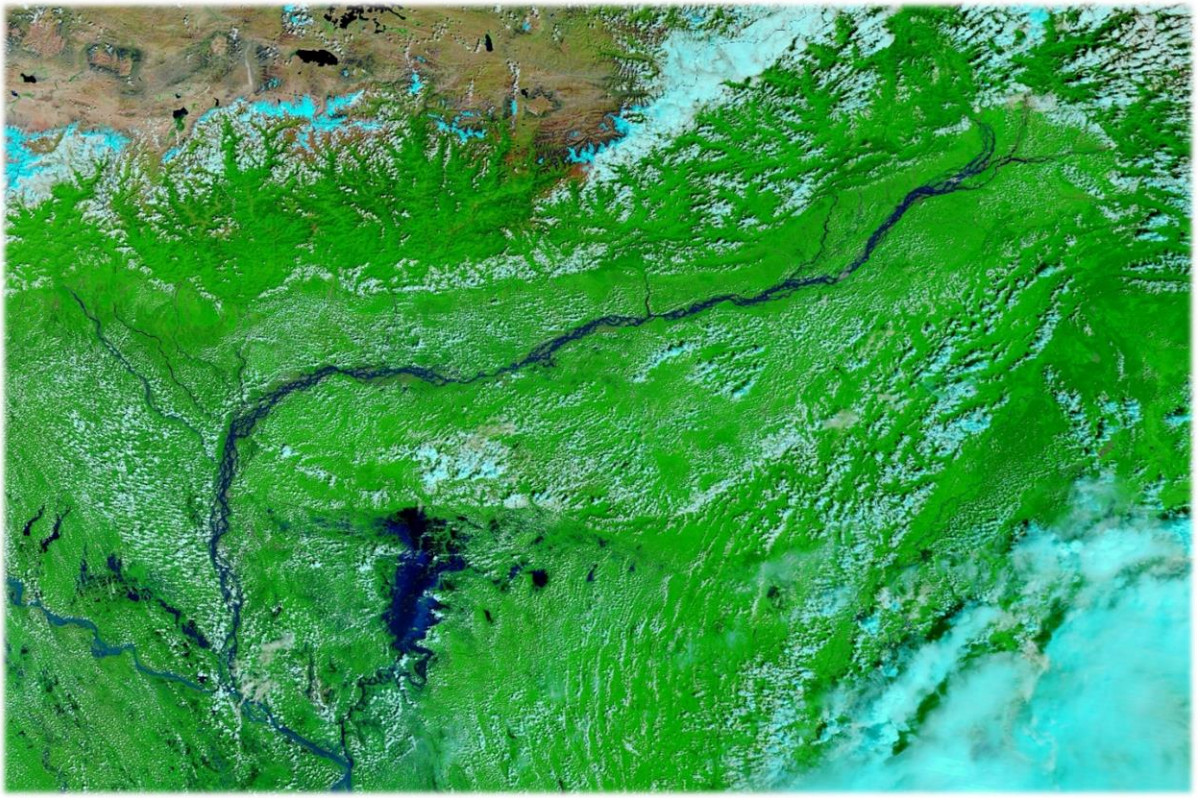
Modelling of Water Resources in the Brahmaputra River Basin



J. A. Bok

April 2016

Utrecht University

Modelling of Water Resources in the Brahmaputra River Basin

MSc Thesis

April 2016

Author: Koos Bok

Student number: 3537218

E-mail: j.a.bok@students.uu.nl

First supervisor: dr. ir. Geert Sterk (UU)

Second supervisor: ir. Patricia Lopez Lopez (UU, Deltares)

MSc programme: Master Earth Surface and Water

Faculty of Geosciences

Department of Physical Geography

Utrecht University

Utrecht University

Abstract In some parts of the world, network density of local hydro-meteorological station data is lacking. A solution to this is to use satellite or reanalysis data. In the present study, different interpolation techniques are applied to estimate precipitation and temperature using local observed data and compared with global reanalysis WFDEI forcing data in the Brahmaputra River Basin. The different forcing datasets were used to drive the global-scale hydrological PCR-GLOBWB model to estimate discharge on a roughly 10×10 km spatial and daily temporal resolution. Discharge model results were compared with observed records from the Bahadurabad gauging station. Results show that the interpolated in-situ temperature forcing data is generally warmer than the WFDEI forcing dataset. The interpolated in-situ precipitation forcing data consists of generally higher quantities than the WFDEI forcing dataset, especially during the monsoon period. Discharge is underestimated considerably using the WFDEI forcing dataset. Using the interpolated in-situ forcing data, the model overestimates the discharge slightly. Several dataset combinations and data modifications were applied. The run with interpolated in-situ precipitation forcing data in the upper part of the basin and WFDEI precipitation forcing data applied in the lower part of the basin, while using the interpolated local temperature forcing data, produced the best result. This study concludes that the WFDEI forcing dataset is useful and can provide reasonable estimates in combination with local hydro-meteorological information.

Keywords PCR-GLOBWB – WFDEI – Brahmaputra River Basin – Discharge – Monsoon

Table of Contents

1. Introduction

Bangladesh, located in the Ganges-Brahmaputra-Meghna delta, acts as the drainage outlet for the whole region. Annually, floods inundate large parts of the country imposing problems on water-related issues. The location and the orography of the region makes the Brahmaputra River Basin especially vulnerable to extreme events. Annual floods inundate approximately 20% of the country for a short time during the monsoon season. During extreme floods, for example in 1998, more than 60% of the country was inundated for nearly 3 months (Chowdhury, 2003; Mirza et al., 2003). In 1987 and 1988, similar events occurred however with a smaller coverage and smaller duration. More recently, the years 2004, 2007, and 2012 experienced extreme flooding. These events are primarily caused by intense monsoon precipitation (Mirza, 2003).

The average discharge of the Brahmaputra is measured to be approximately 20,000 m^3/s (Datta & Singh, 2004; Immerzeel, 2008) with measured extreme discharges of up to 100,000 m^3/s. The mean monthly discharge is highest in July and lowest in February. The high flow period that causes floods start in May and generally ends in the last weeks of October (Sarma, 2005).

The variability in flood impact has a profound effect on social and agricultural activities. Despite water abundancy in the region, water scarcity has its effect as well (Gain & Giupponi, 2015). Water scarcity is defined as unfavourable trends in water supply and/or demand caused by climate variability and socio-economic factors, i.e. population growth and increased food intake per capita (Immerzeel & Bierkens, 2012), which is inherent to natural spatial and temporal variability (Postel et al., 1996). Climate change and increase in population induces water stress, affects food security, endangers access to safe drinking water and public health, and threatens environmental well-being (Taylor, 2009).

The hydrological impact of climate change on the Ganges-Brahmaputra-Meghna Basin is expected to be particularly strong (Gain et al., 2011). It is projected to include: increase in flooded area for annual peak discharge by at least 23–29% (Mirza et al., 2003); significant increase in both peak discharge and flood duration for both pre-monsoon and monsoon seasons (Ghosh & Dutta, 2012); by mid-century, annual average river discharge increases by 10–40% (Milly et al., 2005); while enduring an intensification in water scarcity during the dry season in coming decades (Gain & Wada, 2014; Immerzeel et al., 2010). Immerzeel and Bierkens (2012) point out that the risk in severe flooding lie in the occurrence of extreme rainfall in combination with a higher mean sea level.

The beneficial impacts of projected increases in annual discharge will be tempered by adverse impacts of increased variability on water supply and flood risk, in particular in heavily populated low-lying deltas such as the Ganges-Brahmaputra-Meghna delta (Mirza et al., 2003). Immerzeel et al. (2010) stated that the Brahmaputra is most susceptible to reductions of flow, threatening the food and water security of an estimated 26 million people. As water scarcity will increase in dry seasons in future years (Gain & Giupponi, 2015), this signifies the need for strengthening long term water management policies and adaptation measures in Bangladesh to reduce increased flood hazard and water insecurity, as well as international basin-wide co-operation (Ahmad & Ahmed, 2003). This study is only a small step in creating a better global dataset to achieve this long-term goal.

