecasting mode with all the rainfall products are summarized in Table 3. It is evident from Figs. 9 and 10 and Table 3 that the flow forecasting efficiency of all the models reduces gradually with increase in lead times. However, in the WANN, NARX and WNARX models, the forecasting performances do not deteriorate significantly for up to 3-days lead times. With all the rainfall forcings

Table 2
Performance evaluation measures of different models in simulation mode with input Option [4] during training and testing.

	ARMAX		ANN		WANN		NARX		WNARX	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing	Training	Testing
					IMD					
r (−)	0.79	0.81	0.74	0.83	0.82	0.88	0.92	0.94	0.92	0.94
NSE (%)	63.14	63.24	53.90	66.65	66.44	72.82	83.83	86.84	83.28	88.89
RSR (−)	0.72	0.61	0.80	0.58	0.68	0.52	0.48	0.36	0.48	0.33
MAE (m^3/s)	852.47	791.48	887.81	622.76	790.80	538.49	510.33	396.53	466.75	334.81
Evol (%)	−0.75	−8.53	−2.32	−3.66	−0.30	−9.23	−4.64	−6.90	1.65	−4.83
					TRMM					
r (−)	0.77	0.75	0.80	0.79	0.81	0.84	0.92	0.92	0.94	0.94
NSE (%)	60.02	48.68	63.63	62.74	65.76	66.47	83.68	83.01	87.60	88.95
RSR (−)	0.75	0.72	0.71	0.61	0.69	0.58	0.48	0.41	0.42	0.33
MAE (m^3/s)	910.08	980.93	856.46	659.26	824.46	596.24	443.75	413.28	448.89	354.21
Evol (%)	−1.39	27.61	1.88	3.14	−1.65	−9.68	−5.65	−9.08	1.12	−1.03
					TRMM-RT					
r (−)	0.72	0.67	0.77	0.75	0.76	0.80	0.93	0.92	0.94	0.91
NSE (%)	51.52	26.15	58.45	55.48	57.27	61.07	86.49	83.51	87.58	85.12
RSR (−)	0.82	0.86	0.76	0.67	0.77	0.62	0.43	0.41	0.42	0.40
MAE (m^3/s)	1029.95	1253.89	908.37	780.57	965.17	784.74	438.88	410.34	425.95	409.06
Evol (%)	−1.73	46.63	7.20	7.02	1.35	−1.29	0.83	−1.16	−2.71	−2.63

and lead times, the WNARX model performs with *NSE* > 78%, |*Evol*| < 9%, *r* > 0.89, *RSR* ⩽ 0.47 and *MAE* < 501 m^3/s; whereas the NARX model performs with *NSE* > 74%, |*Evol*| < 16%, *r* > 0.86, *RSR* ⩽ 0.51 and *MAE* < 631 m^3/s. The WANN model performs with *NSE* > 45%, |*Evol*| < 12%, *RSR* ⩽ 0.74, *MAE* < 782 m^3/s, and *r* > 0.74; and the ANN model performs with *NSE* > 31%, |*Evol*| < 9%, *RSR* ⩽ 0.82, *MAE* < 911 m^3/s, and *r* > 0.58. Hence, this analysis along with Figs. 9 and 10 reveal that the WNARX is the best model under all the input and forecasting scenarios followed by NARX, WANN and ANN models. Moreover, the NARX model gives comparable performance with the WNARX model; and the ANN and WANN models may not be used for field application due to their poor performance levels.

As envisaged from Table 3, at 1-day ahead forecasting, the various error measures for the ANN model for all rainfall products considered herein varies with *NSE* ⩾ 59.07%, |*Evol*| ⩽ 4.72%, *r* ⩾ 0.77, *RSR* ⩽ 0.64, and *MAE* ⩽ 727.38 m^3/s; whereas for the WANN model, these are: *NSE* ⩾ 55.83%, |*Evol*| ⩽ 6.52%, *r* > 0.76, *RSR* ⩽ 0.66, and *MAE* ⩽ 716.01 m^3/s; which may not be acceptable for field condition. However, for the NARX model, these ranges are: *NSE* = 80.52–82.07, |*Evol*| = 4.06–12.79%, *r* = 0.90–0.92, *RSR* = 0.42–0.44, and *MAE* = 378.81–541.43 m^3/s; and for the WNARX model, these are: *NSE* = 82.01–88.02, |*Evol*| = 4.41–8.67%, *r* = 0.91–0.94, *RSR* = 0.35–0.43, and *MAE* = 354.55–453.39 m^3/s. Overall, it is surmised that, for 1-day ahead forecasting lead time

