Inference of basin flood potential using nonlinear hysteresis effect of basin water storage: case study of the Koshi basin

Pennan Chinnasamy

ABSTRACT

Current flood forecasting tools for river basins subject to extreme seasonal monsoon rainfall are of limited value because they do not consider nonlinearity between basin hydrological properties. The goal of this study is to develop models that account for nonlinearity relationships in flood forecasting, which can aid future flood warning and evacuation system models. Water storage estimates from the Gravity Recovery and Climate Experiment, along with observed discharge and rainfall data were used to develop two multivariate autoregressive monthly discharge models. Model-I was based on rainfall only, while Model-II was based on rainfall and water storage estimates for the Koshi subbasin within the Ganges River basin. Results indicate that the saturation of water storage units in the basin play a vital role in the prediction of peak floods with lead times of 1 to 12 months. Model-II predicted monthly discharge with Nash–Sutcliffe efficiency (NSE) ranging from 0.66 to 0.87, while NSE was 0.4 to 0.85 for Model-I. Model-II was then tested with a 3-month lead to predict the 2008 Koshi floods – with NSE of 0.75. This is the first study to use 'fixed effects' multivariate regression in flood prediction, accounting for the nonlinear hysteresis effect of basin storage on floods.

Key words | flood prediction, Ganges basin, GRACE, Koshi, nonlinear regression, water storage

Pennan Chinnasamy
International Water Management Institute,
Colombo,
Sri Lanka
E-mail: *p.chinnasamy@cgiar.org*

INTRODUCTION

Developments in river discharge prediction methods, to aid flood forecasting, have been ongoing for many decades in various flood affected basins, globally (WMO 1992; Porporato & Ridofi 2001; Boughton & Droop 2003; Liao *et al.* 2003). Even though current methods cover a wide range from very simple, empirical and statistically based, completely black-box to very physically based detailed and complex conceptual models (WMO 1992), limitations exist. Limitations in long time series of hydrological observations and limitations in understanding of flow mechanisms and nonlinearity have hindered the development of statistical flood models (e.g., WMO 1992; Hsu *et al.* 1995; Raman & SunilKumar 1995; Mukherjee & Mansour 1996; Chau *et al.* 2005; Reager *et al.* 2014, 2015; Sproles *et al.* 2015) and physically based simulation models (e.g., Garrote & Bras 1995; Refsgaard & Knudsen 1996; Campolo *et al.* 1999). However, only some methods have explored the site-specific nonlinearities in river discharge mechanisms (Porporato & Ridofi 2001; Liao *et al.* 2003).

The relationship between monsoon rainfall and discharge mechanisms may be nonlinear in many basins, e.g., in basins where the contribution of storage is important to discharge (Moore *et al.* 2006; Reager & Famiglietti 2009; Andermann *et al.* 2012; Reager *et al.* 2014). Porporato & Ridofi (2001) indicated that the incorporation of nonlinearities has increased the capability of their models in modeling the dynamics of river flow and in practical river discharge forecasting. The complexity of nonlinear modeling increases with the number of properties the model is dependent on, when the model is dependent on one property (single

doi: 10.2166/nh.2016.268

variate, e.g., rainfall) or multiple properties (multivariate, e.g., combination of soil moisture, antecedent discharge, rainfall, etc.). However, due to the challenges in identifying and modeling the various discharge mechanisms, there remain limitations in current methods (Porporato & Ridofi 2001), especially in site-specific monsoonal flood analysis (e.g., regions in the Ganges basin). If such studies are successful in effectively predicting discharge and floods, their methodology in identifying multivariate mechanisms can then be incorporated in other regions with similar severe monsoonal floods. However, data availability is the key concern in this type of modeling.

In recent years, remote sensing platforms have aided in estimation of various hydrological variables at the basin scale that can be used to improve flood prediction tools for regions where observation data are scarce or unavailable. Access to such remotely sensed data is less costly than field measurements. Remotely sensed data from the Gravity Recovery and Climate Experiment (GRACE) satellite mission have been successful in estimating terrestrial water storage (TWS) trends in many regions of the globe (see, e.g., Rodell *et al.* 2007; Famiglietti & Rodell 2013), especially for the Indian subcontinent (see, e.g., Rodell *et al.* 2009; Chinnasamy *et al.* 2013; Chinnasamy & Agoramoorthy 2015). Using GRACE, TWS changes were estimated for Gujarat (Chinnasamy *et al.* 2013), North Eastern India (Rodell *et al.* 2009), and Tamil Nadu (Chinnasamy & Agoramoorthy 2015) and suitable groundwater resource management plans were recommended. Since TWS variations are either unavailable or too expensive in the Koshi basin, GRACE storage estimates are employed here.

The primary objective of this study is to introduce a new unique river forecasting model (for a basin affected by severe monsoon floods) based on explanatory variables including the traditional hydrological variables; rainfall and discharge as well as water storage estimates obtained from remote sensing (GRACE) data. Improvement in the accuracy of the model due to the inclusion of water storage estimates will be assessed. To achieve this, a base model (Model-I) will be identified, using only observed rainfall data, and its accuracy assessed against observed discharge. The model is coupled with GRACE TWS data (Model-II) and evaluated against observed data to quantify the improvement in model predictions for the Koshi subbasin (as a case study) of the Ganges River basin. In addition, the ability of the models to predict the 2008 Koshi floods, with 1 to 12 month lead time, will be evaluated. Finally, a discussion will be provided on how the resulting model can be used by land managers and disaster management teams to both aid in future flood prediction models and to quantify the value of groundwater flood storage for flood management. This study will be the first to employ 'fixed effects' multivariate nonlinear regression modeling to develop a flood prediction model that accounts for the nonlinear hysteresis (unlike many past models that consider linearity) effect of basin storage on flood discharge. This model framework provides the ability to predict river discharge in regions with similar flow mechanisms and regions affected by seasonal floods and in regions with limited observation data.

STUDY AREA

Of the Himalayan rivers that flow into the Ganges, the Koshi is the largest and has the world's largest river-built alluvial fan extending to 180 km long and 150 km wide. The Koshi River basin is a trans-boundary basin encompassing portions of the following countries: China, Nepal, and India. Of the total catchment area of 87,311 km^2, 33% (28,300 km^2) is located in China, 45% (39,407 km^2) in Nepal, and 22% (19,604 km^2) in India, respectively (Figure 1). The Koshi River generally flows south and merges with the Ganges at Katihar district in Bihar, India, after flowing a total distance of 729 km from source to confluence (Reddy *et al.* 2008).