1.1. Research Challenge

Due to the location of Bangladesh in the basin, upstream meteorological and hydrological data is required to establish a consistent model. However, the contemporary transboundary issues limit the availability of local data to drive local hydrological models at a high resolution. Thus, a scarcity of readily available hydro-meteorological station data exists in the Brahmaputra River Basin.

A possible solution to this challenge is to use global water resources reanalysis datasets into large-scale models. For example, the WATCH Forcing Data methodology applied to ERA-Interim reanalysis dataset (WFDEI) (Weedon et al., 2014). This dataset is suitable for driving hydrological models and land surface models with the range 1970 up to and including 2014. Two example global models used for processing such global data are the Variable Infiltration Capacity model (VIC) (Liang et al., 1994) and the PCRaster Global Water Balance model (PCR-GLOBWB) (Van Beek & Bierkens, 2009). The VIC model is based on energy balances, while the PCR-GLOBWB model is based on water balances.

1.2. Research Objective

The goal of the study was to compare the WFDEI global reanalysis forcing data (henceforth denoted as E2O) with in-situ observed data in the Brahmaputra River Basin, as well as model discharge result using the mentioned datasets. An important aspect was to evaluate the added value of the E2O dataset for basin-scale hydrological modelling and to assess improvements on the E2O dataset. Additionally, combinations of global and in-situ forcing data were assessed to achieve the best result. The study used the PCR-GLOBWB model with a five arc minute spatial resolution (roughly 10×10 km) and a daily temporal resolution.

2. The Brahmaputra River Basin

2.1. Basin Characteristics

The river basin area is defined from the source on the Tibetan Plateau to the confluence with the Ganges River after which it flows into the Bay of Bengal (Figure 1) (e.g. Immerzeel, 2008; Sarma, 2005). It drains an approximate area of 530,000 km^2 of which 50.5% lies in China, 33.6% in India, 8.1% in Bangladesh and 7.8% in Bhutan (Immerzeel, 2008). Along its 2900 km course, it flows through diverse environments. Immerzeel (2008) distinguished three distinct physiographic zones (percentage coverage in brackets): the Tibetan Plateau (44.4%) with elevations of 3500 m above sea level and higher, the Himalayan mountain belt (28.6%), and the floodplain (27.0%) with elevations of less than 100 m above sea level.

The Brahmaputra River originates from the Chema Yundung glacier in the Kailash range in South-west Tibet at an elevation of 5300 m.a.s.l. (Sarma, 2005). On the cold and dry Tibetan Plateau, the river is known as Tsangpo or Yarlung Zhanbo and flows an easterly course of 1625 km to the Himalayan belt with a general slope of 1.63 m km/s3. The slope of the (Dihang) river increases greatly while passing through the Himalayan Mountains: up to 16.8 m km/s3 (Figure 2).

Through the steep mountains, it enters the Assam Valley. The Assam Valley is confined by the Himalayas to the North and East and by the Meghalaya mountain reach to the South. The Himalayan Mountains is geologically young and active region and therefore the amount of available sediment is considerably high (Sarker et al., 2003). The valley receives large amounts of rainfall and has a warm humid climate (see Section 2.2.); therefore, the intensity of weathering is also high. This results in high amounts of transported sediment in the drainage network (Sarma, 2005). Upon entering the Assam Valley however, the slope decreases rapidly. This congests the channels resulting in a highly braided channel pattern (Ghosh & Dutta, 2012). After confluencing with two major tributaries further south, the river is called Brahmaputra and flows a south and westerly course of about 900 km through its alluvial plain. The slope consistently decreases over the Assam Valley down to 0.079 m km/s3 at the border

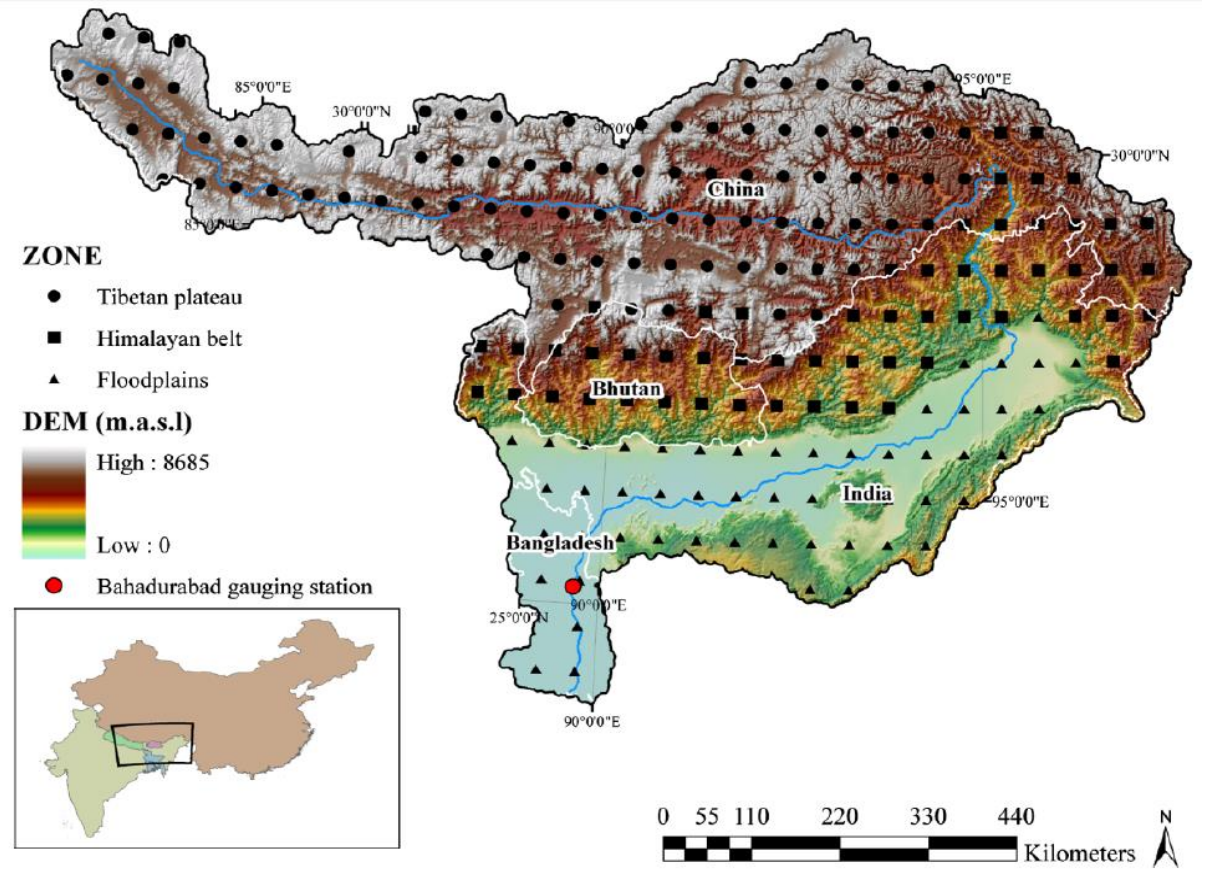


Figure 1. An overview of the Brahmaputra basin (Immerzeel, 2008).