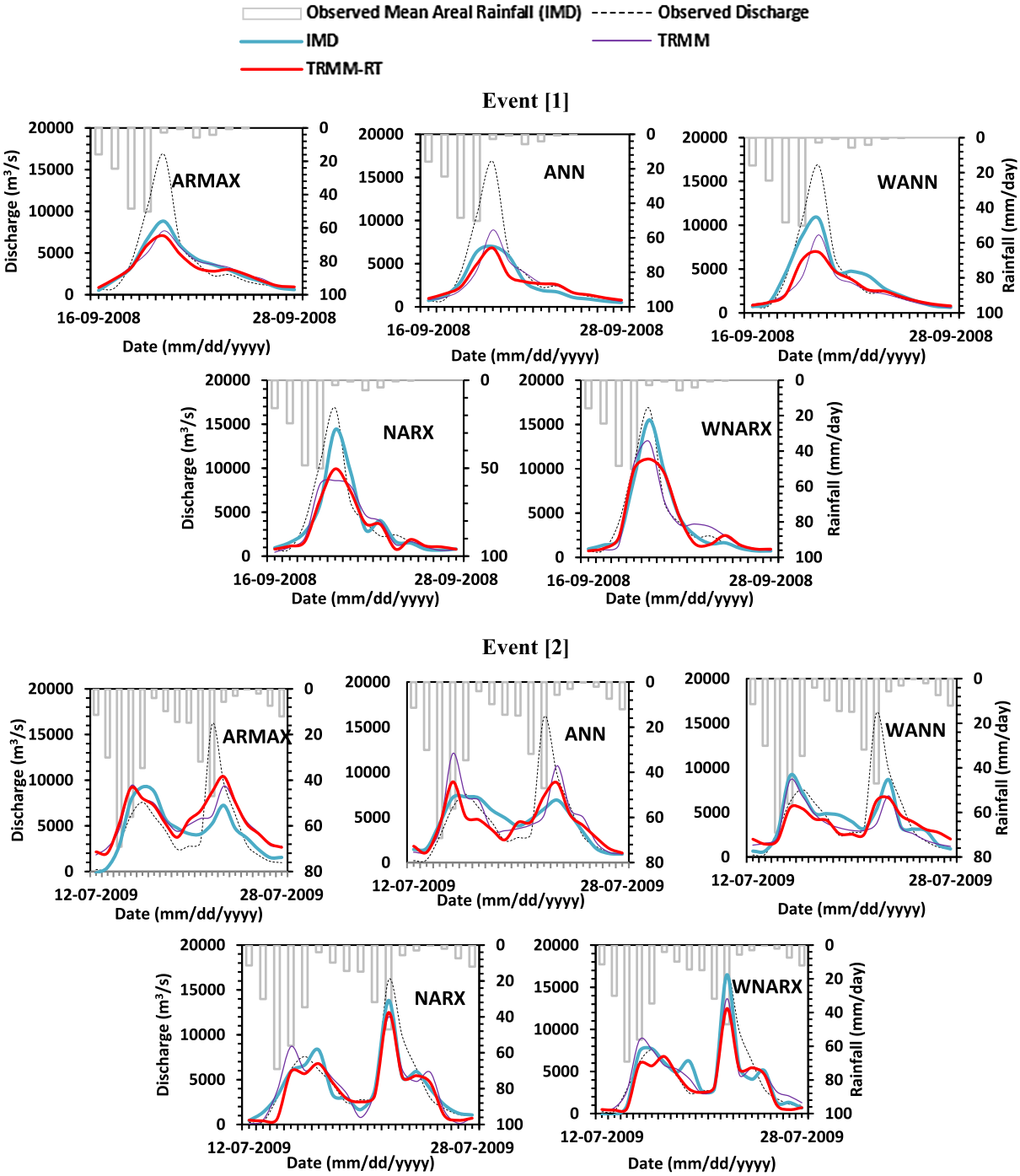

Fig. 8. Simulation of two peak flood events by all the models for input Option [4] with different rainfall products.

with the use of different rainfall products, the performance of both the ANN and WANN model decreases in the order: IMD > TRMM > TRMM-RT; whereas for the NARX model it is: TRMM > TRMM-RT > IMD; and for the WNARX, model it is: IMD > TRMM > TRMM-RT. Since only the TRMM-RT product is available in real-time, this product is only useful for real-time streamflow forecasting than other two rainfall products. Although both the NARX and WNARX models give comparable performances for 1-day ahead forecasting with the TRMM-RT product, however, the WNARX model is preferable over the NARX model due to its consistent performance in terms of various error measures (see Table 3). Such a best performance of the WNARX model may be attributed to the ability of this model to account for the long-term memory of the rainfall–runoff transformation process of the catchment, in terms of the dynamic recurrent input to the network that has the combined properties of ANN and wavelet transform to minimize the bias with the rainfall and runoff inputs.

Similarly, at 2-days ahead forecasting, various performance measures indicate that, with the use of different rainfall products, the performance in flood forecasting increases in the order ANN < WANN < NARX < WNARX (see Table 3). The various error measures for the ANN and WANN models with different rainfall products reveal that these two models cannot be used for 2-day ahead flood forecasting. Conversely, for the NARX model, the corresponding ranges of error measures are quite satisfactory with NSE = 75.15–77.03, $|Evol|$ = 0.21–15.65%, r = 0.87–0.88, RSR = 0.48–0.50, and MAE = 457.32–630.67 m^3/s; whereas for the WNARX model it is: NSE = 78.45–84.45, $|Evol|$ = 0.96–6.27%, r = 0.89–0.91, RSR = 0.40–0.47, and MAE = 382.32–501.01 m^3/s. In real-time discharge forecasting, the WNARX model is consistent