The Koshi basin in Nepal has an elevation drop from the divide at 8,848 m (Mt. Everest) to 60 m, in the alluvial plains. Due to the varying climate and topography, the precipitation is unevenly distributed in both space and time, varying from 1,755 mm in the central mountain to 210 mm in the trans-mountain region. The eastern mountain region has 1,418 mm yr^{-1} rainfall on average, and the southern part of the basin is, on average, wetter than the trans-Himalayan northern part. The climate varies from a cold north to a tropical south. The basin has four major seasons: pre-monsoon (March to May), monsoon (June to

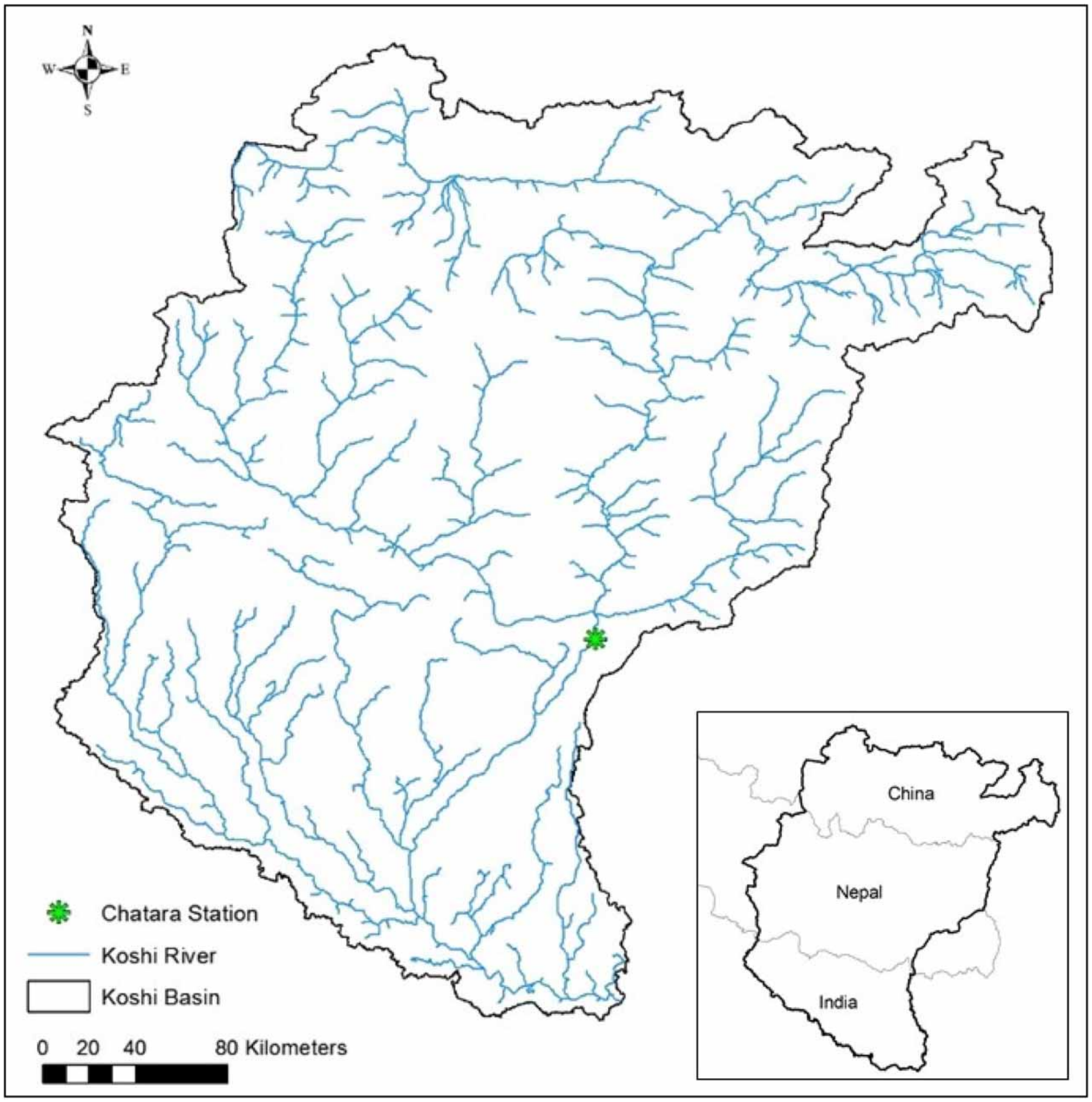

Figure 1 | Koshi basin with the location of Chatara stream gauge station. Inset shows the location of the Koshi basin with respect to China, Nepal, and India.

September), post-monsoon (October to November), and winter (December to February).

Floods are one of the most common natural disasters in the Ganges basin, and are caused by intense seasonal monsoon precipitation (from June to September). Flood frequency and intensity in the Ganges have increased considerably over the past decade (Fushimi *et al.* 1985; Mool *et al.* 2001; Shrestha & Bajracharya 2013). Dhar & Nandargi (2002) listed 10 Ganges basin rivers that were particularly affected by frequent floods – a total of 1,300 flood events occurred in these rivers over the last 14 years of the 20th century. In addition, the confluence of the Ganges and the Brahmaputra, in Bangladesh, is subject to annual flooding (Shrestha & Bajracharya 2013).

Of the Ganges tributaries, the Koshi River is the most flood prone (Reddy *et al.* 2008). The highest Koshi discharge, of 24,000 $m^3 s^{-1}$, was recorded during the 24 August, 1954 flood event (Jain *et al.* 2007). Responding to the devastation caused by the 1954 Koshi floods, the government of India and the government of Nepal built a barrage at Bhimnagar (located in Nepal) to tame the Koshi river floods. However, Bihar state was still affected by Koshi floods in 1978, 1987, 1998, 2004, and 2008, thus earning the name Bihar's 'river of sorrow' (Jain *et al.* 2007). The 1998, 2004, and 2008 floods caused a total loss of 1.1, 1.4, and 0.5 billion USD due to public property and crop damage.