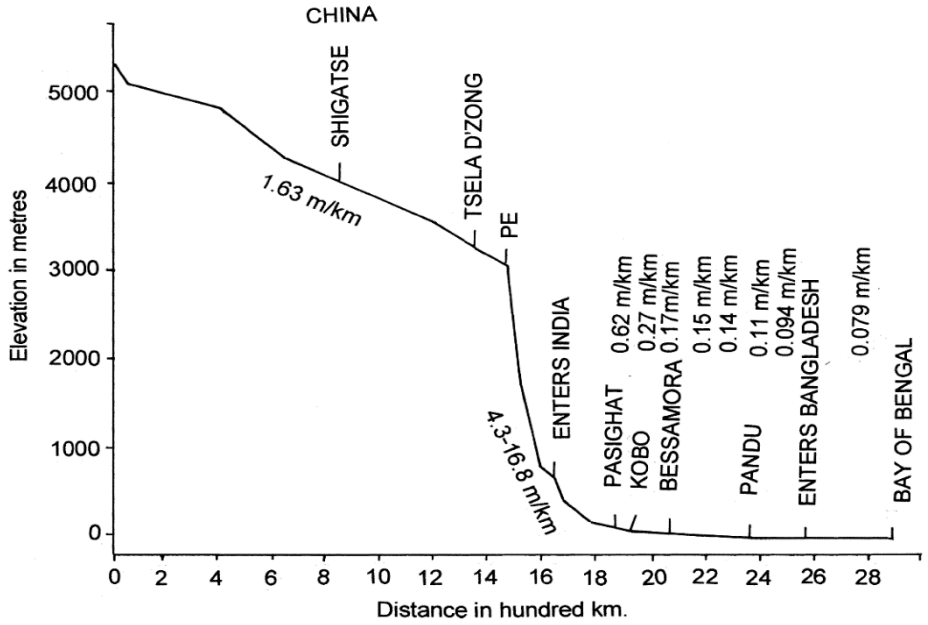


Figure 2. The longitudinal profile of the Brahmaputra River (modified after WAPCOS [1993] in: Sarma, 2005).

with Bangladesh (Sarma, 2005). The average width of the Brahmaputra riverbed varies between six and 18 km with few constraint reaches.

The basin is characterised by high seasonal variability in flow, sediment transport and channel configuration (Goswami, 1985). The integrated drainage network of the Brahmaputra and its tributaries consist of various river dimensions and types. Especially in the floodplain, one can find straight, sinuous, meandering, tortuous, braided, anastomosing, anabranching, reticulate and intermediate types. Subsequently, numerous palaeochannels exist in the region (Sarma, 2005).

From the Bangladesh border, the Brahmaputra River (locally known as the Jamuna River) turns southward with a reach of more than 300 km until the confluence with the Ganges River. The average annual sediment transport through this part of the river is nearly 600 M ton/year (Sarker et al., 2003). After the convergence, the river is known as the Padma River. Downstream, it merges with the Meghna River after which it is named Lower Meghna River and continues into the Bay of Bengal.

About 92.5% of the combined Ganges-Brahmaputra-Meghna basin area lies beyond the boundary of Bangladesh (Mirza, 2003). On average, annual floods inundates 20.5% of Bangladesh (about 3.03 million ha) (Mirza, 2003). In exceptional years, floods may even inundate about 70% of the country, as occurred during the floods of 1988 and 1998 (Ahmed & Mirza, 2000).

2.2. Meteorological Characteristics

Immerzeel (2008) published a chart with climate normals from 1961 to 1990 for the three physiographic zones in the Brahmaputra basin (Figure 3A). The Tibetan Plateau is characterised by the coldest average winter and summer temperature of -10°C and 7°C, respectively. In the mountain belt, the average winter temperature is around 2°C while the summer temperature is around 15°C on average. The lower Brahmaputra river basin or floodplain has the highest average winter and summer temperatures: 17°C and 27°C, respectively. In all zones, the temperature variation is largest during winter. Precipitation is concentrated during the monsoon from June to September for all distinguished regions (Figure 3B). The wettest region is the floodplain and receives annual precipitation of 2354 mm (Immerzeel, 2008). However, the rain distribution is heavily affected by the large-scale orography prominent effects on atmospheric flow patterns (Figure 4). Consequentially, annual rainfall is measured to be up to 5000 mm (Sarma, 2005) or even higher (Nepal & Shrestha, 2015). Precise simulation of the spatial and temporal behaviour of flow patterns in complex local topography is difficult to reflect in models due to the lack of observational data (Beniston, 2003; Immerzeel, 2008)

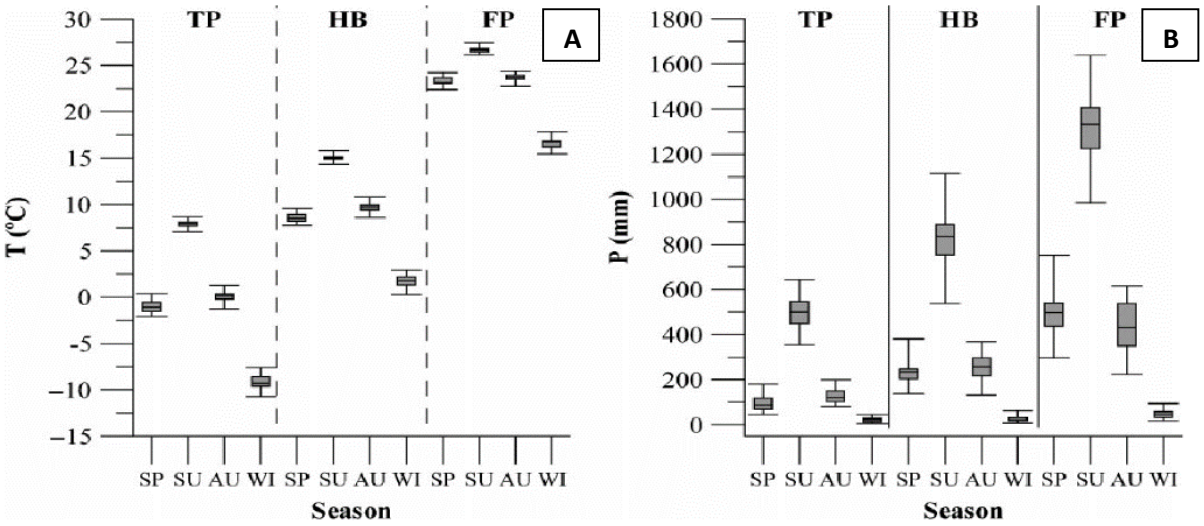


Figure 4. Box-whisker plots with seasonal climate normals (1961–1990) of temperature (A) and precipitation (B) for the Tibetan Plateau (TP), the Himalayan mountain belt (HB), and the floodplain (FP). Categorised for the spring (SP [=March, April, May]), summer (SU [=June, July, August], autumn (AU [=September, October and November], and during the winter (WI [=December, January, February]) (Immerzeel, 2008).

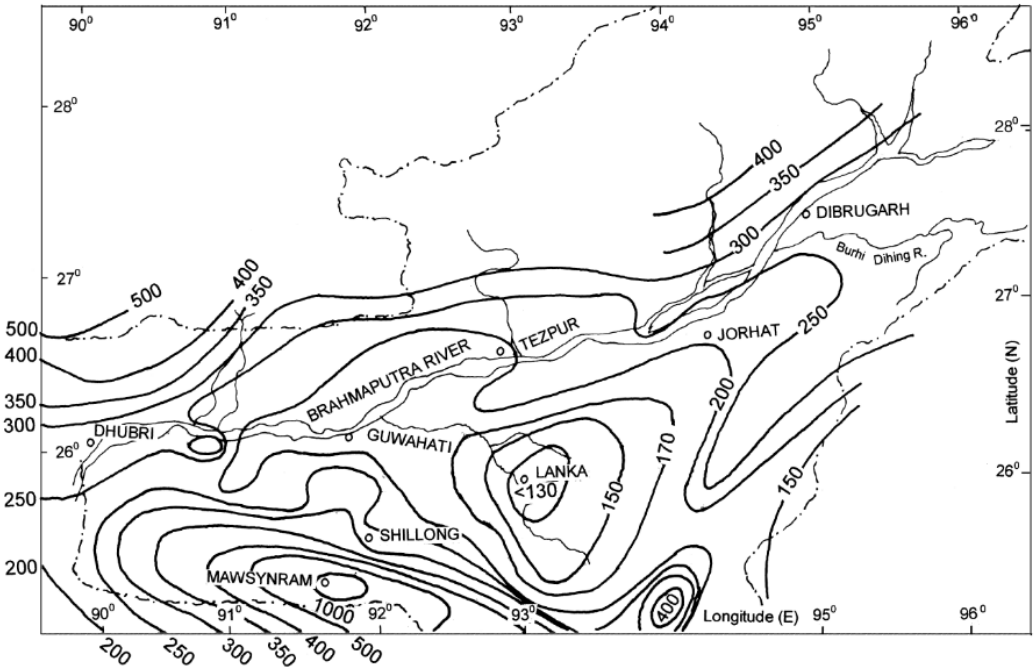


Figure 3. A normal annual isohyetal map of the Brahmaputra basin in the Assam Valley, India (unit in cm). (After WAPCOS [1993] in: Sarma [2005]).

