Hydrological modeling as a tool for water resources management of the data-scarce Brahmaputra basin

Pulendra Dutta and Arup Kumar Sarma

ABSTRACT

A robust hydrological assessment is a challenging task in regions of limited hydro-climatological information. This level of uncertainty is further augmented in studies of flood hydrology for regions like the Brahmaputra River basin, where spatial variations of topography, land use, soil, and weather components are very high. The present study describes the development of a suitable hydrologic model for the data-scarce transboundary Brahmaputra River basin occupying an area of more than $5,42,000$ km^2. The main objective is to provide hydrologic assessment of the Brahmaputra River basin, even at locations having hardly any historical records. The Soil and Water Assessment Tool (SWAT) model is calibrated and validated using observed discharge of three sections located on the main stem. The results show a fair strength of the statistical parameters. Moreover, the model has been found to produce a satisfactory replica of historical flows at the tributaries with a fair value of correlation ($R^2 = 0.77$) at Golaghat. The results of this model would facilitate the ability of the local authorities with science-based elements to carry out decisions on the management of water resources at the main basin, and even at the sub-basin level.

Key words | Brahmaputra basin, data-scarce, hydrologic model, transboundary

Pulendra Dutta (corresponding author)
Arup Kumar Sarma
Department of Civil Engineering,
Indian Institute of Technology,
Guwahati 781039,
India
E-mail: *pulendutta@gmail.com*

INTRODUCTION

Expanding over four different countries, the Brahmaputra River basin occupies an area of $5,42,450$ km^2 with the outlet (23.576N: 89.442E) in Bangladesh. The Brahmaputra is one of the most highly sediment-laden rivers of the world (Goswami 1985). Also, this river basin is located in an area of high structural instability as a large number of earthquakes have been evidenced in the Himalayan catchment through which it flows (Lahiri & Sinha 2012). The Brahmaputra carries the largest amount of water and silt of all Indian rivers. However, this is probably the least exploited basin, even though it has enormous potential regarding water resources management and hydropower generation. Hydrologic models are among the available tools used to estimate the parameters required for water resource planning and management.

There have been many case-specific studies on definite parts of the Brahmaputra basin (Goswami 1985; Sarma 2005; Akhtar *et al.* 2011; Ghosh & Dutta 2012; Gogoi *et al.* 2012; Apurv *et al.* 2015; Sahoo & Sreeja 2015). These studies addressed the local issues only, rather than the whole basin. On the other hand, only a few studies have so far been carried out for this important river basin regarding bank line migration (Singh *et al.* 2004), ion chemistry (Hren *et al.* 2007), and applicability of Lacey's regime equations (Swamee *et al.* 2008). None of these studies addressed the hydrologic assessment of the basin. However, Rao *et al.* (2009) attempted a hydrologic assessment, but the entire Brahmaputra basin was not covered in their study. In another study, although the whole basin was considered by Aktar *et al.* (2015), they failed to establish whether the hydrologic

model developed would be applicable for deriving outputs across the basin. Here, the model was calibrated at only one location and no consideration was paid to spatial outputs. As such, we have attempted in this study to understand the hydrology of the entire Brahmaputra River basin so as to facilitate the water resource planners with science-based information at the desired locations.

Understanding the Brahmaputra basin has always been challenging due to its complex characteristics, both hydraulically and hydrologically (Rao et al. 2009). Although hydrologic models like SWAT, VIC (Variable Infiltration Capacity), WMS (Watershed Modelling System), HEC-HMS (Hydrologic Engineering Center–Hydrologic Modeling System) serve the purpose of hydrological analysis, they require correct input information. Even ANN (artificial neural network) can effectively be used to model rainfall–runoff processes (Hasanpour et al. 2014). The Brahmaputra basin is characterized by high spatial variations of topography, land use, soil properties, and weather components. As such, the challenge to model this basin lies in hydro-climatic information across and beyond the national boundary. This is because the data scarcity situations widely impact the hydrologic model results, as evidenced by many studies (Yu et al. 2011; Ploeg et al. 2012; Tabari et al. 2012; Valdivieso & Sendra 2014). Here, the spatial scarcity (Brath et al. 2004; Chen et al. 2018) as well as the temporal scarcity (Chen et al. 2018) of data are the major concerns for obtaining poorer simulation results. In the absence of available data, however, the missing variables may be generated (Nyeko 2015) and/or explicitly trained as constraints in mathematical programming (Pande et al. 2012). Unfortunately, these approaches often provide less satisfactory simulation results. In this study, we, therefore, have tried to incorporate the presently available data at many sources locally or globally.

The weather variables are the main components to widely impact the model results, but sufficient observed data are unavailable to model this river basin. Indeed, the observed weather records in large parts of the world are often scarce, discontinuous, and frequently contain discrepancies (Hughes 2006; Koutsouris et al. 2016). Even if longer records are available in the gauge observations, they do not afford a reliable spatial depiction (Yatagai et al. 2009) and, as such, evaluations, especially for rainfall characteristics, are important prior to applications in water resource

schemes (Nyatuame et al. 2014). In this study, evaluations of various datasets are carried out based on the simulation results of the SWAT model, set up independently for different input datasets. Subsequently, a proper dataset suitable for the data-scarce Brahmaputra basin is identified.

The process of calibration and validation of a hydrologic model is not possible at the ungauged sections where observed data are not available. Also, the calibration of the large Brahmaputra basin at a single site (Aktar et al. 2015) is not judicious due to its heterogeneous characteristics. However, the present study adopted a multi-site calibration technique for addressing better intra-watershed simulations.

The flow production in the present basin is estimated by implementing the SWAT model, which has previously been successfully applied by many researchers across the globe. Some such studies include flow and sediment simulation (Rosenberg et al. 1999; Srinivasan et al. 2005; Zhang et al. 2008; Valdivieso & Sendra 2014; Daggupati et al. 2015; Dahal et al. 2016), as well as quantitative and qualitative assessment of water quality (Fohrer et al. 2001; Grizzetti et al. 2003; Bekiaris et al. 2005; Abbaspour et al. 2007; Guse et al. 2015; Chen et al. 2018). Another reason for using SWAT is that it is computationally efficient and capable of continuous simulation over long periods of time (Borah & Bera 2004; Arnold et al. 2012). SWAT is a continuous, semi-distributed process-based river model, and describes interactions between the elementary units of various hydrological processes (Lempert & Ostrowski 2002). As the model developers need to make many critical decisions during development and calibration (Zhang et al. 2009; Arnold et al. 2015), the present study is done by utilizing various combinations of datasets in order to ensure accurate, real-world situations.

The Brahmaputra River basin has huge potential with regard to water resources projects, although severe hazards like flood and bank erosion that result in huge loss of life and property every year are also much pronounced. Developing a comprehensive management policy is the current priority, but requires adequate data regarding flow, sediment, soil moisture, water quality, etc. We, therefore, adopted hydrologic models as a tool in order to estimate these data for the Brahmaputra basin, although their establishment was a major challenge due to the lack of input data. Furthermore, we could not find any observational records for the majority of the tributaries and sub-tributaries

that join into the main stem of the Brahmaputra River, within Indian territory in particular. Therefore, modeling these tributary basins individually is really a challenging task, which can probably only be solved through modeling practice at a comprehensive level covering the entire transboundary Brahmaputra River basin, for which certain data are at least available. The present study presents a modeling-based approach to understand the dominant processes controlling the water balance to bridge the gap of base-line knowledge of water resources in the data-scarce Brahmaputra basin. The main aim of this modeling exercise is to identify a suitable dataset for carrying out hydrologic assessment across the basin, including locations having no historical records. This study would provide information required for management of water resources across the basin including its sub-basins.