Despite the aforementioned damage and socioeconomic impacts on the Koshi region, floods have not received adequate attention from policy and management perspectives (Shrestha & Bajracharya 2013; Chinnasamy *et al.* 2015a). Flood prediction in the region is at the early stage and is mostly represented by warnings that occur a few hours before the actual flood (Reddy *et al.* 2008; Shrestha & Bajracharya 2013). Annual floods can be predicted using past trends in hydrological and climatological data; however, flood forecasting models are limited in regions where severe monsoon rainfall flooding is prevalent (Guo *et al.* 2012).

METHODOLOGY

Stream flow and rainfall data

Monthly river discharge data from Chatara (Station number 695) were obtained from the Nepal Department of Hydrology and Meteorology (DHM) for the 11 year period: 2000 to 2010. Of the various existing DHM stations, only data from Chatara (Figure 1) were used, because this station is located on the main stem of the Koshi River, and hence captured the high flow conditions better than the other gauge stations. For the purposes of this initial study, daily average discharge data were converted to monthly averages. The stream discharge data are not available from India, and thus the downstream location most used for this model was in Nepal. Therefore, the effective study area was reduced from 87,311 km^2 to 67,707 km^2.

Monthly rainfall data were collected from 127 DHM rain gauge stations across the Koshi basin. Monthly DHM data were used, as GRACE is at monthly resolution. DHM collects rainfall at daily intervals, and hence the data were converted to monthly totals for each station. Finally, the monthly rainfall for the Koshi basin was estimated from the average (using Theissen polygon method) of all the stations in the basin. No kriging or isohyetal analysis was done.

Flood source identification

Developing a better understanding of flood components will aid in formulating proactive flood management tools (e.g., models). This improved understanding combined with the time lagged flood prediction models can greatly reduce future flood impacts (Mool *et al.* 2001). Figure 2 illustrates a hypothetical plot of rainfall versus discharge, which can be used to understand the hydrological processes for stream flow generation. In this conceptual model, any deviations from linearity (i.e., hysteresis) indicate that other hydrological processes in addition to rainfall, contribute to the discharge. An anticlockwise hysteresis loop in the discharge versus rainfall plot implies that a considerable contribution from groundwater storage to stream discharge exists (as shown in Figure 2). Groundwater storage fills during the pre-monsoon months, and as a result, less rainfall is transferred to baseflow discharge in pre-monsoon months. During post-monsoon months, groundwater storage contributes to discharge in addition to contributions from rainfall. Thus, higher discharge is noted (nonlinearity) producing an anticlockwise hysteresis loop. In such instances, the groundwater storage peaks before the peak discharge, and as a result, floods occur when the groundwater flux contributes with high rainfall. On the other hand, a clockwise hysteresis loop indicates rapid increase in discharge with increase in rainfall, thus indicating a flashy system. The relationship between rainfall and discharge is not homogenous or linear in the Koshi basin due to strong

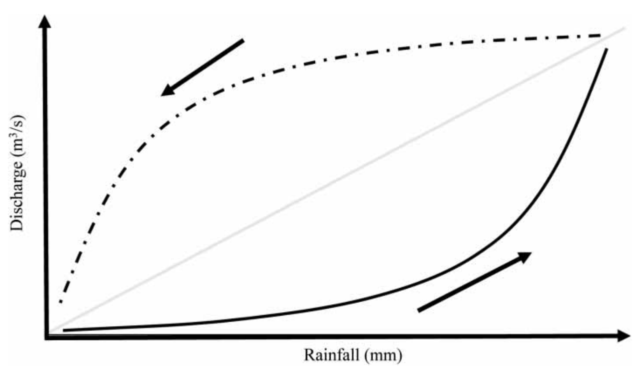

Figure 2 | Explanatory diagram for the catchment hysteresis anticlock-wise loop, with the solid line indicating first half of the hysteresis leg (prior to monsoon) and the dashed line indicating the second half of the hysteresis leg (post-monsoon). Any deviation from the grey solid linear line indicates that other sources (e.g., groundwater) are contributing to discharge.

precipitation seasonality (Barros *et al.* 2006; Bookhagen & Burbank 2010; Andermann *et al.* 2012). Hence, basin-wide monthly rainfall data and monthly discharge data were used to plot discharge versus rainfall graphs to estimate hysteresis trends. In particular, the 2004 and 2007/8 floods were analyzed using this method.

TWS data from GRACE

The GRACE mission was launched by combined efforts from the National Aeronautics and Space Administration (NASA) and the German Aerospace Center Deutschen Zentrum für Luftund Raumfahrt (DLR) on the 17 March, 2002. GRACE has been estimating TWS from April 2002 to date. The GRACE mission, with two satellites working in unison, is the first platform that can estimate global scale TWS at monthly intervals. The global mass anomalies (due to changes in groundwater and surface water volume) cause changes to the gravitational field, which, in turn, causes perturbations to the GRACE satellite orbits. While one satellite undergoes changes in the orbit, the other satellite records these perturbations as changes in TWS in units of equivalent water thickness in centimeters (Swenson & Wahr 2006; Landerer & Swenson 2012). The GRACE TWS data include storage due to the sum of snow water, soil water (soil moisture), surface water (e.g., dams) and in groundwater (both deep and shallow).

A scaled version of the GRACE data, with gain factors provided by Landerer & Swenson (2012), were used here. The current GRACE data are based on the level V (RL05) spherical harmonics and have improved leakage and measurement errors and a spatial resolution (1° by 1°), which were limited in previous GRACE versions (Landerer & Swenson 2012), thus making GRACE estimates feasible for regional and even watershed scale water storage analysis (e.g., as in Billah *et al.* 2015). Further information regarding the GRACE data solutions (Cheng *et al.* 2011), degree one coefficients (Swenson *et al.* 2008), and glacial isostatic adjustments (Geruo *et al.* 2013) performed on the data is provided in the GRACE data download page. Monthly TWS data, in 1° by 1° grid-cells format (with 100 by 100 km resolution near the equator), are available from the NASA Jet Propulsion Laboratory, and can be accessed online at the NASA website (ftp://podaac-ftp.jpl.nasa.gov/allData/tellus/L3/land_mass/). It is to be noted that the GRACE data are only available as monthly anomalies with a baseline average (from January 2004 to December 2009) removed. The baseline was introduced by the GRACE data processing team to understand long-term TWS trends. There are now many studies which have validated GRACE water storage data (e.g., Rodell *et al.* 2009; Chinnasamy *et al.* 2013, 2015b, 2015c; Chinnasamy & Agoramoorthy 2015; Chinnasamy & Sunde 2015; Chinnasamy & Agoramoorthy 2016).