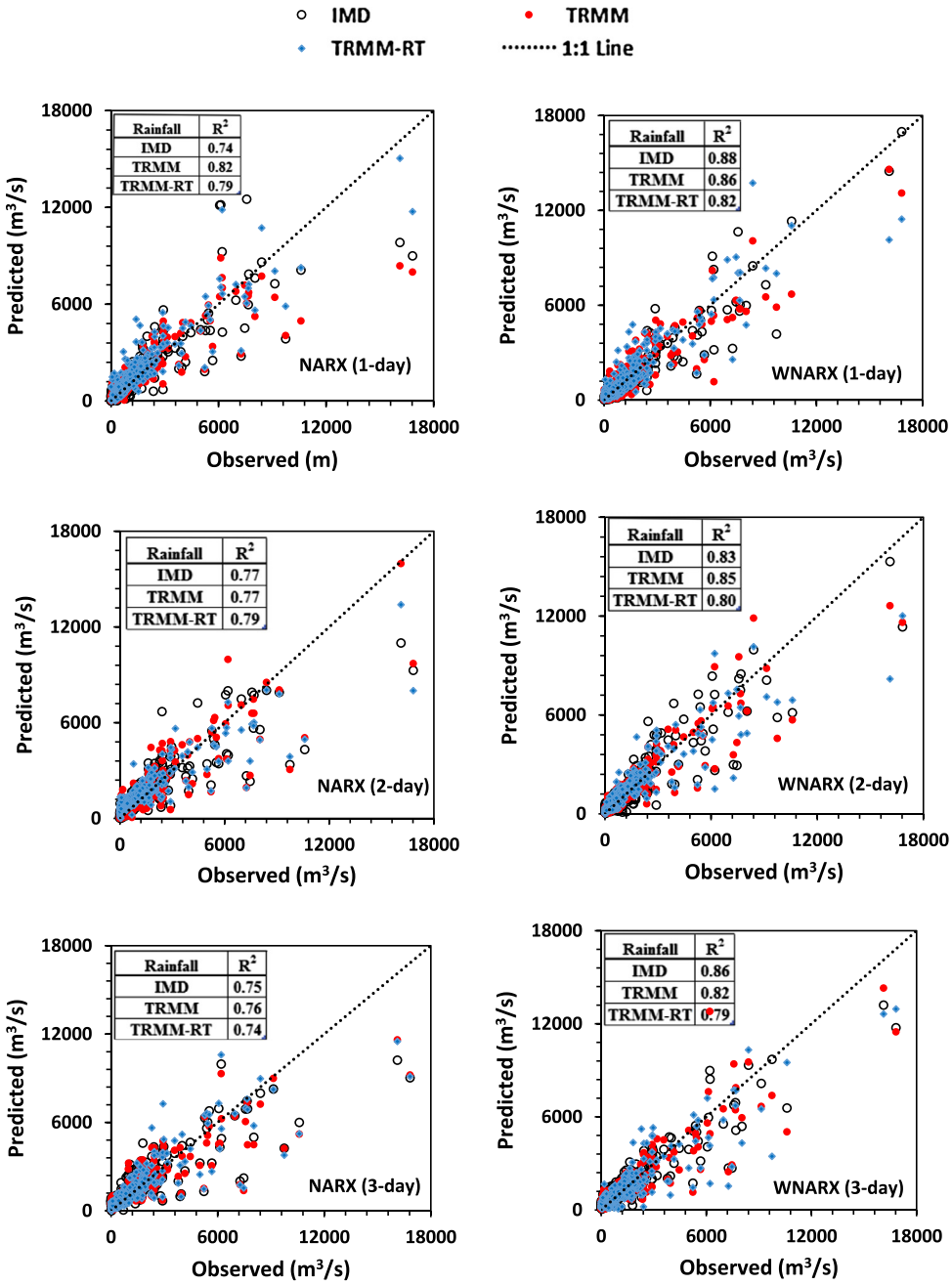

Fig. 9. Scatter plots for lead time monsoonal discharge forecasts for the testing period (2008–2010) by different models with different rainfall products.

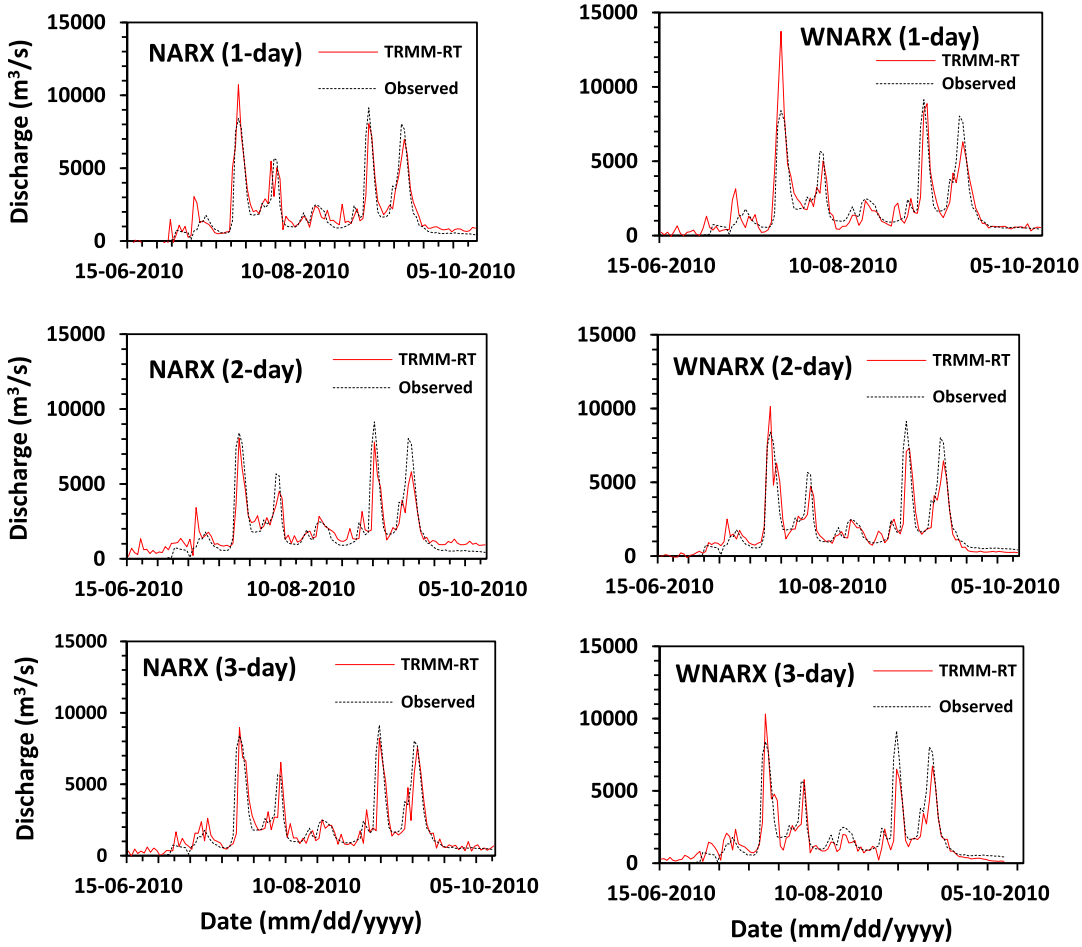

Fig. 10. A typical monsoonal discharge forecast for the year 2010 at lead-times of 1-, 2- and 3-days using the NARX and WNARX models with the TRMM-RT rainfall product.

Table 3
Error statistics of the multi-step-ahead discharge forecasts by the models during testing.