STUDY AREA

The proposed study is intended for the transboundary Brahmaputra River basin (Figure 1). The Brahmaputra River originates in southern Tibet at an elevation of 5,300 m. Out of its total length of 2,880 km, the Brahmaputra covers a major part of its journey in Tibet as 'Tsangpo' River (also popularly known as 'Yarlung Zangbo' in China) flowing 1,625 km in Tibet parallel to the main range of the Himalayas before entering India through Arunachal Pradesh where it is popularly known as the 'Siang'

River. It flows for about 35 km and is joined by two other major tributaries, Dibang and Lohit, at a place in the west of Sadiya, Assam, India. From this confluence point, the river is known as the Brahmaputra until it crosses Assam, traveling 720 km and entering Bangladesh as Jamuna. After traveling a distance of around 95 km from this location, the Brahmaputra is joined by Surma, another major river of Bangladesh. From this point, the Brahmaputra is popularly pronounced as Meghna which then finally merges into the Bay of Bengal to end its journey. Although the main river flows through China, India, and Bangladesh, its basin boundary encompasses Bhutan as the fourth nation (Figure 2(a)).

MATERIALS AND METHODS

In this approach, the results of the SWAT model obtained by using various weather datasets are presented. Gassman et al. (2007) reported on SWAT as a physically based model that simulates the physical processes through input parameters like topography, land use, climate variables, and soil properties. In this study, the basin boundary is delineated through the use of Shuttle Radar Transmission Mission (SRTM) digital elevation model (DEM) of 90 m spatial resolution. In order to form a single DEM, mosaic in ArcGIS 10.4 was carried out for 20 DEMs so as to cover the entire basin. The stream network (Figure 2(b)) is generated for the threshold drainage area of 20,000 km².
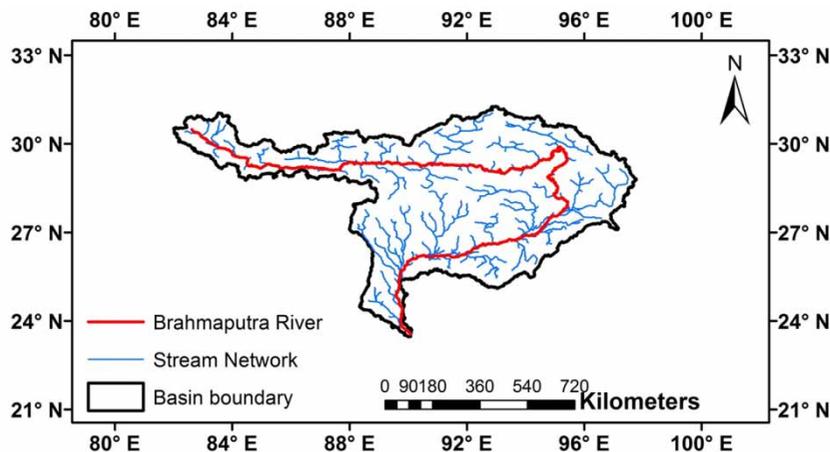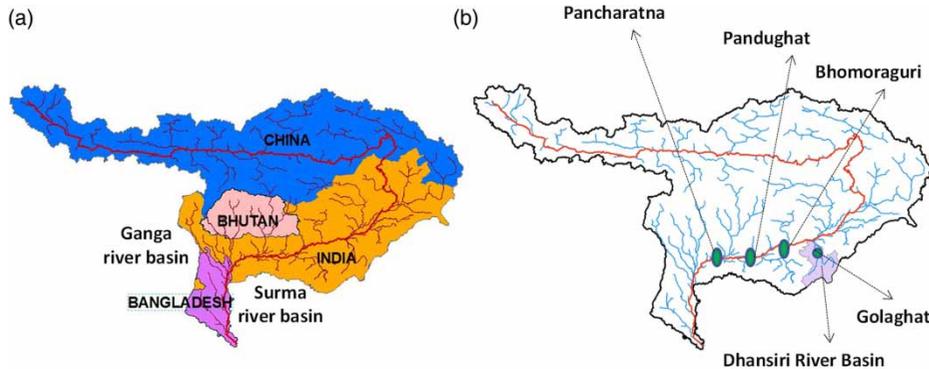

Figure 1 | Study area showing basin boundary of the Brahmaputra River and its stream network.
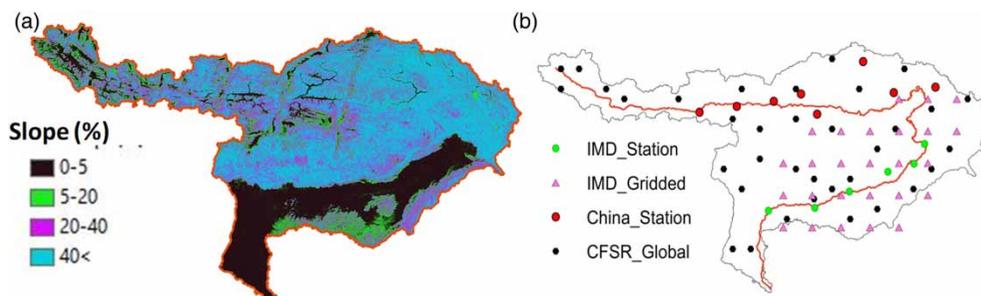
Figure 2 | Brahmaputra River basin: (a) co-nations sharing the basin area and (b) salient locations under consideration in this study (Golaghat, Bhomoraguri, Pandughat, and Pancharatna are the locations of available discharge).

In SWAT, the Brahmaputra basin watershed is divided into 41 sub-watersheds, which are then further subdivided into 1,578 hydrologic response units (HRUs). The HRUs represent an area consisting of dominant land use, soil characteristics, topography, and management practices. The present study utilizes 0.5 km resolution MODIS-based global data for the land use and to capture the land cover of the present basin. A soil map having 0.9 km spatial resolution, and provided by the Food and Agriculture Organization (FAO) has been used in the present study. Based on this soil map, we found as many as 39 soil types in the entire basin. The Brahmaputra basin is characterized by a very high variation in topography (Figure 3(a)).

Water balance is the driving force behind all the processes in SWAT because it impacts plant growth and the movement of sediments, nutrients, pesticides, and pathogens (Arnold et al. 2012). The hydrologic cycle depends upon weather variables such as rainfall, air temperature, solar radiation, wind speed, and relative humidity, which ultimately control the water balance of a watershed. The present study uses four sets of climate data obtained from various sources, as shown in Figure 3(b). The first set of weather variables is the Climate Forecast System Reanalysis (CFSR) data, provided by Texas A&M University (TAMU) as gridded global, high resolution coupled atmosphere–ocean–land surface–sea ice system weather data. These weather data consist of six variables: precipitation (pcp), maximum/minimum temperature (tmp), relative humidity (hmd), solar (slr), and wind (wnd), and are available for 36 years (1979–2014). Out of numerous grid points available within (21°–33°) North latitude and (82°–98°) East longitude, only 35 stations are selected so that at least one station falls on each of the major sub-basins, based on its spatial extent. The second weather datasets include observed data at six gauge stations in Indian territory, and collected from the Indian Meteorological Department (IMD). These datasets consist of four variables (pcp, max/min tmp, wnd), and were collected for a period of 25

Figure 3 | Brahmaputra River basin: (a) slope map and (b) weather stations considered in the present study. These stations are IMD (large circles within India), IMD gridded (triangles), China (large circles within China), and CFSR (small circles).

years (1991–2015). The third dataset includes the measured weather data at a few stations in the Tibet region of China (see Results and discussion section), and contains all the six variables spanning records of 22 years (1991–2012). Finally, the IMD gridded weather data generated by Dr. Balaji Narasimhan, Associate Professor, IIT Madras have been used as the fourth weather dataset. These data contain only three variables (pcp, tmp min/max) and are available up to the year 2005. Any variables if not available in the later three datasets are copied from the nearest corresponding CFSR dataset. The missing values are filled with negative 99, which means SWAT is directed to generate the missing values based on the available records.