Stream discharge prediction using time lagged-autoregressive model

Since a nonlinear hysteresis is expected in basins with streamflow contributions from water storage units, nonlinear relationships are considered here. Initially, the following autoregressive model is considered:

$$Q(t) = \beta_o P(t)^{\beta_1} \tag{1}$$

Applying logarithms we obtain:

$$\ln(Q(t)) = \ln(\beta_o) + \beta_1 \ln(P(t)) \tag{2}$$

where $Q(t)$ is the discharge in month t, $P(t)$ is precipitation, the intercept term is $ln(\beta_0)$ and the slope term is β_1. As noted in Figure 2, the catchment hysteresis may have two legs – one prior to monsoon and the other post-monsoon. Thus, the exponent β_1 in Equation (1) may change from positive as the floods increase from January to September, to negative, as the season progresses back to December. Therefore, a dummy variable, *Season*, is introduced in Equation (2), which will take a 0 value for months from January to August and 1 for September to December. The resulting equation is of the form:

$$\ln(Q(t)) = \ln(\beta_o) + \beta_1 \ln(P(t)) + \beta_2 Season + \beta_3 Season * \ln \mathrm{P(t)} \tag{3}$$

where, β_2 and β_3 are model coefficients.

To analyze the potential value of adding other explanatory variables to model-I, such as TWS and rainfall as stream discharge predictors, the following model was considered:

$$\ln(Q(t)) = \ln(\beta_o) + \beta_1 \ln(P(t)) + \beta_2 Season + \beta_3 Season* \ln \mathrm{P(t)} + \beta_4 \ln(\mathrm{TWSA}) + \beta_5 \mathrm{Season}* \ln(\mathrm{TWSA}) \quad (4)$$

where *TWSA* is the basin wide average of the GRACE terrestrial water storage anomaly, β_4 and β_5 are model coefficients.

Lead time in months was introduced in Equations (3) and (4) to investigate if discharge can be predicted ahead of time, as follows:

$$\ln(Q(t)) = \ln(\beta_o) + \beta_1 \ln(P(t-\tau)) + \beta_2 Season + \beta_3 Season* \ln \mathrm{P}(\mathrm{t}-\tau) \quad (5)$$

$$\ln(Q(t)) = \ln(\beta_o) + \beta_1 \ln(P(t-\tau)) + \beta_2 Season + \beta_3 Season* \ln \mathrm{P}(\mathrm{t}-\tau) + \beta_4 \ln(\mathrm{TWSA}(\mathrm{t}-\tau)) + \beta_5 Season* \ln(\mathrm{TWSA}(\mathrm{t}-\tau)) \quad (6)$$

where, τ is lead time in months and takes the values 1 to 12 months, depending on the lead time prediction needed.

RESULTS AND DISCUSSION

Rainfall, discharge, and GRACE TWSA

The annual average rainfall for the Koshi basin over the study period (2000 to 2010) indicated no significant change in the rainfall pattern. Hence, the time period used for the analysis does not appear to be marked by any unusual changes in annual rainfall, and thus the study results can be used to predict floods during other average rainfall periods. A trend analysis on the discharge and TWSA time series also showed negligible trends over the study period.

Flood source identification

A plot of monthly TWSA versus discharge is shown in Figure 3. The lower envelope of the relationship in Figure 3 indicates the minimum discharge response for a given storage anomaly. It was noted that higher water storage was mostly associated with peak flows in the basin. The relationship between TWSA and discharge can thus be used to proactively identify and potentially discharge the excess storage (e.g., by pumping for agricultural use) in pre-monsoon months in order to reduce the magnitude of peak flood in the subsequent monsoon months. In order to confirm the importance of the water storage component in the Koshi discharge, the hysteresis behavior was analyzed using Equations (4) and (5).

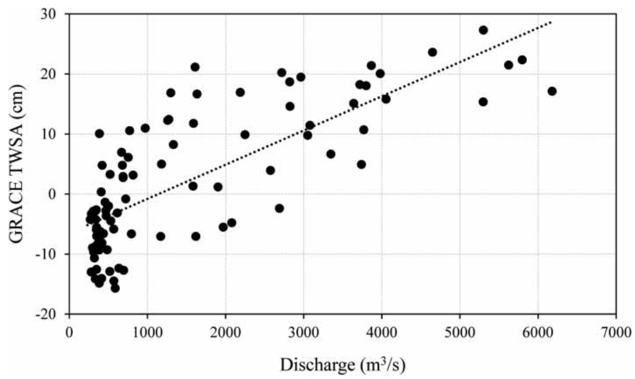

Figure 3 | Monthly GRACE TWSA versus monthly discharge for the Koshi basin from 2002 to 2009.

The chronology of the rainfall and discharge exhibits a well-defined trend during pre-flood and flood months in the Koshi basin. The discharge versus rainfall plot for the 2004 flood (Figure 4) was plotted using data from January to December 2004. The plot shows an overall anticlockwise hysteresis, with the discharge increasing with rainfall up until July 2004, and then slowly decreasing by December 2004. Over an annual timeframe, the monthly discharge increases with increases in rainfall during pre-monsoon

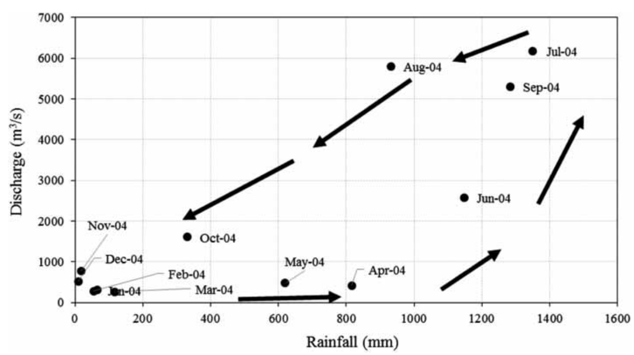

Figure 4 | Monthly discharge versus rainfall for the 2004 Koshi flood.

months (March–May) and monsoon months (June–September). During the post-monsoon (October and November) and winter months (December–February), there is a decrease in the discharge–rainfall relationship, thus showing an overall anticlockwise hysteresis loop. Hence, for the same rainfall, the discharge is higher in post-monsoon months when compared to pre-monsoon months, indicating that the groundwater storage, which filled up during the monsoon season, contributed along with the rainfall to produce high discharge.