Models	Criteria	Lead time (day)			Lead time (day)			Lead time (day)		
		1	2	3	1	2	3	1	2	3
			IMD			TRMM			TRMM-RT	
ANN	*r* (−)	0.82	0.75	0.58	0.77	0.71	0.59	0.77	0.72	0.60
	NSE (%)	65.94	51.06	31.07	59.47	50.19	34.94	59.07	51.02	35.54
	RSR (−)	0.59	0.68	0.82	0.64	0.71	0.81	0.64	0.70	0.80
	MAE (m^3/s)	686.67	724.97	910.64	694.25	764.68	862.3	727.38	790.16	866.18
	Evol (%)	4.72	3.03	2.02	2.85	4.66	2.81	−4.42	1.98	−8.81
WANN	*r* (−)	0.84	0.84	0.75	0.81	0.78	0.75	0.76	0.74	0.68
	NSE (%)	68.90	61.36	60.54	63.48	56.70	51.91	55.83	54.00	44.92
	RSR (−)	0.60	0.59	0.68	0.60	0.66	0.69	0.66	0.68	0.74
	MAE (m^3/s)	659.73	616.24	679.09	642.32	672.10	645.86	716.01	781.74	779.81
	Evol (%)	−6.52	−12.21	−11.08	−1.27	−3.36	−6.33	4.47	7.05	3.23
NARX	*r* (−)	0.90	0.88	0.86	0.92	0.87	0.87	0.91	0.88	0.87
	NSE (%)	80.52	77.03	74.62	82.07	75.64	74.60	81.16	75.15	76.15
	RSR (−)	0.44	0.48	0.50	0.42	0.50	0.51	0.44	0.50	0.49
	MAE (m^3/s)	397.13	457.32	468.27	378.81	487.41	537.58	541.43	630.67	484.27
	Evol (%)	−4.06	0.21	0.50	−12.02	−1.47	5.17	12.79	15.65	0.58
WNARX	*r* (−)	0.94	0.91	0.93	0.93	0.92	0.91	0.91	0.89	0.89
	NSE (%)	88.02	82.31	84.69	85.53	84.45	80.80	82.01	78.45	78.69
	RSR (−)	0.35	0.42	0.39	0.38	0.40	0.44	0.43	0.47	0.46
	MAE (m^3/s)	354.55	501.01	448.80	440.43	382.32	454.40	453.39	439.18	495.04
	Evol (%)	−8.67	6.27	3.20	−7.28	−5.35	0.92	4.44	0.96	−5.72

Table 4
Probability of Detection (*POD*) and False Alarm Ratio (*FAR*) for lead time discharge forecasting using the TRMM-RT rainfall during testing.

	Lead time (Days)			Lead time (Days)			Lead time (Days)			Lead time (Days)		
	1	2	3	1	2	3	1	2	3	1	2	3
	±10% error bound			±20% error bound			±10% error bound			±20% error bound		
	$Q > Q_{90}$ WNARX + TRMM-RT						$Q_{75} < Q < Q_{90}$ WNARX + TRMM-RT					
POD (%)	41.67	44.44	30.00	69.23	56.67	46.88	63.89	67.50	48.89	80.00	82.22	64.58
FAR (%)	54.55	42.86	40.00	35.71	26.09	21.05	45.24	35.71	29.03	31.91	21.28	16.22
	NARX + TRMM-RT						NARX + TRMM-RT					
POD (%)	45.83	25.81	32.00	68.97	48.48	46.15	50.00	56.25	51.43	79.41	75.00	66.67
FAR (%)	52.17	38.46	57.89	25.93	15.79	45.45	69.05	56.10	51.35	43.75	41.30	36.59

Q denotes discharge; Q_{75} and Q_{90} are discharge value of 75th and 90th percentile.

Table 5
PBias of the multi-step-ahead discharge forecasting.

Lead time (Days)			Lead time (Days)		
1	2	3	1	2	3
WNARX			NARX		
$Q > Q_{90}$					
6.48	13.68	21.64	6.37	23.73	16.64
$Q_{75} < Q < Q_{90}$					
−7.62	−6.62	5.65	−16.09	−6.38	−3.27

in its performance among the four models with the TRMM-RT product.

At 3-days lead time discharge forecasting, the reproduction pattern of typical hydrographs (Fig. 10) and the supporting performance measures (Table 3) clearly indicate that there is a little scope of using the real-time rainfall product of TRMM-RT in flood forecasting with the ANN and WANN models. However, the 3-days lead time flood forecast is obtained with *NSE* > 74% by the WNARX and NARX models. The various performance measures for all the rainfall products with the NARX model varies in the ranges of: *NSE* = 74.60–76.15, |*Evol*| = 0.50–5.17%, *r* = 0.86–0.87, *RSR* = 0.49–0.51, and *MAE* = 468.27–537.58 m^3/s; whereas for the WNARX model it is: *NSE* = 78.69–84.69, |*Evol*| = 0.92–5.32%, *r* = 0.89–0.93, *RSR* = 0.39–0.46, and *MAE* = 448.80–495.04 m^3/s. These results indicate the potential of TRMM-RT to reproduce streamflow with acceptable accuracy in the real-time up to 3-days lead time. While using the TRMM-RT rainfall, Nourani et al. (2013) and Akhtar et al. (2009) could only achieve reliable flood forecasting up to 3-days lead time with preliminary spatial rainfall data selection, and pre-processing by runoff travel time and flow length, respectively. In these attempts, the TRMM and TRMM-RT products performed well only when the time-lagged observed discharge is used as one of the model inputs. However, the modeling approaches proposed herein, without using upstream river flow information and without any prior effort on spatial data selection, offers promising outputs with the satellite-based rainfall data. The overwhelming model performances reveal the reliability of the satellite rainfall products for extended streamflow forecast horizons with longer lead times.

The flow forecasting ability of the TRMM-RT rainfall product with the WNARX and NARX models is also evaluated by using the error statistics for high (>90th percentile) and moderate (between the 75th and 90th percentile) flows separately (Table 4). It can be surmised from Table 4 that for the high flows, the models could forecast 1-day lead time discharge with *POD* < 50% within the ±10% error bound; and within ±20% error bound, the corresponding values of *POD* are ~70%. Similarly, for different lead time forecasts, the *FAR* varies from 40.00 to 57.89% within ±10% error bound; and within ±20% error bound, it is 15.79–45.45%. For the moderate flow range, *POD* varies from 48.89–67.50% within ±10% error bound for the discharge forecasts; however, the corresponding *FAR* is less for the WNARX model (*FAR* = 29.03–45.24%) as compared to the *FAR* of 51.35–69.05% for the NARX model at 1–3 days lead times. Table 4 reveals that, within ±20% error bound, the WNARX and NARX models produce *POD* in the range of 64.58–82.22% for the moderate flow range; and the *FAR* varies from 16.22 to 31.91% for the WNARX model as compared to 36.59–43.75% by the NARX model, indicating slightly better flow forecasting by the WNARX model at 1–3 days lead-times. Note that, although 20% error bound of observed flow is towards the higher side; however, this is corresponding to only about ±0.15 m of water level (Central Water Commission, 1989; Tiwari and Chatterjee, 2010a) of the Mahanadi River studied herein, which is the largest river in eastern India. Hence, consideration of 20% error level in flow forecasting is justified for this river.