The ArcSWAT ArcGIS extension is a graphical user interface for the SWAT, which is a semi-distributed, time-continuous watershed simulator operating on a daily time step (Arnold et al. 1998). Here, a total of six SWAT models have been developed by using the same DEM, land use map, soil map, and slope classifications in ArcSWAT2012. However, different weather datasets were utilized for each model. The first SWAT model is developed through the use of only the 35 global CFSR stations' data; whereas, the others are developed by using a combination of weather datasets, as indicated in Table 1.

The first four SWAT models were simulated for 22 years (1991–2012) on a monthly time step basis whereas the last two models (MODEL_5 and MODEL_6) were restricted to simulation for only 15 years (1991–2005), due to unavailability of the IMD gridded weather data beyond 2005. Therefore, the calibration and validation periods of these two SWAT models were chosen differently than the previous ones.

SWAT-CUP is an interface to SWAT, and any calibration/uncertainty or sensitivity program can easily be linked to the later models. It provides a decision-making framework that incorporates a semi-automated approach (SUFI2) using calibration and incorporating sensitivity and uncertainty analysis (Arnold et al. 2012).

The key parameters are necessary to be identified prior to calibration (Ma et al. 2000), and the present study identifies nine sensitive parameters (Table 2) during the iterations of each SWAT model. As the watershed models suffer from various uncertainties in regard to model, input, and parameters, the present study has adopted an uncertainty analysis for the parameters identified through the sensitivity analysis. Table 2 shows the range of sensitive parameters to mark the uncertainty level along with the best fit values at which calibration and validation results are obtained. Normally, three to five iterations are sufficient (Abbaspour et al. 2015) to conclude the best fit of the calibrated flows with respect to the measured ones. Out of four locations (Figure 2(b)) of available discharges, the outlets on the main stem, namely, (i) Bhomoraguri (26.606417N; 92.843865E), (ii) Pandughat (26.176872N; 91.687088E), and (iii) Pancharatna (26.201282N; 90.575692E) are used for model calibrations and validations. Golaghat (26.50278N; 93.95194E), on the Dhansiri tributary basin, is used for cross verification of the model results. The evaluations of model performance are carried out using certain statistical approaches: (a) the coefficient of determination (R^2), which indicates the strength of the relationship between the observed and simulated flow values; (b) Nash–Sutcliffe coefficient (NS), which shows how well the observed values fit the simulated values; (c) P-factor, which represents the

Table 1 | List of SWAT models developed in the present study and their data usage

S.N	Model #	Land use	Soil	Weather	CP	VP
1	MODEL_1	MODIS-based Global Land Cover map with 0.5 km spatial resolution	FAO soil map of 0.9 km spatial resolution	CFSR	1999–2008	2009–2012
2	MODEL_2			CFSR + IMD Stn	1999–2008	2009–2012
3	MODEL_3			CFSR + China Stn	1999–2008	2009–2012
4	MODEL_4			CFSR + (IMD + China) Stn	1999–2008	2009–2012
5	MODEL_5			CFSR + IMD grid	1993–2001	2002–2005
6	MODEL_6			CFSR + IMD grid + China Stn	1993–2001	2002–2005

The calibration periods (CP) and validation periods (VP) are also shown in this table.

Table 2 │ List of sensitive parameters for the present models

SI	Parameter	Description	Range		Best fit
			Min	Max	
1	R__SOL_AWC	Average available soil water content	0.125	0.348	0.196
2	R__ALPHA_BF	Base flow recession factor in days	0.054	0.503	0.246
3	V__GW_DELAY	Ground water delay time in days	0	128.800	79.162
4	V__GW_REVAP	Ground water revap coefficient	0.110	0.184	0.135
5	R__CN2	SCSII curve number	−0.545	−0.173	−0.354
6	V__SMTMP	Snowmelt base temperature in °C	0.970	4.920	1.276
7	V__ESCO	Soil evaporation compensation factor	0.828	0.897	0.888
8	V__GWQMN	Threshold water depth in shallow aquifer required for return flow to occur	1.457	2.753	1.875
9	A__REVAPMN	Threshold water depth in shallow aquifer required for revap to occur	0	127.083	111.516

Note: R: relative; V: replace; A: absolute.

percentage of measured data bracketed by the 95% prediction uncertainty (95PPU); and (d) R-factor, which is the measure of uncertainty.

Initially, all the models were simulated, calibrated, and validated at a single-outlet basis, for discharges (collected from Central Water Commission, (CWC)) at the three main stem outlets individually. Based on these results, MODEL_6 was identified as the suitable model for the present condition and thus forwarded for the spatial calibration/validation. This is because the performance of the default model should not be too drastically different from the measurement. If so, often calibration can be of little help (Abbaspour *et al.* 2015). Finally, the model calibrated and validated at the main stem outlets is cross-verified for outputs at the tributary basin to check whether the same would produce satisfactory results at the sub-basin level or not.

RESULTS AND DISCUSSION

The Brahmaputra is one of the largest river systems in the world and, as such, researchers have faced a huge challenge to understand the physical processes (Pahuja & Goswami 2006). It is, therefore, the need of the moment to develop a comprehensive water management cooperation mechanism for water sharing between the co-basin nations – India, China, Bhutan, and Bangladesh (Sharma 2014). In one approach, Scholars from Bangladesh, China & India

(2014) articulated some issues about water resources of the Himalayan region and were of the opinion that these could be brought before the rest of the world to forward constructive discussions and information exchange by international experts, media, and civil society. Thus, sharing data between the riparian countries can only be achieved through an inter-governmental understanding (Thu & When 2016); but, there is no such agreement known to us between the co-nations. Therefore, obtaining flow data of the river from the upper riparian country was not possible. However, for modeling purposes, we have used some precipitation data along with a few other weather components of China (Tibet), obtained through academic collaboration. This section of the paper highlights the results of the present modeling study along with the subsequent analyses.