The discharge versus rainfall plot for the 2008 flood (Figure 5) was plotted using data from January to December 2008. Similar to the 2004 floods, the 2008 season showed the discharge rising up until September 2008, and then falling rapidly by December 2008. When compared to the 2004 floods, the 2008 floods peaked a month later in August, in which the embankment breach occurred (downstream of the gauge location). In addition, the lag was lesser in 2008 when compared to that in 2004 floods, indicating the water discharge fell sharply after the peak flood occurred.

This anticlockwise hysteresis loop implies that the groundwater storage component is an important factor in flood generation for the Koshi basin. Therefore, groundwater storage should also be included while predicting floods in the Koshi basin, as in Equations (4) and (5).

The 2004 and 2008 floods had an average hysteresis (deviation from linearity) of 1,380 and 512 $m^3 s^{-1}$, respectively. This hysteresis is equivalent (Figure 5) to an average TWSA of 15 and 3 cm, respectively. Therefore, the 2004 and 2008 floods had an excess of 13 and 2 million cubic meters of water in the basin, respectively. If these volumes had been used prior to the peak flow months, the discharge would have been at nearly average levels and flooding damage would have been averted. Thus, if this excess water storage had been utilized before the build up to the flood, the flood water could recharge the storage and, as a result, the peak discharge would have decreased and the resulting flood damage reduced. Identifying suitable methods for rapid utilization of water storage before peak floods can thus become a key component in future flood mitigation plans. This also requires identifying a window of time to utilize water, for which flood prediction tools are important.

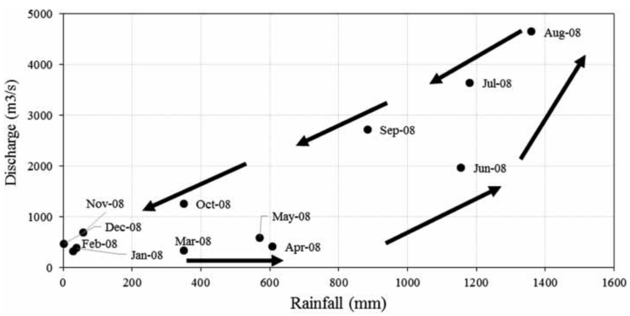

Figure 5 | Discharge versus rainfall for the 2008 Koshi flood.

GRACE flood prediction vs. observed flood flow

Table 1 and Figures 6 and 7 illustrate the goodness of fit of both models (I and II) using the Nash–Sutcliffe efficiency (NSE). Overall, Model-II (with rainfall and TWSA) had better agreement than Model-I (with only rainfall) which reinforces the importance of water storage units to stream discharge contribution for the Koshi basin. All 12 lead time models under Model-II provide satisfactory goodness of fit with NSE values in the range of 0.66 to 0.87. However, the results of the model with only a practical 3-month lead time are discussed here. The 3-month lead time can provide ample time for government agencies to proactively plan for flood management activities across the basin.

Table 1 | Model coefficients for each model at a 5-month lead-time

Lead time (months)	NSE (Model-I: rainfall)	NSE (Model-II: rainfall and TWSA)
0	0.42	0.73
1	0.56	0.68
2	0.70	0.70
3	0.62	0.75
4	0.48	0.75
5	0.50	0.70
6	0.55	0.66
7	0.64	0.69
8	0.79	0.80
9	0.85	0.87
10	0.73	0.86
11	0.47	0.76
12	0.40	0.69

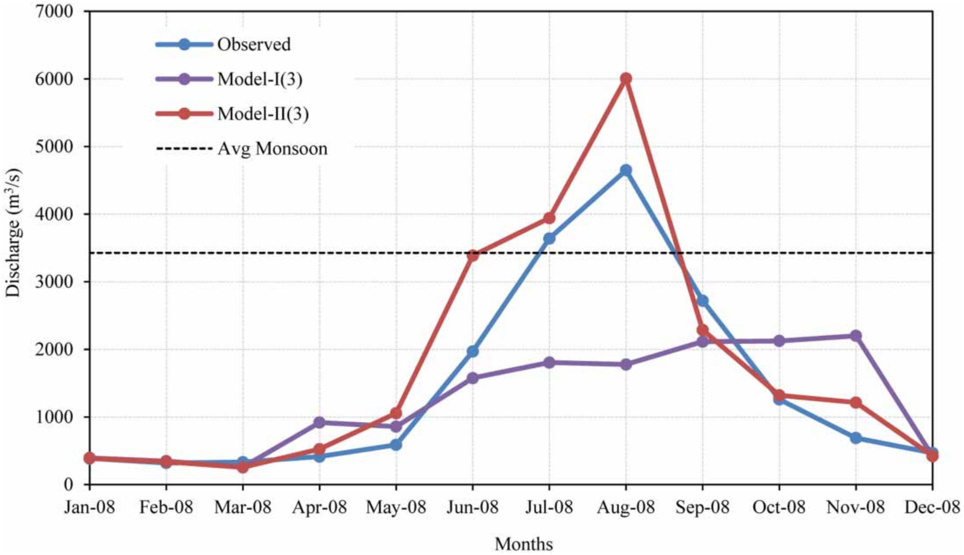

Figure 6 | Stream discharge prediction model results for the 2008 Koshi flood.

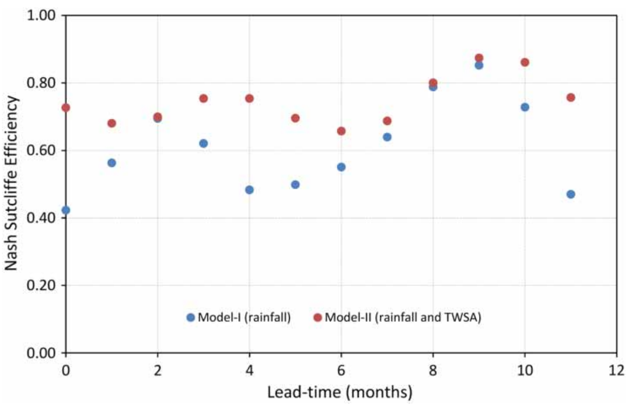

Figure 7 | NSE for Model-I and Model-II.