Some 60–70% of the annual rainfall precipitates during the monsoon (Immerzeel, 2008; Mirza et al., 2003), with a further 20–25% during the pre-monsoon from March through May (Nepal & Shrestha, 2015; Sarma, 2005). Clusters of successive rainy days with around 100 mm per day are standard during the rainy seasons. Precipitation is characterised by an increasing trend from east to west along the Himalayas with its consequences for the dependent river basins (Immerzeel et al., 2009). The eastern part of the Brahmaputra basin experiences less precipitation due to the Meghalaya Mountain reach inflicting with large-scale wind directions during the monsoon (Figure 4). Additionally, snowmelt contributes significantly to the total river discharge (Immerzeel et al., 2009).

2.3. Hydrological Characteristics

Downstream water availability is sensitive to changes in snow and glacier extent (Immerzeel & Bierkens, 2012). Immerzeel et al. (2010) estimated that the discharge generated by snow and glacier melt in the Brahmaputra basin is 27% of the total discharge generated in the basin.

As mentioned above, precipitation occurs during the pre-monsoon and monsoon seasons with high intensity and quantity. These events cause quick hydrological response in the form of flood waves (Ghosh & Dutta, 2012). On average, the Brahmaputra River experiences four to five flood waves annually during the monsoon period (Datta & Singh 2004; Karmaker & Dutta 2010), causing floods in the floodplain. About 40% of the fluvio-deltaic plain of the Brahmaputra River basin is prone to flooding (Immerzeel, 2008). The physical factors contributing to this phenomenon include snow and glacier melt, the El Niño Southern Oscillation (ENSO) induced conditions, loss of drainage capacity due to the siltation of principal distributaries, backwater effect, unplanned infrastructure development, deforestation and the synchronisation of flood peaks of the major rivers of the delta (Mirza et al., 2003). Mirza (2003) compared three extreme floods (1987, 1988 and 1998) in Bangladesh and found intense monsoon precipitation was the primary cause of flooding.

Bangladesh generally experiences four main types of floods (Mirza et al., 2003): flash, riverine, rain and storm-surge. Eastern and northern areas of Bangladesh adjacent to its border with India are vulnerable to flash floods. Rivers in these regions are characterised by sharp rises and high flow velocities resulting from storm events occurring in neighbouring India. Riverine floods occur when the major rivers and their (dis-)tributaries flood due to increases in upstream river discharge. Rain floods are caused by intense local rainfall of long duration in the monsoon months. Heavy pre-monsoon rainfall causes local run-off to accumulate in topographic depressions. Local rain accumulates in ponds by rising water levels in adjoining rivers. Coastal areas of Bangladesh, which consist of large estuaries, extensive tidal flats, and low-lying offshore islands, are vulnerable to storm-surge floods, which occur during cyclonic storms. Cyclonic storms usually occur during April–May and October–November.

Floods are quite common during the (pre-) monsoon, while in the low-flow periods the river becomes a highly braided river with a large number of mid-channel and lateral bars (FAO, Aquastat, 2011). Due to the braided nature of the river, measuring the discharge can be difficult and only possible at some nodal points. At these constraint reaches, the bed level of the river generally undergoes aggradation and degradation throughout the year. The average discharge of the Brahmaputra is measured to be approximately 20,000 m^3/s (Datta & Singh, 2004; Immerzeel 2008). Sarma (2005) elaborates on the discharge characteristics. The average monthly discharge is highest in July and lowest in February. The high flow period that causes floods in the region starts in May and generally ends in the last weeks of October with peak flows generally in the range of 60,000 to 70,000 m^3/s with measured extreme peak discharges of 100,000 m^3/s. From November to April, discharge is relatively low: 5000 to 6000 m^3/s. Large variations in discharge are noticed during a flood, with the maximum increase of about 17,000 m^3/s in 24 hours (June 7–8, 1990) and 24,000 m^3/s in 48 hours (June 7–9, 1990) (Sarma, 2005). The maximum measured discharge reduction over 24 hours was 12,000 m^3/s at September 21–22, 1977.

3. Research Methodology

In this study, different datasets were used: the WFDEI global water resources reanalysis forcing data from the eartH2Observe project (E2O), and datasets based on locally observed data. The study will use the PCR-GLOBWB model with a 5 arc minute spatial resolution (roughly 10×10 km) and a daily temporal resolution. A spatial extent was created which encompasses the Brahmaputra River Basin (Appendix A). The model was not calibrated for this basin.

First, multiple forcing datasets were created from in-situ data from local stations using different interpolation techniques. Precipitation and temperature input were compared using timeseries and maps. Furthermore, an evaluation was done to investigate the influence of the different datasets in terms of estimated discharge. The calculated discharge patterns were analysed with daily, monthly and mean-monthly timeseries using the observed record at the Bahadurabad gauging station as reference. Additionally, verification metrics were used as indication of the performance of the model results. Finally, different combinations of datasets and modifications of the E2O forcing data were investigated.

3.1. Hydrological Model: PCR-GLOBWB

In this study, the PCR-GLOBWB model was used. PCR-GLOBWB is a deterministic process-based macro-scale lattice-based model describing global water balances of the terrestrial hydrology (Van Beek et al., 2011; Van Beek & Bierkens, 2009). Similar to other global hydrological models, it is essentially a "leaky-bucket" type of model (Bergström, 1995) applied on a cell-by-cell basis. The model is coded in the dynamic scripting language PCRaster (Wesseling et al., 1996). The PCRaster Environmental Modelling language is a high-level computer language that uses spatio-temporal operators with intrinsic functionality for constructing spatio-temporal models. It enables a very efficient manipulation of raster-based maps and has several built-in hydrological functions, such as accumulating and routing water and