Simulation results

It is very important to use adequately correct data while establishing a hydrologic model, without which reasonably valid outputs can rarely be achieved. Figure 4 shows a scatter plot between the simulated and observed flows for the models at Pancharatna (Figure 2(b)), as an example. Similar trends were also observed at the other two outlets, namely, Bhomoraguri and Pandughat. This plot provides an idea of how the simulated flow is different from the observed flow. A perfect match between the two represents a straight 1:1 line.
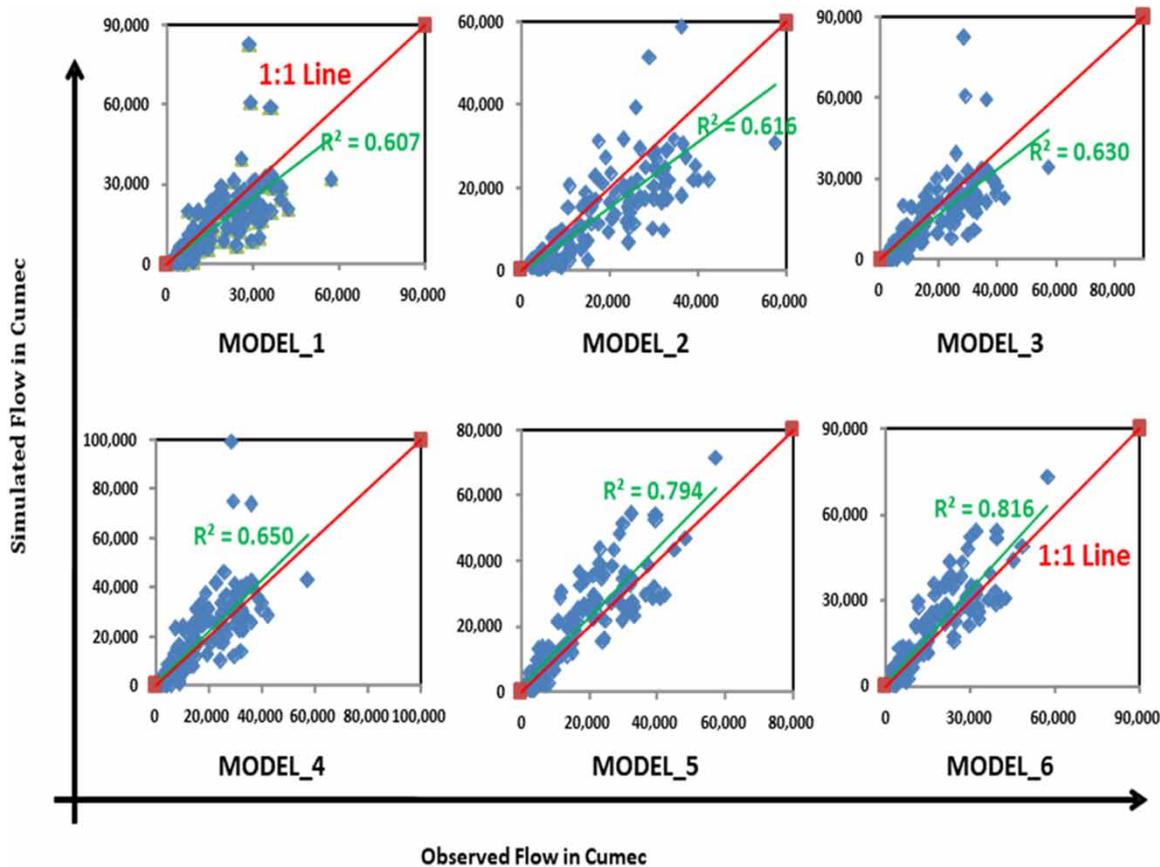
Figure 4 │ Scatter plot of monthly simulated flows (m³/s) against the corresponding measured discharges (m³/s) at Pancharatna for all SWAT models. Here, simulated values indicate the monthly outputs obtained from model run (1991–2012, for the first four models and 1991–2005, for the last two models) without calibration and validation.

It is evident from Figure 4 that the MODEL_1 which uses only the CFSR global stations' datasets exhibits the least correlation ($R^2 = 0.607$) value, indicating the highest disagreement between the measured and simulated flows out. The flows out of MODEL_2 and MODEL_3 that use observed stations' data of IMD and China, respectively, also did not show a good correlation with the observed flow data. Here, although there are improvements in model outputs, significantly wide changes have not been observed, due to non-capture of all the observed weather datasets by the second and third SWAT models. This is because only the weather station close to the geometric center of sub-watersheds is picked up by SWAT. On combining all the observed weather datasets with the global CFSR dataset as used in MODEL_4, however, the correlation further improved with R^2 value being 0.65.

The simulation result of MODEL_5 has surprisingly improved as the IMD gridded data are added to the

global CFSR weather data. Furthermore, MODEL_6 provided the best correlation ($R^2 = 0.816$) value among all models developed in this study. This analysis indicates MODEL_6, which uses the global CFSR dataset in conjunction with the IMD gridded, as well as the observed data over China, would provide the best performances in simulating the discharges, and it, therefore, may be taken up as the default model for the present basin.

Calibration and validation results

Single-outlet calibration/validation

The calibration and validation of the SWAT models are done in the SWAT-CUP platform that tries to capture most of the measured data within 95% prediction uncertainty (95PPU) during iteration.

The two indices P-factor and R-factor are used for the goodness of fit between the simulated and the observed values. The P-factor is the fraction of observed data bracketed by the 95PPU band and varies from 0 to 1, where 1 indicates 100% bracketing by 95PPU. On the other hand, R-factor is the thickness of the model envelope (i.e., 95PPU) and should be around 1, but a value of <1.5 is acceptable (Abbaspour *et al.* 2007) for this index.

Initially, all the SWAT models are calibrated and validated on a single-outlet basis. In the present analysis, we show the calibrated and validated results at Pancharatna outlet only, as a continuous graph (Figure 5) for simplicity and clarity of presentation. Here, we show the time series plot for 'non-dimensional flow' rather than the actual values due to the data secrecy policy of the concerned authority. The first four models, i.e., MODELs_1, 2, 3, and 4, were calibrated for ten years (1999–2008) and validated for four years (2009–2012) by skipping eight years (1991–1998) for initiation of the hydrologic response. Meanwhile, the last two models, i.e., MODELs_5 and 6, were calibrated for nine years (1993–2001) and validated for four years (2002–2005). Although the more the skipping periods the better the initial response, we could not make it longer for the last two models due to limitations of model input data, and a two-year (1991–1992) period only was provided for them.

It is understood from the results that MODEL_1 provides the least satisfactory output, as evidenced by the calibration statistics in Table 3. Here, the value of R^2 (0.78) is fair, but the other values like NS (0.68), P-factor (0.60), and R-factor (0.63) are relatively inferior for acceptance of the model. On checking the validation statistics, however, slightly better performances were noticed. The statistics of MODEL_2 and MODEL_3 are found to comply satisfactorily. However, MODEL_4 provided the best statistics among the first four models, during both the calibration and validation. Surprisingly, MODEL_5 which showed good agreement during the simulation (Figure 4) failed to produce good statistics during calibration and validation, although they were carried out by using the same sensitive parameters (Table 2) as the previous ones. Among all models, MODEL_6 is found to perform the best, since it provides a very good strength of statistics for all the parameters, even during both the calibration and validation (Table 3). This fair result is due to using a proper combination of input datasets in this model (MODEL_6), as compared to those used in the other models (MODELs_1–5). The R-factor (0.97) value during calibration (Table 3) is nearly equal to the desired value (1.00). It indicates a very good measure of parameter uncertainty. As such, this model has only been extended for subsequent analyses.

Multi-outlet calibration/validation

The calibration at a single site is judicious for the small watershed only, because of homogeneous watershed characteristics. On the other hand, a large watershed needs a multi-site calibration in order to represent heterogeneous characteristics. The single-site calibration if done on a large watershed may result in a combination of under/overestimation for values (Qi & Grunwald 2005), due to consideration of homogeneous characteristics over the entire basin. To have a better intra-watershed spatial accuracy, the present study includes multi-site calibration/validation at the three main-stem outlets simultaneously. This spatial accuracy is done only for the best SWAT model as identified earlier (sections Simulation results and Single-outlet calibration/validation), i.e., MODEL_6, assuming that the other models would not lead to better performances, even during the multi-site calibration and validation.

The time series plots are shown in Figure 6 as a continuous graph for simplicity in the presentation. Due to the non-symmetry of available discharge data length, different calibration periods were chosen for the outlets (Figure 2(b)). However, the same validation periods for all the three sites were chosen (2002–2005).