In order to understand the ability of the above models in predicting floods, the modeled discharge was compared against observed average monsoon (non-flood) discharge – which was estimated to be 3,428 $m^3\,s^{-1}$ (averaged from 2000 to 2010 monsoon discharge values). Therefore, the potential of flooding can be inferred if the models predict a monthly discharge above 3,428 $m^3\,s^{-1}$. The modeled monthly discharge is plotted along with actual monthly discharge for the 2008 Koshi floods in Figure 6. For the 2008 floods, Model-I could not predict the occurrence of floods, while Model-II was capable of flood prediction. Even though the predicted discharge was 1,350 $m^3\,s^{-1}$ above the observed flow (Figure 6), Model-II was successful in predicting above average monsoon discharge. This information can be useful for flood management activities to invest in proactive measures.

The above NSE analysis indicates that, of the models considered, the models which incorporate GRACE information were the best stream discharge predictive tool for the Koshi basin. Of the lead times introduced in Model-II, a 9-month lead time had the highest NSE (0.87), while a 6-month lead time had the lowest NSE (0.66). However, all NSE factors were satisfactory for the model that included GRACE TWSA estimates.

A past study by Reager *et al.* (2014) also showed the value of TWSA in discharge prediction for the Missouri River basin flood (2010–2011) and Columbia River basin flood (2011). In addition, the study indicated that the autoregressive model can be used along with Global Land Data Assimilation System's basin-wide routed runoff estimates (Rodell *et al.* 2004) in ungauged basins. However, only results from the linear regression were reported, due to the complications they faced while considering nonlinearity. Many other studies have used simulation models that required intense computing algorithms and associated computing costs. This current study used cost-effective methods while taking into account the nonlinearity. Camporese *et al.* (2014) simulated the nonlinear hydrological dynamics for an experimental mountain headwater catchment in northeastern Italy, using a process-based coupled surface–subsurface flow model. Kia *et al.* (2012) developed an artificial neural network model using MATLAB to predict floods in the Johor River basin located in the Malaysian peninsular.

The current study used the discharge and rainfall data in addition to the GRACE TWSA data to achieve a better correlation between actual and modeled discharge. The GRACE TWSA and R package, used for estimating model coefficients, was a free source and hence the current model is cost-effective and less data intensive when compared to many previous studies. These attributes and the satisfactory flood prediction capability of our model can serve as baseline models for future complex models that can effectively predict peak floods. In addition, unlike other studies that do not consider basin water storage parameters, a better understanding of the contribution of TWSA to stream discharge can be valuable for flood management.

It is also important to note that nonlinearities may differ between sites, due to differences in the dominant processes that impact discharge and floods. As a result, it is recommended that site-specific nonlinearities be deduced from the long-term discharge data and GRACE data (as conducted in the current study) before employing them in the flood prediction models.

The analysis presented in this study, for the first time, provides a novel methodology in developing a holistic fixed effect multivariate nonlinear regression stream discharge model that accounts for contributions from storage compartments (i.e., surface and groundwater stores) across the basin. This model can be applied globally to similar basins that have significant discharge contributions from water storage units, especially water flux from the groundwater storage. Potential sites, to test our methodology, will include the alpine basins of the Canadian Rockies, the Missouri River basin (Reager *et al.* 2014), the Columbia River basin (Sproles *et al.* 2015), the Indus River basin, the Ganges River basin, the Brahmaputra Basin, the Mekong River basin, and the Yellow River basin. These basins serve most of the world's population and hence flood prediction tools are of utmost importance.

LIMITATIONS AND FUTURE DIRECTIONS

Stream discharge data were not available for the downstream locations near the confluence point in the Koshi Basin, which are in India. The data used in this study were collected from the stream gauge located in Chatara (Nepal), and are the best downstream data available for the Koshi basin. As a result, the effective size of the basin is reduced from 87,311 km^2 to 67,707 km^2. GRACE data have been used earlier for smaller basins (e.g., Billah *et al.* 2015) by using the scaled version of the GRACE data. Similarly, for this study the scaled version of the GRACE data were used, which is at an average resolution of 12,300 km^2 at the equator. The error with GRACE data is low for large basins (area >200,000 km^2), while smaller basins may have significant GRACE errors while quantifying terrestrial water storage. Since the current study estimated stream discharge as a function of the GRACE estimates, it is assumed that any errors in the GRACE data would be homogenous for the study period, and thus would only alter the model coefficients. The overall goal of this study was to identify a new nonlinear regression model that best describes the relationship between groundwater storage and stream discharge. GRACE-derived water storage estimates were used in the model as they were the only data available for this region as observation data are limited as it is logistically difficult to monitor the Himalayan region, due to trans-boundary issues (e.g., international policies for data sharing, etc.). Future research is therefore needed to validate the GRACE observations in the study region before inclusion in operational applications. Such validation would require high spatiotemporal observation of water storage (both surface and groundwater) in the region, and such data are currently unavailable. Hence, there is a need to collate data from different agencies in the region and initiate new observation networks. Such data will result in scientifically validated models that can predict floods with high accuracy. Therefore, once GRACE resolution is upgraded in future missions, future studies may use this study's results as a baseline to improve the discharge model and estimate the errors (if any) due to the GRACE resolution.

The accuracy and sensitivity tests noted by previous studies (e.g., Swenson & Wahr 2006) indicate that the

GRACE data have an accuracy of ± 2 cm for the region under this study. In addition, since only the anomaly of the data were used (and are the only data available) for this study, the effect of this accuracy was assumed to be averaged out in the long-term analysis, i.e., 10 years. The accuracy and sensitivity of GRACE data are getting better with new tools and methods, which are constantly being upgraded by the GRACE team. Once the newer version of the data is available, future studies can engage in using the method described in this paper, with better version data to obtain improved results, if any.

In addition, due to the monthly resolution of GRACE data, it is to be noted that the flood analysis was done using monthly mean discharge data, which may result in averaging out of peak flow characteristics. However, there were no other data available currently that could be used for the model's objectives. The author urges future researchers to include other observation data and remote sensing data, when and if available, to understand more about the proposed methodology in the future. It is to be noted that future GRACE missions are aimed at providing data at higher temporal resolution. Therefore, this study's model can be upgraded by incorporating the future GRACE data into the nonlinear regression models.

One important future direction of this study is to serve as a baseline paper in identifying the importance of groundwater storage on peak discharge in mountainous regions, especially in Nepal, where observation data are scarce. Many new methods are being tested for flood prediction in Nepal, and therefore, their results can be compared in the future against results from this paper. In addition, since most of the methods do not rely on groundwater storage, results from this paper can be used in unison with their methods, wherein groundwater compartment might be an important factor to consider.