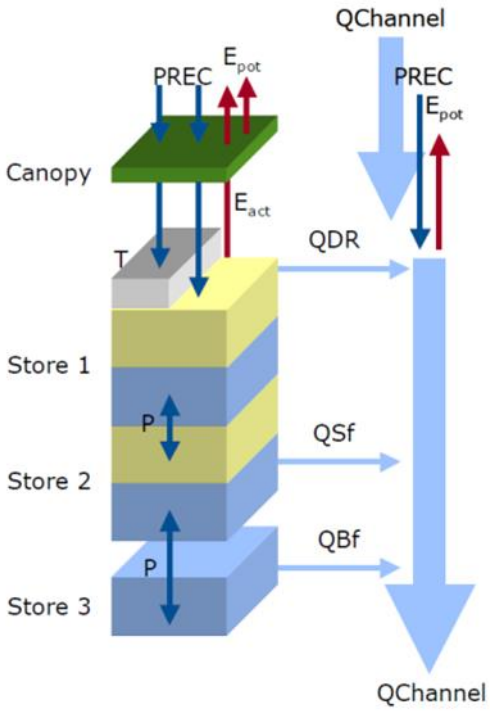


Figure 5. Model concept of PCR-GLOBWB (Van Beek & Bierkens, 2009). The left-hand side represents the vertical structure for the soil hydrology representing the canopy, soil column (stores 1 and 2), and the groundwater reservoir (store 3). Precipitation (PREC) falls as rain if the temperature (T) is above 0°C or as snow if below 0°C. Snow accumulates on the surface, and melt is temperature controlled. Vertical transport in the soil column is accounted for using percolation or capillary rise (P). Potential evapotranspiration (E_{pot}) is broken down into canopy transpiration and bare soil evaporation. This is reduced to actual evapotranspiration (E_{act}) on the basis of soil moisture content. On the right-hand side, drainage from the soil column to the river network occurs via direct runoff (QDR), interflow (QSf) or subsurface stormflow (QBf) as total specific discharge from one cell to the subsequent grid cell. Drainage accumulates as discharge (QChannel) along the drainage network and is subject to a direct gain or loss depending on the precipitation and potential evapotranspiration acting on the open water surface.

sediments over drainage networks. The model has been applied in various water resource balance studies, such as: calculating groundwater balances (Sutanudjaja et al., 2011; Wada et al., 2010), consumptive use of surface water and groundwater resources (e.g. irrigation and industry) (De Graaf et al., 2014; Gleeson et al., 2012; Wada et al., 2012, 2014), impact of climate change on irrigation (Wada et al., 2011, 2013), river discharge (Sperna Weiland et al., 2010; Bierkens & Van Beek, 2009), and impact of climate change on river discharge (Gain et al., 2011; Sperna Weiland et al., 2012). PCR-GLOBWB calculates for every grid cell and for every time step the water storage in two vertically stacked soil layers and for one underlying groundwater layer. A schematic overview is given in Figure 5. Water exchange between layers (capillary rise and percolation), and interaction between the top layer and the atmosphere (snowmelt, evapotranspiration and rainfall) are calculated, as well as snow storage and canopy interception. The process description presented in the remaining part of this Section is only a short overview and only describes modelled processes related to this study. For more elaborate information on the PCR-GLOBWB model, the author recommends Van Beek (2008), Van Beek and Bierkens (2009), Van Beek et al. (2011), and Appendix A of Sutanudjaja et al. (2011).

Climate forcing is applied at a daily resolution and assumed constant over the grid cell. Precipitation falls as liquid or solid depending on the air temperature 2 m above the ground. Excess precipitation either adds to the snow pack, adds to the liquid pore space in the snow pack, or infiltrates into the first soil layer. Snow accumulation and melt are temperature driven and modelled according to the Hydrologiska Byråns Vattenbalansavdelning (HBV) snow model (Bergström, 1995). Precipitation can be intercepted by the canopy (with finite storage capacity) and any intercepted water is subject to open water evaporation. Excess precipitation is added to the snowpack if the air temperature is less than 0°C ($T_{air} < 273\ K$). Above 0°C ($T_{air} \geq 273\ K$), precipitation and melt water are stored as liquid water in the available pore space in the snow cover, or passed on to the first soil layer. The evapotranspiration can be determined using either the Hamon method (Hamon et al., 1954) or the Penman-Monteith approach (Allen et al., 1998).

Instead of using per-cell values for vegetation and soil types, the model recognizes sub-grid variability by taking into account a fractionized land type coverage, considering tall and short vegetation, open water, and different soil types. Short vegetation extracts water from the top layer only, while tall vegetation extracts water from both soil layers. Calculation of the fraction of saturated soil (indicated with x) to assess direct runoff is based on the improved Arno scheme (Hageman & Gates, 2003) and the digital elevation model with surface elevations of the 1×1 km Hydro1K dataset (USGS EROS Data Center, 2006), and is derived by Van Beek and Bierkens (2009) as:

$$x = 1 - \left(\frac{W_{max} - W_{act}}{W_{max} - W_{min}}\right)^{\frac{b}{b+1}} \quad (1)$$

where W_{act}, W_{max} and W_{min} indicate cell-averaged total soil moisture storage and maximum and minimum storage capacities, respectively, all in [m]. These parameters refer to the area in the cell not specified as open water and are based on the FAO gridded soil map of the world (FAO, 1998). The parameter b [-] is a dimensionless shape factor that defines the distribution of soil water storage within the cell and is calculated based on the distribution of maximum rooting depths. This in turn is derived from the 1×1 km distribution of vegetation types from Global Land Cover Characteristics Database (GLCC) (Hagemann, 2002; USGS EROS Data Center, 2002).

3.1.1. Surface Runoff

As mentioned above, precipitation and melt water are stored as liquid water in the available pore space in the snow cover or passed on to the first soil layer. As such, the input to the first soil layer consists of both non-intercepted precipitation and snowmelt. Melt water is first stored in the snow pack up to a maximum storage capacity that is related to snow depth (in snow water equivalent) and an average snow porosity. Water stored in the snow pack may refreeze or is subject to evaporation. Snowmelt in excess of the snow water storage capacity is added to the precipitation. The sum of non-intercepted precipitation and excess snowmelt P_n infiltrates into the first soil layer if the soil is not saturated, while surface runoff occurs if the soil is saturated.

The actual water content corresponds to fractional saturation of the soil x, see Eq. (1). Surface runoff Q_{dr} is related to cell-averaged moisture storage W_{act} and net input P_n. However, the input is distributed first in the recipient grid cell (up to its maximum moisture storages) before new surface runoff is distributed to the subsequent grid cell. The direct runoff is given by (Sutanudjaja et al., 2011; Van Beek & Bierkens, 2009):

$$Q_{dr}(t) = \begin{cases} 0, & \text{if } P_n(t) + W_{act} \le W_{min} \\ P_n(t) - \Delta W_{act} + \Delta W \left[\left(\frac{\Delta W_{act}}{\Delta W} \right)^{\frac{1}{b+1}} - \frac{P_n(t)}{(b+1)\Delta W} \right]^{b+1}, & \text{if } W_{min} < P_n(t) + W_{act} \le W_{max} \\ P_n(t) - \Delta W_{act}, & \text{if } P_n(t) + W_{act} > W_{max} \end{cases} \quad (2)$$

with $\Delta W = W_{max} - W_{min}$, and $\Delta W_{act} = W_{max} - W_{act}$

Eq. (2) states that for a given P_n, Q_{dr} is only generated if $P_n + W_{act}$ exceeds W_{min}. The direct runoff contributes to the total specific discharge discussed in Section 3.1.5. Soil column infiltrated water is discussed in the next Section.

3.1.2. Vertical Water Exchange

As mentioned in Section 3.1.1., the sum of non-intercepted precipitation and excess snowmelt infiltrates into the first soil layer if the soil is not saturated. The downward fluxes q between the layers are equal to the unsaturated hydraulic conductivity k [LT/s3] of the top, Eq. (3) in Table 1. The unsaturated hydraulic conductivity depends on the degree of saturation s. The s is defined using the average soil moisture content of the layer $\bar{\theta}$, saturated soil moisture content θ_s and residual soil moisture θ_r (Eq. [4]). These variables can be obtained by dividing W, W_{max} and W_{min} by the thickness of the layer. The maximum depth of the two upper soil layers are 0.3 and 1.2 m, respectively.

If the relative degree of saturation of the top layer is smaller than that of the second layer ($s_1 < s_2$), an upward flux is generated (i.e. capillary rise) (Eq. [5]). This flux can be sustained by the soil moisture deficit in the top layer and is proportional to the unsaturated hydraulic conductivity of the second layer. For the groundwater layer, the upward flux is described in a similar way, except that (i) the conductivity is the geometric mean of the conductivity of the second and the third layer, (ii) it only occurs given the proximity of the water table, (iii) and that the resulting moisture content of the second layer cannot rise above the field capacity ($s_2 < s_{fc}$) (Van Beek & Bierkens, 2009): see Eq. (6). Here f_5

Table 1. A selection of equations used in the PCR-GLOBWB model for calculating vertical water exchange in the soil column (Van Beek & Bierkens, 2009). Symbols are explained below the table.