Due to its complex characteristics, a large watershed may not be expected to produce good statistics against each outlet during multi-site calibration; however, a balance among the statistical parameters is required for reasonable acceptance of a model. During this spatial analysis (Figure 6(c)), the most downstream outlet (i.e., Pancharatna) is found to conform with good statistics during calibration ($R^2 = 0.84$; NS = 0.74) and validation (P-factor = 0.94; R-factor = 0.91). Similarly, the model is found to produce satisfactory statistics at the other two outlets, namely, Bhomoraguri (Figure 6(a)) and Pandughat (Figure 6(b)). However, the calibration results at Pancharatna, especially the value of R^2 (0.84) are relatively low
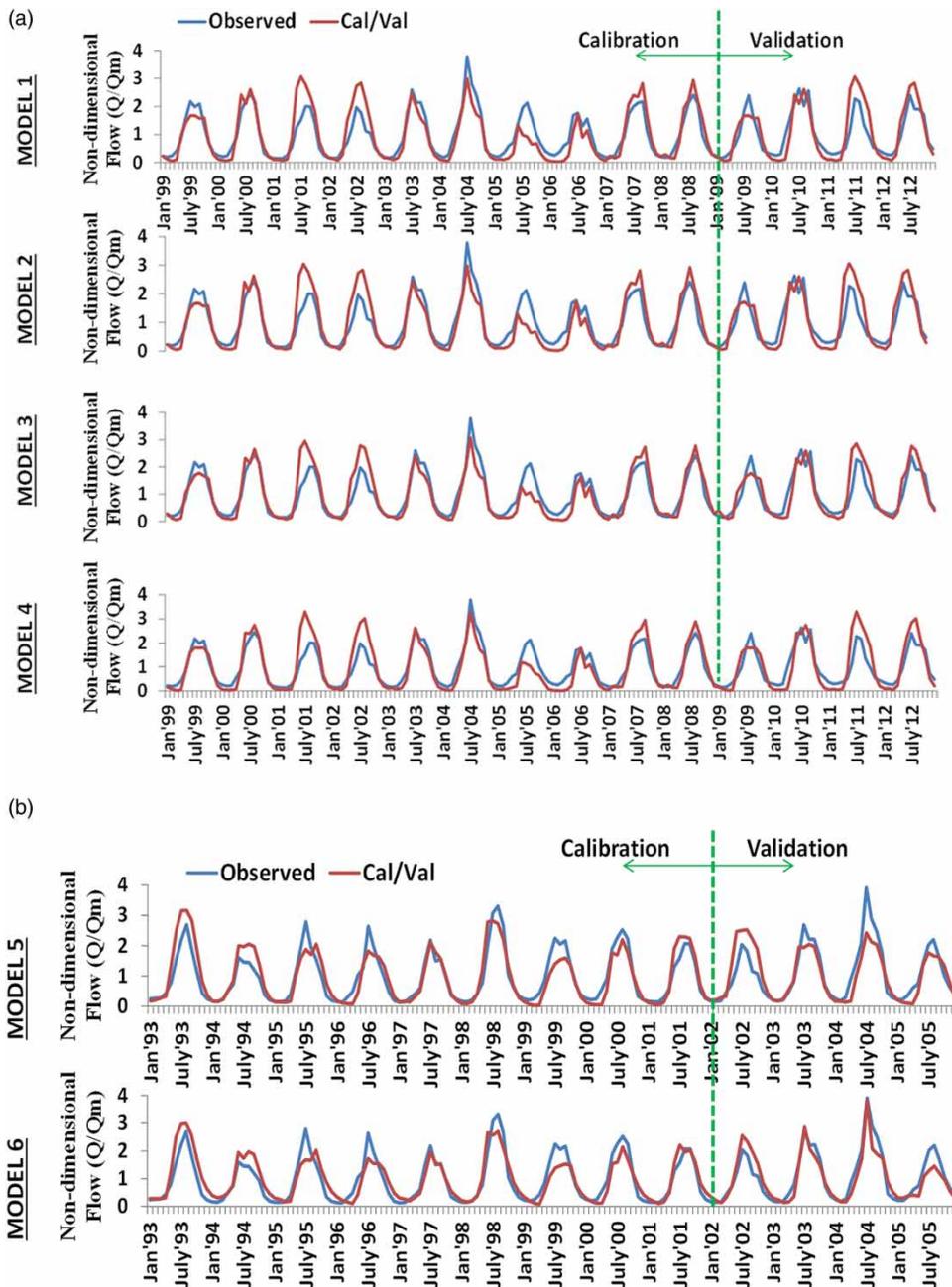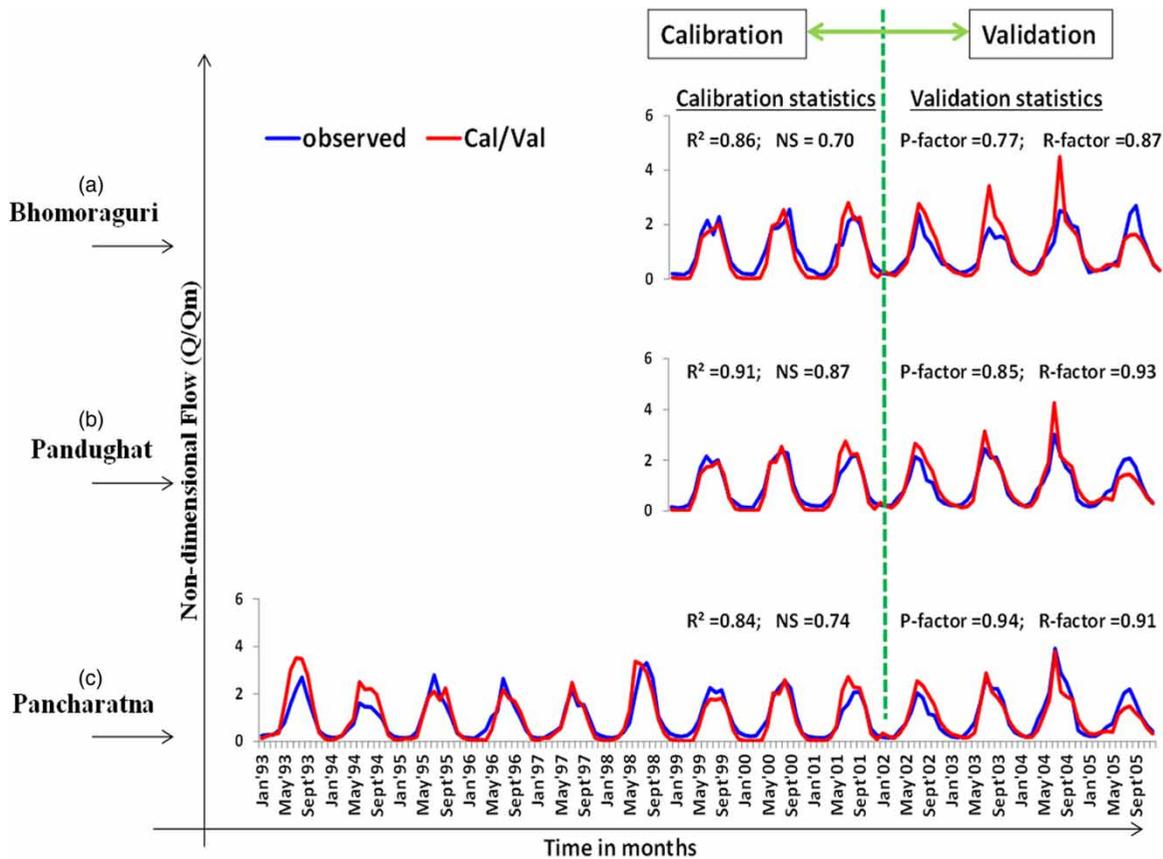
Figure 5 │ (a) Results of single-outlet calibration (1999–2008) and validation (2009–2012) at Pancharatna of the first four SWAT models (i.e., MODEL_1, MODEL_2, MODEL_3, MODEL_4). The calibration and validation results of the monthly models are plotted in a single plot showing the non-dimensional flow (Q/Q_m) values. Here, Q represents flow (simulated or observed) and Q_m represents the mean of five-year flow data selected on a random basis. (b) Results of single-outlet calibration (1993–2001) and validation (2002–2005) at Pancharatna of the last two SWAT models (i.e., MODEL_5 and MODEL_6). The calibration and validation results of the monthly models are plotted in a single plot showing the non-dimensional flow (Q/Q_m) values. Here, Q represents flow (simulated or observed) and Q_m represents the mean of five-year flow data selected on a random basis.