Moreover, the percentage biases (*PBias*) in flow forecasting using the TRMM-RT rainfall are calculated for different percentile ranges of discharge at different lead times. Table 5 presents the *PBias* values for forecasting of high (>90th percentile) and moderate (between the 75th and 90th percentile) flows separately. As envisaged from Table 5, for the moderate flow range, the WNARX and NARX models forecast the discharge with *PBias* < ±10% (acceptable) by the TRMM-RT rainfall (showing underestimation). For the high discharges, both the models forecast 1-day ahead discharge with *PBias* < 10%, and for 2- and 3-days ahead forecasts it is with *PBias* > 14% (showing overestimation). Table 5 also reveals that, in general, the WNARX model tends to give a better performance as compared to the NARX model for both the flow ranges.

The NARX and WNARX solutions for the two extreme flood events at 1–3 days lead times with Option [4] of input data scenario are illustrated in Fig. 11, in which the corresponding IMD observed rainfall is plotted in the secondary axis. It can be surmised from Fig. 11 that the observed rainfall allowed a better fit of the models than the satellite rainfall forcings for peak flow forecasts. The event [1] is best reproduced by the WNARX + IMD-driven forecast with slight over-prediction (*Epeak* = 0.91%). However, the quality of simulation of event [1] by the NARX model is not good as it produces a positive phase lag of one day. In the NARX model, since the meteorological inputs are not accounted with long-term seasonality, there is every possibility that the network fits according to the recent input scenarios. This may cause the shifting of the flow estimates by the model. Conversely, the time to peak error is well-addressed by the WNARX model (*Et* = 0%). The peak discharge in event [2], which seems to be a sharp peak, is well captured by the NARX and WNARX models with zero time to peak error (*Et* = 0%). The NARX + TRMM-RT combination best-captures the peak flow (*Epeak* = −6.69%) at 1-day lead time. The TRMM-RT rainfall, although, is underestimated across the high range of intensities (*PBias* = −19.20%), still reliable

forecast of peak flow is obtained by the WNARX and NARX models. Further, the NARX + TRMM-driven simulations produce the 2-day ahead peak flow forecast with very small underprediction (*Epeak* = −0.91%) (Fig. 11). At 1-day lead-time, the WNARX model provides significant peak forecast of *Epeak* = −10.16% and *Epeak* = 3.75% with IMD and TRMM-driven simulations, respectively. Basically, the peak events considered herein are not captured by the TRMM-RT rainfall, as for high rainfall thresholds, the TRMM-RT rainfall shows more discrepancy (*POD* of <20%) than the TRMM rainfall (Fig. 5). It is noticed that most of the peak flow forecasts are within acceptable limit while using the satellite-based rainfall products driven with the WNARX and NARX models. However, the magnitude of error increases for both the TRMM and TRMM-RT-driven simulations at 3-day lead times, although there is no error in time to peak. Once again these results demonstrate that, for real-time flood forecasting with the hybrid WNARX model, the satellite-based rainfall offers consistent performances similar to the observed rainfall.

In this study, the time lagged observed discharge is beyond the requirement of the model inputs. Both the TRMM and TRMM-RT-based rainfall–runoff modeling give better insight when used standalone. Hence, the short-term autoregressive features of the runoff process are dominant in the WNARX and NARX models to improve the model performances across the whole flow regime as well as the peak events. As far as the validation projects suggested (Huffman and Bolvin, 2013), the quality of rainfall estimates accords with the advanced calibration of the microwave sensors applied to the TMPA products. However, from February 27, 2014, the TMPA satellite has been transitioned to the Global Precipitation Measurement (GPM) with improved sensitivity to detect low precipitation rates (<0.5 mm/h) and falling snow (Hou et al., 2014). A few preliminary assessments of the GPM rainfalls over