CONCLUSIONS

The current study highlights the importance of incorporating nonlinear methods in stream discharge prediction for the Koshi basin, which is highly affected by seasonal floods on an annual basis. The nonlinear, multivariate autoregressive models introduced here used past records of discharge, water storage, and rainfall to predict discharge with a lead time of 1 to 12 months. Importantly, the models employed fixed effects (dummy variables) to account for nonlinear hysteresis, as basin storage can have a nonlinear relationship with flood discharge. Results show that peak flood discharges could have been predicted 3 months in advance, and hence pro-active measures could have been employed to combat flood damage.

Of such measures, the potential of subsurface stores to accommodate excess rainfall water leading to a decrease in flood intensity is often cited (e.g., Revelle & Lakshminarayana 1975; Khan *et al.* 2014). However, it would be necessary to empty groundwater storage several months prior to a flood in order to provide sufficient time and storage volume for the excess water to recharge groundwater stores. The lead time of 3 months could provide sufficient leeway for land managers and farmers to use underground stores in this context – e.g., by extracting water from the aquifers and using it for irrigation and other water-intensive applications. Since this 3-month lead time covers the pre-monsoon period (when water demand is high due to dry weather), the extraction of groundwater can ease dry season water shortage issues. The site-specific constraints, such as costs associated with each technique and the technical feasibility of each technique (e.g., geophysical controls) should be analyzed before choosing a technique. As a result, with better flood forecasting methods and pro-active on the ground flood intensity reduction interventions, flood damage can be greatly averted in the Koshi basin, and possibly up-scaled over other parts of the Ganges basin due to similar hydrological conditions.

ACKNOWLEDGEMENTS

The author gratefully acknowledges the financial support provided by the Consortium of International Agricultural Research Centers (CGIAR) Research Program on Water, Land and Ecosystems (WLE). Sincere gratitude is extended to Dr John T. Reager from the University of California, Irvine, for providing helpful comments for the formulation of the paper. Gratitude is also extended to an anonymous reviewer and Dr Vladimir Smakhtin, Theme Leader at the International Water Management Institute, whose comments greatly impacted the paper. The author thanks Kriubaharan Jeremiah (Ashoka Trust for Research in Ecology and the Environment), Michael G. Sunde (University of Missouri), and Vaskar Dahal (University of Illinois at Urbana-Champaign) for helping with the R programming codes. The author acknowledges Dr Sean Swenson from the Gravity Recovery and Climate Experiment (GRACE) mission for providing the land water storage data. The author declares no conflict of interest.

REFERENCES

Andermann, C., Longuevergne, L., Bonnet, S., Crave, A., Davy, P. & Gloaguen, R. 2012 Impact of transient groundwater storage on the discharge of Himalayan rivers. *Nature Geoscience* **5**, 127–132.

Barros, A. P., Chiao, S., Lang, T. J., Burbank, D. & Putkonen, J. 2006 From weather to climate – seasonal and interannual variability of storms and implications for erosion processes in the Himalaya. *Geological Society of America Special Papers* **398**, 17–38.

Billah, M. M., Goodall, J. L., Narayan, U., Reager, J. T., Lakshmi, V. & Famiglietti, J. S. 2015 A methodology for evaluating evapotranspiration estimates at the watershed-scale using GRACE. *Journal of Hydrology* **523**, 574–586.

Bookhagen, B. & Burbank, D. W. 2010 Toward a complete Himalayan hydrological budget: spatiotemporal distribution of snowmelt and rainfall and their impact on river discharge. *Journal of Geophysical Research: Earth Surface* **115**, F3.

Boughton, W. & Droop, O. 2003 Continuous simulation for design flood estimation – a review. *Environmental Modelling and Software* **18**, 309–318.

Campolo, M., Andreussi, P. & Soldati, A. 1999 River flood forecasting with a neural network model. *Water Resources Research* **35** (4), 1191–1197.

Camporese, M., Penna, D., Borga, M. & Paniconi, C. 2014 A field and modeling study of nonlinear storage-discharge dynamics for an Alpine headwater catchment. *Water Resources Research* **50** (2), 806–822.

Chau, K. W., Wu, C. L. & Li, Y. S. 2005 Comparison of several flood forecasting models in Yangtze River. *Journal of Hydrologic Engineering* **10** (6), 485–491.

Cheng, M., Ries, J. C. & Tapley, B. D. 2011 Variations of the Earth's figure axis from satellite laser ranging and GRACE. *Journal of Geophysical Research: Solid Earth* **116**, B1.

Chinnasamy, P. & Agoramoorthy, G. 2015 Groundwater storage and depletion trends in Tamil Nadu State, India. *Water Resources Management* **29**, 2139.

Chinnasamy, P. & Agoramoorthy, G. 2016 India's groundwater storage trends influenced by tube well intensification. *Groundwater* **54** (5), 727–732.

Chinnasamy, P. & Sunde, M. G. 2015 Improving spatiotemporal groundwater estimates after natural disasters using remotely sensed data – a case study of the Indian Ocean Tsunami. *Earth Science Informatics* **9** (1), 101–111.

Chinnasamy, P., Hubbart, J. A. & Agoramoorthy, G. 2013 Using remote sensing data to improve groundwater supply estimations in Gujarat, India. *Earth Interactions* **17**, 1–17.

Chinnasamy, P., Bharati, L., Bhattarai, U., Khadka, A., Dahal, V. & Wahid, S. 2015a Impact of planned water resource development on current and future water demand in the Koshi River basin, Nepal. *Water International* **40**, 1004–1020.

Chinnasamy, P., Maheshwari, B. & Prathapar, S. 2015b Understanding groundwater storage changes and recharge in Rajasthan, India through remote sensing. *Water* **10**, 5547–5565.

Chinnasamy, P., Misra, G., Shah, T., Maheshwari, B. & Prathapar, S. 2015c Evaluating the effectiveness of water infrastructures for increasing groundwater recharge and agricultural production – a case study of Gujarat, India. *Agricultural Water Management* **158**, 179–188.

Dhar, O. N. & Nandargi, S. 2002 Flood study of the Himalayan tributaries of the Ganga river. *Meteorological Applications* **9**, 63–68.