as compared to Bhomoraguri (0.86) and Pandughat (0.91). This is probably due to different lengths of warm-up period maintained for initiation of watershed responses by the hydrologic models. The warm-up period for Pancharatna

is only two years (1991–1992) against eight years (1991–1998) for the later two outlets. Overall, the spatial calibration/validation results of MODEL_6 may be termed as satisfactory.

Table 3 │ Statistics of single-output calibration/validation at Pancharatna for all SWAT models

Model #	Calibration statistics				Validation statistics			
	R^2	NS	P-factor	R-factor	R^2	NS	P-factor	R-factor
MODEL_1	0.78	0.68	0.60	0.63	0.77	0.70	0.71	0.74
MODEL_2	0.78	0.69	0.60	0.65	0.77	0.70	0.70	0.73
MODEL_3	0.80	0.73	0.59	0.60	0.78	0.74	0.67	0.70
MODEL_4	0.81	0.72	0.74	1.07	0.79	0.64	0.80	1.27
MODEL_5	0.82	0.70	0.60	0.62	0.73	0.66	0.60	0.46
MODEL_6	0.79	0.74	0.90	0.97	0.86	0.84	0.81	0.88

Figure 6 │ Results of multi-outlet calibration and validation results of MODEL_6. The calibration and validation results of the monthly model are plotted in a single plot showing the non-dimensional flow (Q/Q_m) values. Here, Q represents flow (simulated or observed) and Q_m represents the mean of five-year flow data selected on a random basis.

Verification of model results at tributary

The present SWAT models were developed for the large Brahmaputra River basin, but whether this model built for a basin area of more than $5,42,000 \text{ km}^2$ would be suitable for its sub-basins occupying smaller areas, or not, needs to be evaluated. This is done through cross verification of the model results. This analysis is possible for several outputs (discharge, sediment, evapotranspiration, groundwater substances, etc.) of the SWAT model, provided historical

records are available. Unfortunately, the majority of the tributaries have hardly any observational records, and we could collect only one set of discharge data for a tributary. Here, we compared the final MODEL_6 results pertaining to Golaghat, an outlet at Dhansiri tributary (Figure 2(b)), with the measured discharges. It is evident from the plot (Figure 7) that the model simulated values follow a fair correlation ($R^2 = 0.77$) with the measured discharges, in terms of magnitudes and pattern. As such, the present model is expected to provide satisfactory outputs at the other tributaries too. Unfortunately, the model verification was not possible at any other tributary since no historical records are available to us. It is evident from the analyses that the SWAT model could provide satisfactory outputs at both the main stem outlets as well as tributary outlets of the data-scarce Brahmaputra basin.

Model application in water resources management

The SWAT model provides location-specific hydrologic information across a river basin which may be utilized by the water resource managers for planning and designing hydraulic structures. This information belongs to the quantitative and qualitative assessment of river flows. A list of certain parameters is enumerated in Table 4 corresponding to a location (latitude: 29.15485N; longitude: 95.006979E) termed the 'Indo-China Border' (Figure 2(b)).

Here, monthly values are obtained from the outputs of the final run of MODEL_6 which are then averaged during 1991–2005. The SWAT hydrologic model output for river flows at this location varies widely from a very low value ($28.0 \text{ m}^3/\text{s}$) to a very high value ($8,364.8 \text{ m}^3/\text{s}$). Similarly, a broad range of outputs has also been noticed for the sediment concentration with a maximum value of 445.815 mg/L during the month of June. High values of sediment concentration may lead to loss of aquatic habitats, wetlands, and recreation attributes. Moreover, it becomes a concern for human health and erosion hazards. Water quality parameters like nitrogen, phosphorus, nitrates, ammonium, dissolved oxygen, etc. may be a concern for human and plant health. Based on the qualitative analysis for those parameters, the processes and cost of treatment of water, if derived for water supply projects, can be ascertained. Although the present model is capable of providing quite a lot of information regarding water quantity and quality at all desired locations across the basin, the present analysis is shown corresponding to one outlet only, as an example. Such information provided by the hydrologic models at a desired location may be utilized by water resource managers. As such, hydrologic models serve the purpose of deriving information at locations with no historical records, especially at those locations beyond the national boundary for such a transboundary river basin.
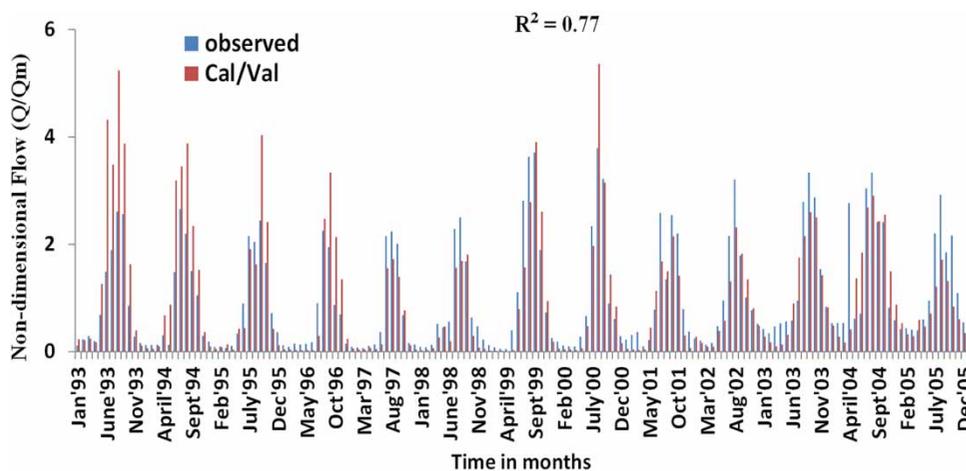
Figure 7 | Performance of MODEL_6 at tributary basin. This plot shows the comparison of non-dimensional flows (Q/Q$_m$) between the outputs of model final run and the observed values at Golaghat outlet in Dhansiri river basin. Here, Q represents flow (simulated or observed) and Q$_m$ represents the mean of five-year flow data selected on a random basis.