Purpose	Equation	
Downward fluxes between soil layers	$q_{1\to2}(t) = k_1\big(s_1(t)\big)$ $q_{2\to3}(t) = k_2\big(s_2(t)\big)$	(3)
Degree of saturation	$s = \dfrac{(\bar{\theta} - \theta_r)}{(\theta_s - \theta_r)}$	(4)
Upward flux between soil layers 2 and 1	$q_{2\to1}(t) = \begin{cases} k_2(s_2) \cdot (1 - s_1), & \text{if } s_1 < s_2 \\ 0, & \text{if } s_1 \ge s_2 \end{cases}$	(5)
Upward flux between soil layers 3 and 2	$q_{3\to2}(t) = \begin{cases} \sqrt{k_{s,3} k_2(s_2)} \cdot (1 - s_2) 0.5 f_5, & \text{if } s_2 < s_{fc} \text{ and } W_3 > 0 \\ 0, & \text{otherwise} \end{cases}$	(6)

In Equation (3): s_1 and s_2 are the degree of saturation of layers 1 and 2, and k indicates the unsaturated hydraulic conductivity [LT-1]. In Equation (4): $\bar{\theta}$ is the average soil moisture content of the layer, θ_s is the saturated soil moisture content, and θ_r is the residual soil moisture. In Equation (6): f_5 is the fraction of the grid cell with a groundwater depth within 5 m, the factor 0.5 is an estimate of the average capillary flux over the area fraction f_5 with a groundwater table within 5 m depth, and W_3 is the water storage in the groundwater layer.

denotes the fraction of the grid cell with a groundwater depth within 5 m. It is thus assumed that capillary rise is at maximum if the groundwater table is at the surface and 0 if it is 5 m or lower below the surface. The factor '0.5' is an estimate of the average capillary flux over the area fraction f_5 with a groundwater table within 5 m depth, and W_3 is the water storage in the groundwater layer. Fluxes between the second soil layer and the groundwater reservoir are mostly downward, except for areas with shallow groundwater tables, where fluxes from the groundwater reservoir to the soil reservoir are possible during periods of low soil moisture content.

3.1.3. Interflow

Interflow is modelled by a simplified approach based on the work of Sloan and Moore (1984) in which the soil is idealized as a uniform, sloping slab with an average soil depth and inclination:

$$Q_i(t) = \left(1 - \frac{\Delta t}{CRT}\right) Q_i(t - \Delta t) + \left(\frac{\Delta t}{CRT}\right) \cdot L_{slope}[q_{12}(t) - q_{23}(t)],$$
$$\text{with } CRT = \frac{L(\theta_s - \theta_{fc})}{2k_{s,2}\tan(\alpha)}, \tag{7}$$

where Q_i is interflow per m slope width [L^{2}T/s3], L_{slope} is the average slope length or drainage distance [L], and Δt is the discrete time step. Please note that $q_{12}(t)$ and $q_{23}(t)$ are the net fluxes between indicated layers [LT/s3]. The parameter CRT [T] is a characteristic response time with θ_{fc} soil moisture content at field capacity, θ_s is the saturated soil moisture content, $k_{s,2}$ is the saturated hydraulic conductivity in layer 2, and α is the average slope between the cells. The average slope is determined from the average of calculated slopes from Hydro1K dataset. Interflow is only calculated if the soil water content in the second layer is above field capacity ($s_2 \geq s_{fc}$). Additionally, interflow is only calculated in mountainous areas. For each grid cell, the fraction of soils with a soil depth smaller than 1.5 m is determined and interflow from this fraction is calculated (Van Beek & Bierkens, 2009).

3.1.4. Baseflow

Besides surface runoff and interflow, the groundwater layer contributes to the total specific discharge via baseflow Q_b. The groundwater reservoir is infinitely large; however, the active groundwater storage is computed by assuming a linear relationship between storage and outflow (Eq. [8]):

$$Q_b(t) = \frac{W_3(t)}{J},$$
$$\text{with } W_3(t) = \left(1 - \frac{\Delta t}{J}\right) W_3(t - \Delta t) + q_{23,n}(t), \tag{8}$$

In this equation, J means a reservoir coefficient which represent the average residence time of water in the groundwater reservoir (Van Beek & Bierkens, 2009). W_3 is the water storage in the groundwater layer, Δt is the discrete time step, and $q_{23,n}(t)$ is the net flux between layer 2 and 3 [LT/s3]. Van Beek and Bierkens (2009) state that large-scale aquifer thickness information is lacking and is therefore unreliable. A constant aquifer thickness of 50 m is assumed.

3.1.5. Surface Water Accumulation and Routing

The total simulated specific runoff of a (local) cell Q_{loc} [LT/s3] consists of surface runoff Q_{dr}, interflow Q_i and baseflow Q_b (also appointed as QDR, QSf and QBf in Figure 5.) (Eq. [9] in Table 2). To acquire the lateral specific discharge q [L^{2}T/s3] (i.e. drainage) (Eq. [11]), Q_{loc} is used to determine the total accumulated local runoff Q_{tot}. This is determined by correcting for direct gain and loss by precipitation P [LT/s3] and reference potential evapotranspiration $ET_{0,wat}$ [LT/s3] over the open water fraction f_{wat} in the cell (sub-grid variability) (Eq. [10]). Open water evaporation occurs at the potential rate, with different crop factors being applied to deep water (lakes and reservoirs) and shallow water (river stretches) as suggested by Allen et al. (1998).

Table 2. A selection of equations used in the PCR-GLOBWB model to determine the total lateral in-flow and accumulated discharge (Sutanudjaja et al., 2011; Van Beek & Bierkens, 2009). Symbols are explained below the table.

Purpose	Equation	
Total local runoff	$Q_{loc} = Q_{dr} + Q_i + Q_b$	(9)
Direct inputs to open water surface	$Q_{wat} = P - ET_{0,wat} k_{c,wat}$	(10)
Lateral influx	$q = \frac{Q_{tot}}{L_{cell}} = \frac{A_{cell}}{L_{cell}} [(1 - f_{wat}) Q_{loc} + f_{wat} Q_{wat}]$	(11)
Total discharge	$\frac{\partial Q}{\partial L_{ch}} + C_1 C_2 Q^{C_2 - 1} \left(\frac{\partial Q}{\partial t}\right) = q$ with $C_1 = \left(\frac{n P_w^{2/3}}{\sqrt{S}}\right)^{3/5}$ and $C_2 = 3/5$	(12)
Manning's equation	$Q = \frac{R^{2/3}\sqrt{S}}{n} A_{ch} \rightarrow$ [substitute R by A_{ch}/P_w] $\rightarrow A_{ch} = Q^{3/5} \left(\frac{n P_w^{2/3}}{\sqrt{S}}\right)^{3/5}$	(13)

In Equation (9): surface runoff Q_{dr}, interflow Q_i and baseflow Q_b, all in [LT/s3]. In Equation (10): precipitation P [LT/s3], reference potential evaporation over water $ET_{0,wat}$ [LT/s3], and $k_{c,wat}$ [-] is the reference crop factor coefficient assumed for water bodies. In Equation (11): total lateral flow Q_{tot} [LT/s3], area of the cell A_{cell} [L²], length of the cell L_{cell} [L], and the open water fraction in the cell f_{wat}. In Equation (12): L_{ch} is the channel's length [L], n is Manning's roughness coefficient [L$^{5/6}$T/s3], P_w is the wetted perimeter [L], S and is the gradient. In Equation (13): R is the hydraulic radius [L] and A_{ch} is the cross-sectional area of the channel [L²].