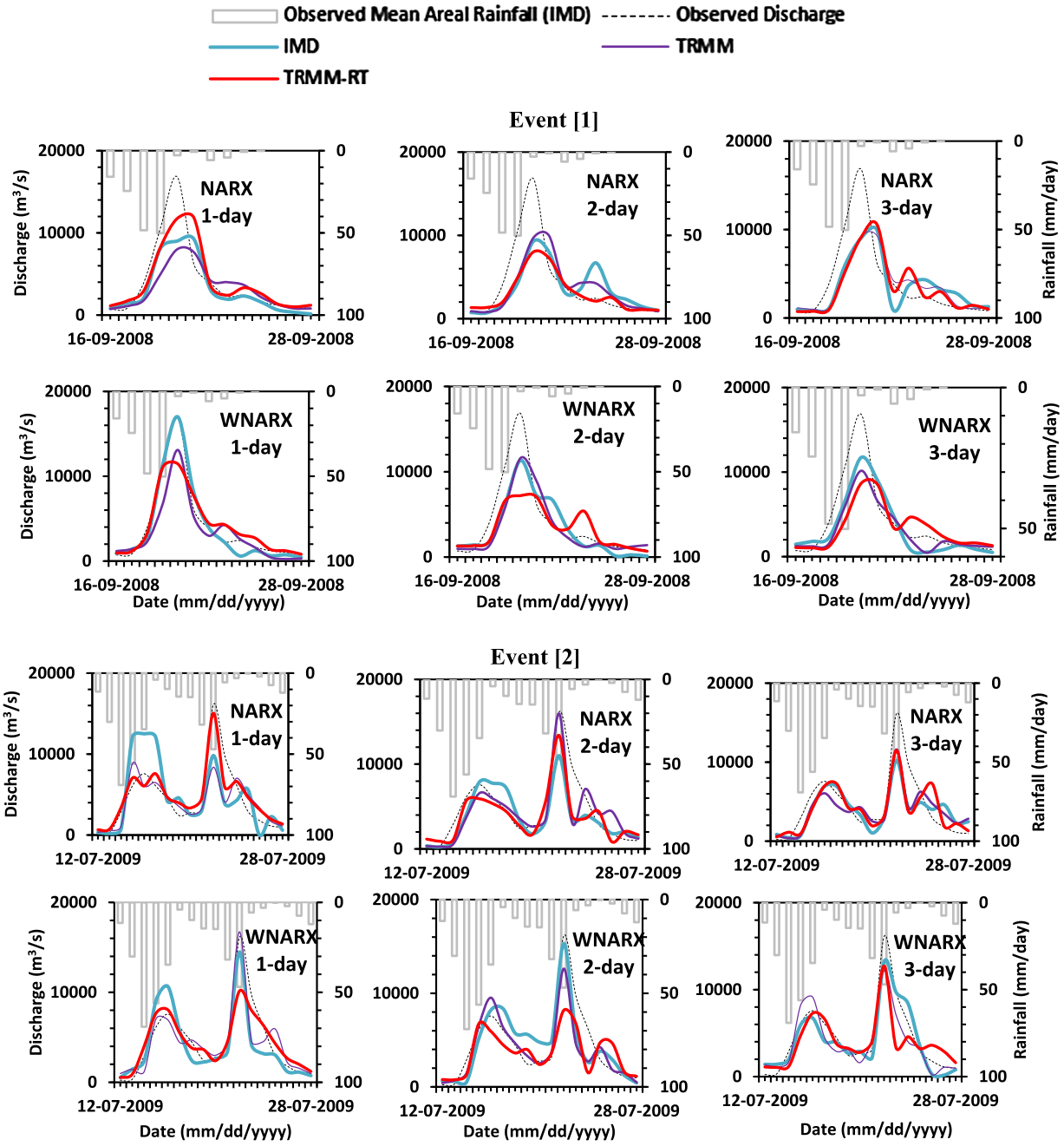

Fig. 11. Reproduction of two peak flood events at different lead times by the NARX and WNARX models.

India and China (Prakash et al., 2015; Liu, 2015; Tang et al., 2016) suggested an improvement in extreme rainfall detection, reduced false alarms, better reproducibility of probability density functions at various precipitation intensities and improved representation of precipitation diurnal cycles. Hence, it is expected that with the improved GPM-based rainfall products, the streamflow forecast capabilities of the WNARX and NARX models would be further enhanced. Moreover, the utility of rainfall forecasts from weather models, such as, the ensemble forecast of the European Center for Medium-range Weather Forecast (ECMWF) may improve the streamflow forecast for higher lead times. As a future scope of study, as carried out by Yuan et al. (2014, 2015) and Emerton et al. (2016), the medium as well as long-term rainfall forecasts could be evaluated in the developed data-driven models for improved streamflow forecasting.

5. Conclusions

To address the issues of limited or non-availability of real-time rainfall and discharge data, and satellite-based highly biased rainfall data products, a novel WNARX model is developed in this study using the hybrid concepts of wavelet transform and dynamic NN. For real-time flood forecasting in the upper Mahanadi River Basin at Basantpur gauging station, the efficacy of the satellite-based rainfall products, such as, TRMM and TRMM-RT, and observed rainfall product of IMD are tested for their use in the ARMAX, ANN, WANN, NARX and WNARX models. The results of this study reveal the following conclusions:

1. As compared to the IMD gauge-based rainfall, the satellite-based TRMM and TRMM-RT rainfall products have significant bias for low, medium and high rainfall magnitudes. The after-real-time estimate of TRMM shows more compatibility with the IMD observed rainfall as it is a gauge-rainfall-adjusted estimate.
2. Regardless of the over- or under-estimations, the satellite-based rainfall products-driven NN-models very well captured the non-linear rainfall–runoff relationship of the basin studied herein. In essence, the model performances based on the available rainfall products vary in the order of: WNARX > NARX > WANN > ANN > ARMAX.
3. Interestingly, the quality of streamflow estimation by the TRMM-RT rainfall product is highly appreciated using the proposed hybrid wavelet-based dynamic NN model in simulation mode. This efficiency of the WNARX model with the highly biased satellite products may be attributed to the inherent nature of the non-linear transfer function which internally corrects the rainfall bias when provided as the feedback to the dynamic networks. Conversely, the conventional ARMAX model does not cope up with the non-linear rainfall–runoff process efficiently with the satellite-based forcings.
4. For simulating the peak floods, both the static ANN and WANN models are not reliable. With the WNARX and NARX models, the TRMM-RT product could be very much useful for operational flood forecasting due to its real-time dissemination.
5. Ultimately, it is concluded that the TRMM-RT rainfall does not require any bias correction if used in the dynamic NN-based rainfall–runoff models to predict river flow. Moreover, the performance of the WNARX model is quite satisfactory for flood event simulations as well as peak flow forecasting with the TRMM-RT rainfall in comparison with other benchmark models. Hence, the WNARX model driven by the highly biased satellite-based TRMM-RT rainfall product could be reliably used for real-time streamflow forecasting up to 3-days lead-time, although the NARX model also performs almost equally well.
6. Overall, the developed WNARX hybrid model with only satellite-based rainfall inputs performs as reliably as with observed rainfall for multi-step-ahead streamflow forecasting. However, for generality, these models have to be tested in other world-river basins since the satellite rainfall products have varied bias levels depending on the topography of the basin.

Acknowledgements

The provision of hydro-meteorological data by Central Water Commission (CWC) and India Meteorological Department (IMD) is highly appreciated. The TRMM and TRMM-RT products are downloaded from <ftp://disc2.nascom.nasa.gov/data/s4pa/TRMM_L3/TRMM_3B42> and <ftp://trmmopen.gsfc.nasa.gov/pub/merged/3B42RT> and the temperature data is obtained from <http://rda.ucar.edu/datasets/ds093.1/>. Moreover, the research funding to the first author provided by the Ministry of Human Resources Department (MHRD), Government of India, is highly acknowledged.

References

Akhtar, M.K., Corzo, G.A., van Andel, S.J., Jonoski, A., 2009. River flow forecasting with artificial neural networks using satellite observed precipitation pre-processed with flow length and travel time information: case study of the Ganges river basin. Hydrol. Earth Syst. Sci. 13, 1607–1618.

Artan, G., Gadain, H., Smith, J.L., Asante, K., Bandaragoda, J., Verdin, J.P., 2007. Adequacy of satellite derived rainfall data for stream flow modeling. Nat. Hazards 43, 167–185.

Asante, K.O., Dezanove, R.M., Artan, G.A., Lietzow, G.R., Verdin, J., 2007. Developing a flood monitoring system from remotely sensed data for the Limpopo Basin. IEEE Trans. Geosci. Rem. Sens. 45, 1709–1714.

ASCE, 2000a. Artificial neural networks in hydrology. I: Preliminary concepts. J. Hydrol. Eng. 5 (2), 115–123.

ASCE, 2000b. Artificial neural networks in hydrology II: hydrologic applications. J. Hydrol. Eng. 5 (2), 124–137.