Table 4 | Parameters for water resources management practices and their annual average values at a certain location (latitude: 29.15485N; longitude: 95.006979E) within the Brahmaputra River basin

Months	Flow value m³/s	Sediment concentration mg/L	Organic nitrogen mg/L	Organic phosphorus mg/L	Nitrate mg/L	Ammonium mg/L	Mineral phosphorus mg/L	Dissolved oxygen mg/L
Jan	100.4	44.075	0.000	0.000	0.001	0.002	0.000	4.758
Feb	28.0	24.060	0.565	0.074	0.001	0.002	0.001	3.646
Mar	44.1	11.129	5.870	0.778	0.003	0.010	0.009	0.937
Apr	711.2	45.566	7.417	1.598	0.059	0.535	0.046	0.271
May	6,878.2	252.269	7.036	1.529	0.014	0.698	0.073	0.215
Jun	8,364.8	445.815	4.586	0.754	0.018	0.408	0.069	1.599
Jul	4,920.0	382.623	8.508	1.097	0.046	0.069	0.032	6.927
Aug	4,994.2	310.185	8.869	1.189	0.053	0.107	0.029	3.836
Sep	4,266.5	278.100	6.373	0.850	0.049	0.044	0.015	5.388
Oct	2,439.8	172.431	0.479	0.064	0.019	0.001	0.003	5.280
Nov	909.9	64.991	0.000	0.000	0.001	0.000	0.000	1.037
Dec	312.4	55.787	0.000	0.000	0.001	0.001	0.000	2.217

The figures are obtained as output from the final run of the SWAT model suitably adopted in the present study, i.e., MODEL_6.

CONCLUSION

There has been a huge challenge for researchers to model the transboundary Brahmaputra River basin due to its complex characteristics and data-scarce situations. However, we can succeed in modeling this important basin, as evidenced by the results discussed above. Here, the situations arising from the limited information and reliability of the available data are handled by developing various SWAT models, in order to decide the proper dataset suitable for the Brahmaputra basin. The conclusions from this study can be drawn as follows. (1) The CFSR weather data provides an underestimation of flow values in the Brahmaputra basin. (2) The observed weather stations data (IMD gridded + China weather) when used in conjunction with the CFSR data provided near-real hydrologic assessment for the basin. This dataset that produces acceptable model-generated stream flow series in the case of both single and multi-outlet validation is considered as the best combination. Therefore, this dataset along with the model parameters can now be used with confidence, for estimating outflow at any desired location of this large data-scarce basin. (3) The SWAT hydrologic model developed for the large Brahmaputra River basin can provide satisfactory monthly outputs for discharges, even at its tributaries. Calibration and validation may be possible for many variables like sediment and water quality parameters, and corresponding to more locations distributed throughout the basin, and even for the daily time series. However, that was not possible in this study because of the non-availability of records to us. This may be extended as scope for future study. (4) For regions lacking historical records, as in the present case, hydrologic models can serve the purpose of providing information regarding flow, sediment, and water quality across the basin. This would enable the water resource planners and managers to decide their plans and actions for extracting the utmost benefits from the basin. (5) The suitable hydro-meteorological dataset identified in this study can be used for GCM downscaling and hence for climate change projection of the Brahmaputra basin.

ACKNOWLEDGEMENTS

REFERENCES

Abbaspour, K. C., Rouholahnejad, E., Vaghefi, S., Srinivasan, R., Yang, H. & Klove, B. 2015 A continental-scale hydrology and water quality model for Europe: calibration and uncertainty of a high-resolution large-scale SWAT model. *Journal of Hydrology* **524**, 733–752.

Abbaspour, K. C., Yang, J., Maximov, I., Siber, R., Bogner, K., Mieleitner, J., Zobrist, J. & Srinivasan, R. 2007 Modelling hydrology and water quality in the pre-alpine/alpine Thur watershed using SWAT. *Journal of Hydrology* **333**, 413–430.

Akhtar, M. P., Sharma, N. & Ojha, C. S. P. 2011 Braiding process and bank erosion in the Brahmaputra River. *International Journal of Sediment Research* **26**, 431–444.

Aktar, N., Hossain, B. M. T. I., Ahmed, T., Khan, M. F. A., Islam, A. K. S. M., Yazdan, M. S., Noor, F. & Rahaman, A. Z. 2015 Climate change impacts on water availability in the Brahmaputra basin. In *5th International Conference on Water and Flood Management*, Dhaka, Bangladesh. Available from: www.researchgate.net/publication/281610569.

Apurv, T., Mehrotra, R., Sharma, A., Goyal, M. K. & Dutta, S. 2015 Impact of climate change on floods in the Brahmaputra basin using CMIP5 decadal predictions. *Journal of Hydrology* **527**, 281–291.

Arnold, J. G., Moriasi, D. N., Gassman, P. W., Abbaspour, K. C., White, M. J., Srinivasan, R., Santhi, C., Harmel, R. D., Griensven, A. V., VanLiew, M. W., Kannan, N. & Jha, M. K. 2012 SWAT: model use, calibration, and validation. *Transactions of the ASABE* **55**, 1491–1508.

Arnold, J. G., Srinivasan, R., Muttiah, R. S. & Williams, J. R. 1998 Large-area hydrologic modeling and assessment: part I. Model development. *Journal of the American Water Resources Association* **34** (1), 73–89.

Arnold, J. G., Youssef, M. A., Yen, H., White, M. J., Sheshukov, A. Y., Sadeghi, A. M., Moriasi, D. N., Steiner, J. L., Amatya, D. M., Skaggs, R. W., Haney, E. B., Jeong, J., Arabi, M. & Gowda, P. H. 2015 Hydrological processes and model representation: impact of soft data on calibration. *Transactions of the ASABE* **58**, 1637–1660. doi:10.13031/trans.58.10726.

Bekiaris, I. G., Panagopoulos, I. N. & Mimikou, M. A. 2005 Application of the SWAT (Soil and water assessment tool) model in the Ronnea catchment of Sweden. *Global NEST Journal* **7** (3), 252–257.

Borah, D. K. & Bera, M. 2004 Watershed-scale hydrologic and nonpoint-source pollution models: review of applications. *Transactions of the ASAE* **47** (3), 789–803.

Brath, A., Montanari, A. & Toth, E. 2004 Analysis of the effects of different scenarios of historical data availability on the calibration of a spatially-distributed hydrological model. *Journal of Hydrology* **291**, 232–253.

Chen, L., Xu, J., Wang, G., Liu, H., Zhai, L., Li, S., Sun, C. & Shen, Z. 2018 Influence of rainfall data scarcity on non-point source

pollution prediction: implications for physically based models. *Journal of Hydrology* **562**, 1–16.

Daggupati, P., Yen, H., White, M. J., Srinivasan, R., Arnold, J. G., Keitzer, C. S. & Sowa, S. P. 2015 Impact of model development, calibration, and validation decisions on hydrological simulations in West lake Erie basin. *Hydrological Processes* **29**, 5307–5320. doi:10.1002/hyp.10536.

Dahal, V., Shakya, N. M. & Bhattarai, R. 2016 Estimating the impact of climate change on water availability in Bagmati Basin, Nepal. *Environmental Processes* **3**, 1–17. https://doi.org/10.1007/s40710-016-0127-5.

Fohrer, N., Haverkamp, S., Eckhardt, K. & Frede, H. G. 2001 Hydrologic response to land use changes on the catchment scale. *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere* **26** (7–8), 577–582.

Gassman, P., Reyes, M., Green, C. & Arnold, J. G. 2007 The soil and water assessment tool: historical development, applications, and future research directions. *Transactions of ASABE* **50** (4), 1211–1250.

Ghosh, S. & Dutta, S. 2012 Impact of climate change on flood characteristics in Brahmaputra basin using a macro-scale distributed hydrological model. *Journal of Earth System Sciences* **121** (3), 637–657.

Gogoi, M. P., Gogoi, B., Hazarika, S. & Borgohain, P. 2012 GIS based study on fluvio-morphology of the river Brahmaputra in part of upper Assam, NE India. *Journal of Frontline Research* **2**, 114–121.

Goswami, D. C. 1985 Brahmaputra River, Assam, India: physiography, basin denudation and channel aggradation. *Water Resources Research* **21** (7), 959–978.