The specific discharge of each cell is then routed over a drainage network that defines flow to one of the eight adjacent cells, according to the local drainage direction (LDD) method. This drainage network either terminates at the ocean or at an inland sink in the case of land-locked basins (i.e. pits). River discharge Q [L^{3}T/s3] is calculated by accumulating and routing the total specific discharge to the subsequent grid cell using the kinematic wave approximation of the Saint-Venant equation (Chow et al., 1988) (Eq. [12]). Eq. (12) is a combination form of the momentum and continuity equation. A numerical solution of the kinematic wave approximation is available as an internal function in PCRaster in which the new discharge Q_{t+1} at every point along the channel is calculated from the discharge from the previous time step. The coefficients C_1 and C_2 are determined using Manning's equation (Eq. [13]) and are passed over the LDD to the downstream cell. At the end of the time step, the calculated discharge is used to retrieve the new stage, which is calculated under the assumption of a rectangular channel with known channel depth and width. The new stage is passed to the next time step to estimate the wetted perimeter P_w for the calculation of C_1.

3.2. Data Description

3.2.1. Meteorological Forcing Data

3.2.1.1. WFDEI Dataset

The global forcing data used in this study is based on the WFDEI (WATCH Forcing Data methodology applied to ERA-Interim data) product (Weedon et al., 2014). This dataset is based on its forbearer WFD dataset (based on ERA-40 reanalysis data) (Uppala et al., 2005) created by the EU WATCH (Water and Global Change) project to directly compare global hydrological model output of different models with the range 1958 to 2001 (Haddeland et al., 2011; Harding et al., 2011, Harding & Warnaars, 2011). The WFDEI dataset is comprised of both satellite and local data. The product consists of precipitation and temperature data. This study uses a finer resolution version of the original WFDEI product. The WFDEI data was downscaled in accordance to the method explained in Weedon et al. (2014). The downscaling is based on the difference in elevation between the low-resolution digital elevation map (DEM) and a high resolution DEM. The downscaled data is calculated for each cell of the high resolution DEM. The

downscaled version has a resolution of 5 arc minutes, approximately 10×10 km. A practical improvement of the WFDEI opposed to the WFD dataset, is that the dataset ranges from 1979 to 2012, which allows for intercomparison with satellite data products (Weedon et al., 2014). This dataset is hereafter indicated as E2O.

3.2.1.2. In-situ Datasets

The region is considered to have poor available data due to the non-data sharing policy between involved countries. Only in recent years this has begun to change. The precipitation and temperature data (Appendix B) was acquired from two main sources. It was either made available by the Institute for Water Modelling (IWM) or from the weather website Weather Underground (2015). This website is not scientific and weather enthusiasts can link their personal weather stations to this website. The data has certain shortcomings. For instance, many measurement stations do not have a complete record and some stations switch from daily to monthly data (or vice versa) in the record. It was opted to filter out unreliable records and choose reliable stations such as airports and governmental buildings.

The gathered temperature data has a combined period of 2004 to 2015. The number of stations with precipitation data is higher. The observed precipitation data in Bangladesh have the best records with the longest data range (1970-2014). The average percentage of missing values for temperature and precipitation are 1.85 ± 2.33% and 9.03 ± 7.03%, respectively. However, the stations in India and China have substantial smaller time periods. These sets of records limit the range and analysis of the in-situ dataset to a study period of four years: from 2009 up to 2014.

In Figure 6 the distribution of the stations is shown. The density of the station network varies over the basin significantly. The lower basin is well represented, especially considering the precipitation data. However, a notable lack of stations is present in the upper part of the basin. The North-Western part of the basin is not represented. The temperature data is more evenly distributed. These stations are regarded as point data from which the data could be interpolated to create a spatial grid. This was done in ArcMAP (Appendix C). Interpolating techniques applied in the study are Inverse Distance Weighting (IDW), Spline, and Ordinary Kriging (OK) (Childs, 2004; Davis & Sampson, 1986).

The IDW and Spline interpolation methods are deterministic interpolation methods. These methods are directly based on the surrounding measured values. The IDW interpolation method determines cell values using a linear-weighted combination of sample points. This method assumes that the variable decreases in influence to an unknown location the further the known location is. The value assigned to the unknown locations is determined by averaging values of specified amount of known locations or of

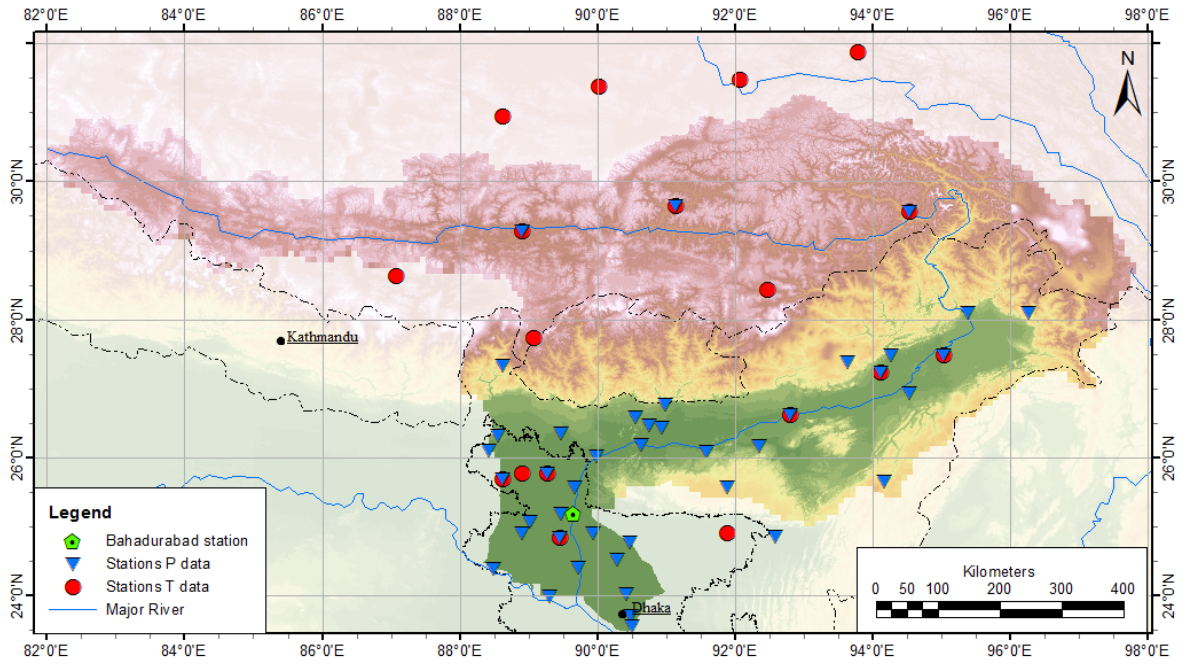


Figure 6. The Brahmaputra River basin with the locations of the stations with observed temperature (T) and precipitation (P) data. Additionally, the location of the Bahadurabad gauging station is shown.

values within a specified search radius. The weight of the values is a function of the distance of the known point to the unknown cell.

The Spline interpolation technique estimates values to minimize surface curvature. Spline bends a surface that passes through the input locations while minimizing the curvature. It fits a mathematical relation to a specified number of known input locations. It is best used to estimate a variable that is smooth in nature, such as temperature and water table heights. However, it needs a sufficient number of known locations.

As opposed to deterministic methods, OK is a geostatistical interpolation method. Geostatistical methods are based on statistical models for the relation among measured points. Given this fact, geostatistical methods, such as OK, are capable to produce a prediction surface as well as provide a measure of certainty of the predictions. OK assumes there is no trend in the data (i.e. no constant mean). OK assumes that the distance between known locations reflects a spatial correlation that can be used to explain variation in the surface. The kriging method applies a 'best fit' to a specified number of locations or to a number of locations in a specified search radius, to determine the output value of the unknown point. This is done by exploratory statistical analysis of the data, creation of a semi-variogram, and the creation of a surface. Kriging is most appropriate when a spatial correlation is present in the data.