Famiglietti, J. S. & Rodell, M. 2013 Water in the balance. *Science* **340**, 1300–1301.

Fushimi, H., Ikegami, K., Higuchi, K. & Shankar, K. 1985 Nepal case study: catastrophic floods. In: *Techniques for Prediction of Runoff from Glacierized Areas* (G. J. Young, ed.). IAHS Publication, Oxford, no. 149, pp. 125–130.

Garrote, L. & Bras, R. L. 1995 A distributed model for real-time flood forecasting using digital elevation models. *Journal of Hydrology* **167** (1), 279–306.

Geruo, A., Wahr, J. & Zhong, S. 2013 Computations of the viscoelastic response of a 3-D compressible Earth to surface loading: an application to Glacial Isostatic Adjustment in Antarctica and Canada. *Geophysical Journal International* **192**, 557–572.

Guo, Y., Liu, S. & Baetz, B. W. 2012 Probabilistic rainfall-runoff transformation considering both infiltration and saturation excess runoff generation processes. *Water Resources Research* **48**, W06513.

Hsu, K., Gupta, H. V. & Sorooshian, S. 1995 Artificial neural network modeling of the rainfall-runoff process. *Water Resources Research* **31** (10), 2517–2530.

Jain, S. K., Agarwal, P. K. & Singh, V. P. 2007 Ganga Basin. In: *Hydrology and Water Resources of India*. Springer, Dordrecht, The Netherlands, pp. 333–418.

Khan, M. R., Voss, C. I., Yu, W. & Michael, H. A. 2014 Water resources management in the Ganges Basin: a comparison of three strategies for conjunctive use of groundwater and surface water. *Water Resources Management* **28** (5), 1235–1250.

Kia, M. B., Pirasteh, S., Pradhan, B., Mahmud, A. R., Sulaiman, W. N. A. & Moradi, A. 2012 An artificial neural network model for flood simulation using GIS: Johor River Basin, Malaysia. *Environmental Earth Sciences* **67** (1), 251–264.

Landerer, F. & Swenson, S. 2012 Accuracy of scaled GRACE terrestrial water storage estimates. *Water Resources Research* **48**, W04531.

Liao, F., Porporato, A., Revelli, R. & Ridolfi, L. 2003 A comparison of nonlinear flood forecasting methods. *Water Resources Research* **39** (5), 1129.

Mool, P. K., Bajracharya, S. R., Joshi, S. P., Kunzang, K. & Gurung, D. R. 2001 *Inventory of Glaciers, Glacial Lakes and Glacial Lake Outburst Floods. Monitoring and Early Warning Systems in the Hindu Kush-Himalayan Region*. Nepal. International Centre for Integrated Mountain Development (ICIMOD), Patan, Nepal.

Moore, R. J., Cole, S. J., Bell, V. A. & Jones, D. A. 2006 Issues in flood forecasting: ungauged basins, extreme floods and uncertainty. *Frontiers in Flood Research* **305**, 103–122.

Mukherjee, D. & Mansour, N. 1996 Estimation of flood forecasting errors and flow-duration joint probabilities of exceedance. *Journal of Hydraulic Engineering* **122** (3), 130–140.

Porporato, A. & Ridofi, L. 2001 Multivariate nonlinear prediction of river flows. *Journal of Hydrology* **248**, 109–122.

Raman, H. & Sunilkumar, N. 1995 Multivariate modelling of water resources time series using artificial neural networks. *Hydrological Sciences Journal* **40** (2), 145–163.

Reager, J. T. & Famiglietti, J. S. 2009 Global terrestrial water storage capacity and flood potential using GRACE. *Geophysical Research Letters* **36**, L23402.

Reager, J. T., Thomas, B. F. & Famiglietti, J. S. 2014 River basin flood potential inferred using GRACE gravity observations at several months lead time. *Nature Geoscience* **7**, 588–592.

Reager, J. T., Thomas, A. C., Sproles, E. A., Rodell, M., Beaudoing, H. K., Li, B. & Famiglietti, J. S. 2015 Assimilation of GRACE terrestrial water storage observations into a land surface model for the assessment of regional flood potential. *Remote Sensing* **7**, 14663–14679.

Reddy, D. V., Kumar, D., Saha, D. & Mandal, M. K. 2008 The 18 August 2008 Kosi river breach: an evaluation. *Current Science* **95**, 1668.

Refsgaard, J. C. & Knudsen, J. 1996 Operational validation and intercomparison of different types of hydrological models. *Water Resources Research* **32** (7), 2189–2202.

Revelle, R. & Lakshminarayana, V. 1975 The Ganges water machine. *Science* **188** (4188), 611–616.

Rodell, M., Houser, P. R., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C.-J., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., Entin, J. K., Walker, J. P., Lohmann, D. & Toll, D. 2004 The global land data assimilation system. *Bulletin of the American Meteorological Society* **85**, 381–394.

Rodell, M., Chen, J., Kato, H., Famiglietti, J. S., Nigro, J. & Wilson, C. R. 2007 Estimating groundwater storage changes in the Mississippi River basin (USA) using GRACE. *Hydrogeology Journal* **15**, 159–166.

Rodell, M., Velicogna, I. & Famiglietti, J. S. 2009 Satellite-based estimates of groundwater depletion in India. *Nature* **460**, 999–1002.

Shrestha, A. B. & Bajracharya, S. R. (eds). 2013 *Case Studies on Flash Flood Risk Management in the Himalayas*. ICIMOD, Kathmandu, Nepal.

Sproles, E. A., Leibowitz, S. G., Reager, J. T., Wigington Jr, P. J., Famiglietti, J. S. & Patil, S. D. 2015 GRACE storage-runoff hystereses reveal the dynamics of regional watersheds. *Hydrol. Earth Syst. Sci.* **19**, 3253–3272.

Swenson, S. & Wahr, J. 2006 Post-processing removal of correlated errors in GRACE data. *Geophysical Research Letters* **33**, L08402.

Swenson, S., Chambers, D. & Wahr, J. 2008 Estimating geocenter variations from a combination of GRACE and ocean model output. *Journal of Geophysical Research: Solid Earth*, **113**, B08410.

WMO – World Meteorological Organization 1992 *Simulated Real-Time Intercomparison of Hydrological Models*. WMO Publications, Geneva, Switzerland.

First received 22 December 2015; accepted in revised form 19 September 2016. Available online 5 December 2016