Grizzetti, B., Bouraoui, F., Granlund, K., Rekolainen, S. & Bidoglio, G. 2003 Modelling diffuse emission and retention of nutrients in the Vantaanjoki watershed (Finland) using the SWAT model. *Ecological Modelling* **169**, 25–38. doi:10.1016/S0304-3800(03)00198-4.

Guse, B., Pfannerstill, M. & Fohrer, N. 2015 Dynamic modeling of land use change impacts on nitrate loads in rivers. *Environmental Processes* **2**, 575. https://doi.org/10.1007/s40710-015-0099-x.

Hasanpour Kashani, M., Ghorbani, M. A., Dinpashoh, Y. & Shahmorad, S. 2014 Comparison of Volterra model and artificial neural networks for rainfall–runoff simulation. *Natural Resources Research* **23**, 341–354. doi:10.1007/s11053-014-9235-y.

Hren, M. T., Chemberlain, C. P., Hilley, G. E., Blisniuk, P. M. & Bookhagen, B. 2007 Major ion chemistry of the Yarlung Tsangpo–Brahmaputra river: chemical weathering, erosion, and CO2 consumption in the southern Tibetan plateau and eastern syntaxis of the Himalaya. *Geochimica et Cosmochimica Acta* **71**, 2907–2935. www.sciencedirect.com

Hughes, D. A. 2006 Comparison of satellite rainfall data with observations from gauging station networks. *Journal of Hydrology* **327**, 399–410.

Koutsouris, A. J., Chen, D. & Lyon, S. W. 2016 Comparing global precipitation data sets in eastern Africa: a case study of

Kilombero Valley, Tanzania. *International Journal of Climatology* **36**, 2000–2014.

Lahiri, S. K. & Sinha, R. 2012 Tectonic controls on the morphodynamics of the Brahmaputra River system in the Upper Assam Valley, India. *Geomorphology* **169–170**, 74–85. doi:10.1016/j.geomorph.2012.04.012.

Lempert, M. & Ostrowski, M. 2002 A hydrologic model to bridge the gap between conceptual and physically based approaches. In: *EGS XXVII General Assembly*, Nice, France. Available from: https://www.researchgate.net/publication/250929031

Ma, L., Ascough, J. C., Ahuja, L. R., Shaffer, M. J., Hanson, J. D. & Rojas, K. W. 2000 Root zone water quality model sensitivity analysis using Monte Carlo simulation. *Transactions of the ASAE* **43** (4), 883–895.

Nyatuame, M., Owusu-Gyimah, V. & Ampiaw, F. 2014 Statistical analysis of rainfall trend for Volta region in Ghana. *International Journal of Atmospheric Sciences* **2014**, 203245. http://dx.doi.org/10.1155/2014/203245.

Nyeko, M. 2015 Hydrologic modelling of data scarce basin with SWAT model: capabilities and limitations. *Journal of Water Resources Management* **29**, 81–94. doi:10.1007/s11269-014-0828-3.

Pahuja, S. & Goswami, D. 2006 *A Fluvial Geomorphology Perspective on the Knowledge Base Brahmaputra. Draft Paper No. 3, Study of Natural Resource, Water and the Environmental Nexus for Development and Growth in Northeast*. World Bank and Guwahati University, India. Available from: www.researchgate.net/publication

Pande, S., Savenije, H., Bastidas, L. A. & Gosain, A. K. 2012 A parsimonious hydrological model for a data scarce dryland region. *Journal of Water Resourced Management* **26**, 909–926. doi:10.1007/s11269-011-9816-z.

Ploeg, M. J. V. D., Haldorsen, S., Leijnse, A. & Heim, M. 2012 Sub permafrost groundwater systems: dealing with virtual reality while having virtually no data. *Journal of Hydrology* **475**, 42–52. doi:10.1016/j.jhydrol.2012.08. 046.

Qi, C. & Grunwald, S. 2005 GIS-based hydrologic modeling in the Sandusky watershed using SWAT. *American Society of Agricultural Engineers* **48** (1), 169–180.

Rao, K. H. V. D., Bhanumurthy, V. & Roy, P. S. 2009 Application of satellite- based rainfall products and SRTM DEM in hydrological modelling of Brahmaputra Basin. *Journal of the Indian Society of Remote Sensing* **37**, 587–600.

Rosenberg, N. J., Epstein, D. J., Wang, D., Vail, L., Srinivasan, R. & Arnold, J. G. 1999 Possible impacts of global warming on the hydrology of the Ogallala aquifer region. *Journal of Climate Change* **42** (4), 677–692.

Sahoo, S. N. & Sreeja, P. 2015 Development of flood inundation maps and quantification of flood risk in an urban catchment of Brahmaputra River. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering* **3**, A4015001.

Sarma, J. N. 2005 Fluvial process and morphology of the Brahmaputra River in Assam, India. *Journal of Geomorphology* **70**, 226–256. doi:10.1016/j.geomorph.2005. 02.007.

Scholars from Bangladesh, China and India 2014 *Brahmaputra: Towards Unity*. Available from: www.Thethirdpole.net

Sharma, N. 2014 *Imperatives for a Framework on Water Sharing Mechanism for Trans-Boundary Brahmaputra River*. Available from: https://www.linkedin.com/pulse

Singh, C. B., ISH, M. & Bhagwan, A. 2004 ANN model development for bank-line migration of river Brahmaputra using remote sensing data. *ISH Journal of Hydraulic Engineering* **10** (2), 133–138. doi:10.1080/09715010.2004. 10514761.

Srinivasan, M. S., Gerald-Marchant, P., Veith, T. L., Gburek, W. J. & Steenhuis, T. S. 2005 Watershed-scale modeling of critical source areas of runoff generation and phosphorus transport. *Journal of the American Water Resources Association* **41** (2), 361–375.

Swamee, P. K., Sharma, N. & Dwivedi, A. 2008 Lacey regime equations for river Brahmaputra. *Journal of Hydraulic Research* **46** (5), 707–710. doi:10.3826/jhr.2008.3130.

Tabari, H., Kisi, O., Ezani, A. & Talaee, P. H. 2012 SVM, ANFIS, regression and climate based models for reference evapotranspiration modeling using limited climatic data in a semi-arid highland environment. *Journal of Hydrology* **444**, 78–89.

Thu, H. N. & When, U. 2016 Data sharing in international transboundary contexts: the Vietnamese perspective on data sharing in the Lower Mekong Basin. *Journal of Hydrology* **536**, 351–364.

Valdivieso, F. O. & Sendra, J. B. 2014 Semi-distributed hydrological model with scarce information: application to a large South American Binational Basin. *Journal of Hydrologic Engineering* **19** (5), 1006–1014.

Yatagai, A., Arakawa, O., Kamiguchi, K., Kawamoto, H., Nodzu, M. I. & Hamada, A. 2009 A 44-year daily gridded precipitation dataset for Asia based on a dense network of rain gauges. *SOLA* **5**, 137–140.

Yu, M. Y., Chen, X. L., Li, L. H., Bao, A. M. & Paix, M. Z. 2011 Streamflow simulation by SWAT using different precipitation sources in large arid basins with scarce raingauges. *Journal of Water Resources Management* **25**, 2669–2681.

Zhang, X., Srinivasan, R. & Van Liew, M. 2009 On the use of multi-algorithm, genetically adaptive multi-objective method for multi-site-calibration of the SWAT model. *Hydrological Processes* **24**, 965–969. doi:10.1002/hyp.7528.

Zhang, X., Srinivasan, R., Zhao, K. & Van Liew, M. 2008 Evaluation of global optimization algorithms for parameter calibration of a computationally intensive hydrologic model. *Hydrological Processes* **23**, 430–441. doi:10.1002/hyp.7152.