Badrzadeh, H., Sarukkalige, R., Jayawardena, A.W., 2015. Hourly runoff forecasting for flood risk management: application of various computational intelligence models. J. Hydrol. 529 (3), 1633–1643.

Berthet, L., Andréassian, V., Perrin, C., Javelle, P., 2009. How crucial is it to account for the Antecedent Moisture Conditions in flood forecasting? Comparison of event-based and continuous approaches on 178 catchments. Hydrol. Earth Syst. Sci. 13, 819–831.

Besaw, L.E., Rizzo, D.M., Bierman, P.R., Hackett, W.R., 2010. Advances in ungauged streamflow prediction using artificial neural networks. J. Hydrol. 386 (1–4), 27–37.

Cannas, B., Fanni, A., See, L., Sias, G., 2006. Data preprocessing for river flow forecasting using neural networks: wavelet transforms and data partitioning. Phys. Chem. Earth Parts A/B/C 31 (18), 1164–1171.

Central Water Commission, 1989. Manual on Flood Forecasting. River Management Wing, New Delhi.

Chang, F.J., Chen, Y.C., 2001. A counterpropagation fuzzy-neural network modeling approach to real time streamflow prediction. J. Hydrol. 245 (1–4), 153–164. http://dx.doi.org/10.1016/S0022-1694(01)00350-X.

Chang, F.J., Chen, P.A., Lu, Y.R., Huang, E., Chang, K.Y., 2014a. Real-time multi-step-ahead water level forecasting by recurrent neural networks for urban flood control. J. Hydrol. 517, 836–846. http://dx.doi.org/10.1016/j.jhydrol.2014.06.013.

Chang, L.C., Shen, H.Y., Chang, F.J., 2014b. Regional flood inundation nowcast using hybrid SOM and dynamic neural networks. J. Hydrol. 519, 476–489. http://dx.doi.org/10.1016/j.jhydrol.2014.07.036.

Coulibaly, P., Anctil, F., Bobee, B., 2000. Daily reservoir inflow forecasting using artificial neural networks with stopped training approach. J. Hydrol. 230 (3–4), 244–257.

Emerton, R., Stephens, E.M., Pappenberger, F., Pagano, T.C., Weerts, A.H., Wood, A.W., Salamon, P., Brown, J.D., Hjerdt, N., Donnelly, C., Baugh, C.A., Cloke, H.L., 2016. Continental and global scale flood forecasting systems. Wiley Int. Rev.: Water 3, 391–418. http://dx.doi.org/10.1002/wat2.1137.

Ezer, T., Atkinson, L.P., 2014. Accelerated flooding along the U.S. East Coast: on the impact of sea-level rise, tides, storms, the Gulf Stream, and the North Atlantic oscillations. Earth's Future 2, 362–382. http://dx.doi.org/10.1002/2014EF000252.

Froidevaux, P., Schwanbeck, J., Weingartner, R., Chevalier, C., Martius, O., 2015. Flood triggering in Switzerland: the role of daily to monthly preceding precipitation. Hydrol. Earth Syst. Sci. 19, 3903–3924.

Hopson, T.M., Webster, P.J., 2010. A 1–10-day ensemble forecasting scheme for the major river basins of Bangladesh: forecasting severe floods of 2003–07. J. Hydrometeorol. 11, 618–641.

Hou, A.Y., Kakar, R.K., Neeck, S., Azarbarzin, A.A., Kummerow, C.D., Kojima, M., Oki, R., Nakamura, K., Iguchi, T., 2014. The global precipitation measurement mission. Bull. Am. Meteorol. Soc. 95, 701–722. http://dx.doi.org/10.1175/BAMS-D-13-00164.1.

Huffman, G.J., Bolvin, D.T., Nelkin, E.J., Wolff, D.B., Adler, R.F., Gu, G., Hong, Y., Bowman, K.P., Stocker, E.F., 2007. The TRMM multisatellite precipitation analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. J. Hydrometeorol. 8 (1), 38–55. http://dx.doi.org/10.1175/JHM560.1.

Huffman, G.J., Bolvin, D., 2013. TRMM and other data precipitation data set documentation. Tech. Rep., Mesoscale Atmospheric Processes Laboratory, NASA Goddard Space Flight Center and Science Systems and Applications Inc.

Jena, P.P., Chatterjee, C., Pradhan, G., Mishra, A., 2014. Are recent frequent high floods in Mahanadi basin in eastern India due to increase in extreme rainfalls? J. Hydrol. 517, 847–862.

Khu, S.T., Liong, S.Y., Babovic, V., Madsen, H., Muttil, N., 2001. Genetic programming and its application in real-time flood forecasting. J. Am. Water Resour. Assoc. 36 (2), 439–452.

Kisi, A., 2009. Wavelet regression model as an alternative to neural networks for monthly streamflow forecasting. Hydrol. Process. 23 (25), 3583–3597.

Kneis, D., Chatterjee, C., Singh, R., 2014. Evaluation of TRMM rainfall estimates over a large Indian river basin (Mahanadi). Hydrol. Earth Syst. Sci. http://dx.doi.org/10.5194/hess-18-2493-2014.

Kumar, D.N., Raju, K.S., Sathish, T., 2004. River flow forecasting using recurrent neural networks. Water Resour. Manage. 18 (143–161), 2004.

Lammers, R.B., Shiklomanov, A.I., Vorosmarty, C.J., Fekete, B.M., Peterson, B.J., 2001. Assessment of contemporary Arctic river runoff based on observational discharge records. J. Geophys. Res. 106, 3321–3334.

Lin, T.N., Horne, B.G., Tino, P., Giles, C.L., 1996. Learning long-term dependencies in NARX recurrent neural networks. IEEE Trans. Neural Networks 7 (6), 1329–1337.

Liu, Z., 2015. Comparison of integrated multi-satellite retrievals for GPM (IMERG) and TRMM multi-satellite precipitation analysis (TMPA) monthly precipitation products: initial results. J. Hydrometeorol. 151216150524009. http://dx.doi.org/10.1175/JHM-D-15-0068.1.

Mohammadi, K., Eslami, H.R., Dardashti, Sh.D., 2005. Comparison of regression ARIMA and ANN models for reservoir inflow forecasting using snowmelt equivalent (a case study of Karaj). J. Agric. Sci. Technol. 7, 17–30.

Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Binger, R.L., Harmel, R.D., Veith, T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. ASABE 50 (3), 885–900. http://dx.doi.org/10.13031/2013.23153.

Morlet, J., Arens, G., Fourgeau, E., Glard, D., 1982. Wave propagation and sampling theory – Part I: Complex signal and scattering in multilayered media. Geophysics 47 (2), 203–221. http://dx.doi.org/10.1190/1.1441328.

Napolitano, G., See, L., Calvo, B., Savi, F., Heppenstall, A., 2010. A conceptual and neural network model for real-time flood forecasting of the Tiber River in Rome. Phys. Chem. Earth 35, 187–194.

Nash, J.E., Sutcliffe, J.V., 1970. River flow forecasting through conceptual models. I. J. Hydrol. 10, 282–290.

Nourani, V., Komasi, M., Mano, A., 2009. A multivariate ANN-wavelet approach for rainfall–runoff modeling. Water Resour. Manage. 23, 2877–2894.

Nourani, V., Baghanam, A.H., Adamowski, J., Gebremichael, M., 2013. Using self-organizing maps and wavelet transforms for space-time pre-processing of satellite precipitation and runoff data in neural network based rainfall–runoff modeling. J. Hydrol. 476, 228–243.

Prakash, S., Mitra, A.K., Pai, D.S., AghaKouchak, A., 2015. From TRMM to GPM: how well can heavy rainfall be detected from space? Adv. Water Resour. 5 (5). http://dx.doi.org/10.1016/j.advwatres.2015.11.008.

Perumal, M., Sahoo, B., 2007. Limitations of real-time models for forecasting river flooding from monsoon rainfall. Nat. Hazards 42 (2), 415–422. http://dx.doi.org/10.1007/s11069-006-9082-1.

Perumal, M., Moramarco, T., Barbetta, S., Melone, F., Sahoo, B., 2011. Real-time flood-stage forecasting by variable parameter Muskingum stage hydrograph routing method. Hydrol. Res. (Formerly Nord. Hydrol.) 42 (2–3), 150–161. http://dx.doi.org/10.2166/nh.2011.063.

Petra, S.F., Naef, F., 2010. More frequent flooding? Changes in flood frequency in Switzerland since 1850. J. Hydrol. 381, 1–8.

Sehgal, V., Sahay, R.R., Chatterjee, C., 2014a. Effect of utilization of discrete wavelet components on flood forecasting performance of wavelet based ANFIS models. Water Resour. Manage. 28, 1733–1749.

Sehgal, V., Tiwari, M.K., Chatterjee, C., 2014b. Wavelet bootstrap multiple linear regression based hybrid modeling for daily river discharge forecasting. Water Resour. Manage. 28 (10), 2793–2811.

Seo, Y., Kim, S., Singh, V.P., 2015. Multistep-ahead flood forecasting using wavelet and data-driven methods. J. Civ. Eng. 19, 401–417.

Shoaib, M., Shamseldin, A.Y., Melville, B.W., Khan, M.M., 2014. Hybrid wavelet neuro-fuzzy approach for rainfall–runoff modeling. J. Comput. Civ. Eng. 1–16. http://dx.doi.org/10.1061/(ASCE)CP.1943-5487.0000457.

Singh, P., Deo, M.C., 2007. Suitability of different neural networks in daily flow forecasting. Appl. Soft Comput. 7, 968–978.

Tang, G., Ma, Y., Long, D., Zhong, L., Hong, Y., 2016. Evaluation of GPM Day-1 IMERG and TMPA Version-7 legacy products over Mainland China at multiple spatiotemporal scales. J. Hydrol. 533, 152–167. http://dx.doi.org/10.1016/j.jhydrol.2015.12.008.

Tiwari, M.K., Chatterjee, C., 2010a. Uncertainty assessment and ensemble flood forecasting using bootstrap based artificial neural networks (BANNs). J. Hydrol. 382 (1–4), 20–33.

Tiwari, M.K., Chatterjee, C., 2010b. Development of an accurate and reliable hourly flood forecasting model using wavelet–bootstrap–ANN hybrid approach. J. Hydrol. 394 (3–4), 458–470.

Tiwari, M.K., Chatterjee, C., 2011. A new Wavelet–Bootstrap–ANN hybrid model for daily discharge forecasting. J. Hydroinf. 13 (3), 500–519.

Tong, K., Su, F., Yang, D., Zhang, L., Hao, Z., 2014. Tibetan Plateau precipitation as depicted by gauge observations, reanalyses and satellite retrievals: Tibet precipitation depicted by gauge, reanalyses, and satellite data. Int. J. Climatol. 34 (2), 265–285. http://dx.doi.org/10.1002/joc.3682.

Trenberth, K.E., 1999. Conceptual framework for changes of extremes of the hydrological cycle with climate change. Clim. Chg. 42, 327–339.

Valipour, M., Banihabib, M.E., Behbahani, S.M.R., 2013. Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir. J. Hydrol. 476, 433–441.

Willmott, C.J., Matsuura, K., 2005. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Clim. Res. 30, 79-2.

Yong, B., Ren, L.L., Hong, Y., Wang, J.H., Gourley, J.J., Jiang, S.H., Chen, X., Wang, W., 2010. Hydrologic evaluation of multisatellite precipitation analysis standard precipitation products in basins beyond its inclined latitude band: a case study in Laohahe basin, China. Water Resour. Res. 46, W07542. http://dx.doi.org/10.1029/2009WR008965.

Yong, B., Ren, L.L., Hong, Y., Wang, J.H., Gourley, J.J., Jiang, S.H., Chen, X., Wang, W., 2012. Assessment of evolving TRMM-based multisatellite real-time precipitation estimation methods and their impacts on hydrologic prediction in a high latitude basin. J. Geophys. Res. 117, D09108. http://dx.doi.org/10.1029/2011JD017069.

Yuan, X., Wood, E.F., Liang, M., 2014. Integrating weather and climate prediction: toward seamless hydrologic forecasting. Geophys. Res. Lett. 41, 5891–5896. http://dx.doi.org/10.1002/2014GL061076.

Yuan, X., Wood, E.F., Ma, Z., 2015. A review on climate-model-based seasonal hydrologic forecasting: physical understanding and system development. Wiley Int. Rev.: Water 2, 523–536. http://dx.doi.org/10.1002/wat2.1088.