Generally, IDW is smoother than the Spline method and the Spline method is smoother than the OK method. Due to the fact that the known locations are not well distributed over the entire basin, the interpolation method is used for extrapolation as well. The three datasets distinguished will hereafter be identified as In-situ OK (IOK), In-situ Spline (ISpline), and In-situ IDW (IIDW).

3.2.2. Observed Discharge Data

Measured discharge data was taken from the Bahadurabad gauging station (25°18'N, 89°66'E), which is North of the confluence of the Brahmaputra River with the Ganges River (Figure 1 and Figure 6). The raw data was provided by the Bangladesh Water Development Board and the Institute for Water Modelling (BWDB, 2015) and consisted of two datasets. The datasets have an overlap period in 2012. The data represents daily and irregular timed 3-hourly rated discharge data (Appendix D).

On aggregating the data, it was chosen to use the daily averaged discharge values. The mean deviation of average discharge of a single day to all 3-hourly measurements of that same day is 0.36%. Secondly, it was assumed that the datasets with overlap periods were better represented by the younger dataset when the difference was minor. Thus in the case of an overlap period in the 2012, dataset 2 was chosen over dataset 1. This resulted in an observed discharge dataset covering a time period of 1985 to 2014 on a daily time scale (Figure 7).

According to literature, the average discharge of the Brahmaputra is approximately 20,000 m^3/s (Datta & Singh, 2004; Immerzeel 2008). The average discharge of this dataset is determined to be 21,674.21 m^3/s. The median discharge is 14,463.96 m^3/s and the maximum value present in this database is 103,166.9 m^3/s. However, there are gaps in the dataset and these gaps occur mostly during low flow conditions. An extreme example is displayed in Figure 8. The original dataset 1 and 2 are primarily used for flood forecasting by flood forecasting agencies (e.g. Flood Forecasting & Warning Centre of the Bangladesh Water Development Board) and research institutes (e.g. Institute for Water Modelling). As such, low flow periods are deemed less important. Therefore, it can be safely assumed that the actual average discharge would be lower. Of the completely aggregated dataset, 3.35% of days have missing discharge values. Over the last ten years, the percentage of missing values is 9.26%.

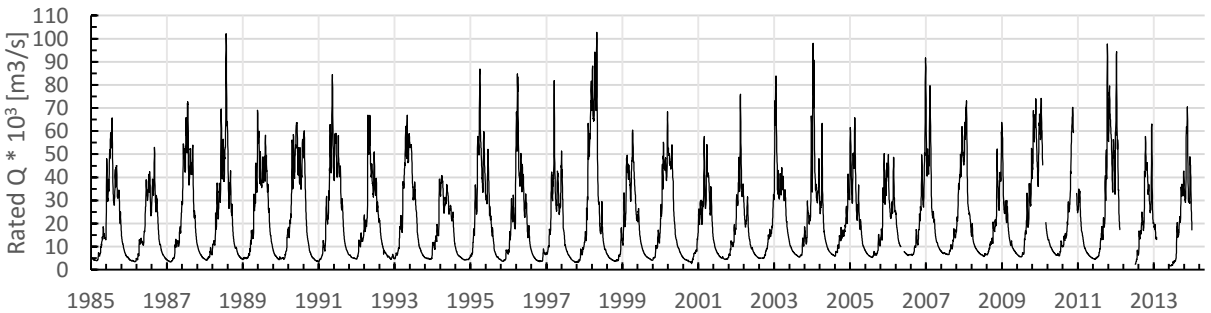


Figure 7. The aggregated measured discharge over the period 1985-2014. The years with extreme flooding (1988, 1998, 2004, 2007, and 2012) can clearly be observed.

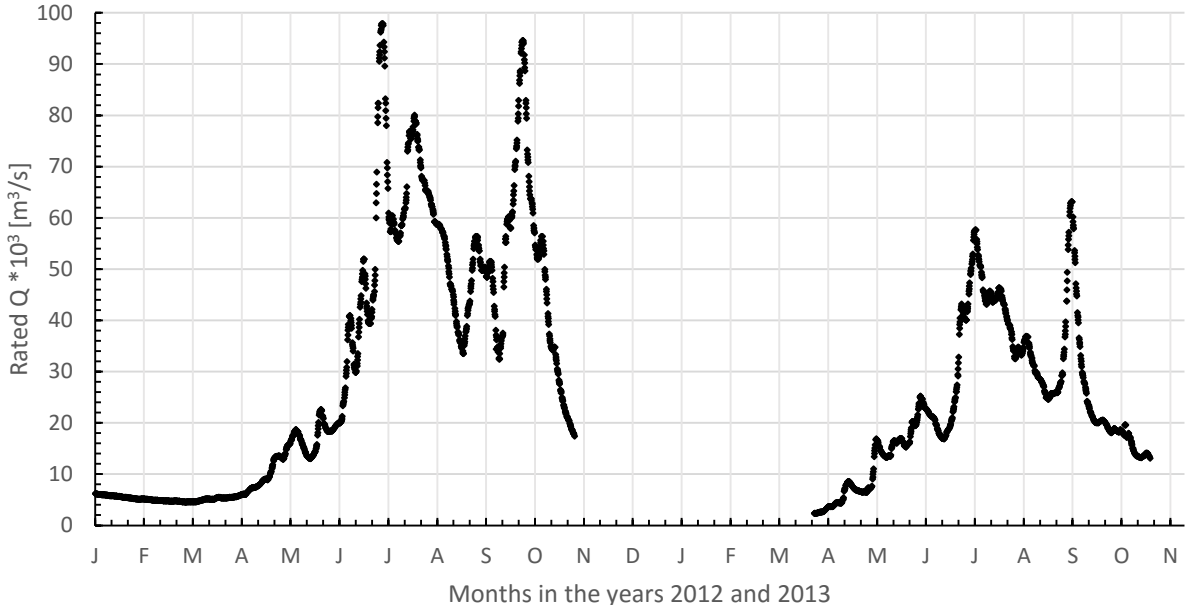


Figure 8. An example of missing data in the observed discharge data at the Bahadurabad gauging station.

3.3. Combined and Modified Forcing Datasets

3.3.1. Bias-Correction Method

Different combinations of datasets and modifications of the E2O forcing data were investigated to improve on the E2O dataset. Bias correction have been applied in many studies and in the studies of Huffman et al. (1997) and Adler et al. (2000) effort was made to create a (multi-)satellite-gauge-model to combine satellite and gauge station precipitation records to improve the prediction of precipitation in sections not covered by gauging stations. The studies used different techniques, one using an additive bias correction while the other study performed a ratio bias correction. Vila et al. (2009) combined these techniques and opted for the minimum correction.

To elaborate, the difference between additive bias correction and observed values, and the difference between ratio bias correction and observed values were calculated. The next step chooses the minimum absolute correction of the two methods and adds to the former E2O value, and performs this adjustment per cell and per timestep. The benefit of choosing the minimum forcing correction is that this results in a minimum discharge change according to these two methods. A second advantage of applying this method is that a correction value is created per cell and per timestep. The disadvantage of using this approach is that the correction is applied to an already interpolated field. The correction fields are created with conditional statements resulting in Boolean maps. Two example Boolean maps are shown in Figure 9. The equations involved in this process are shown in Eq. (14):

$$P_{E2O,new} = \begin{cases} P_{E2O} + (\bar{P}_{IOK} - \bar{P}_{E2O}) = A, & \text{if } |A - P_{E2O}| < |B - P_{E2O}| \\ P_{E2O}\left(\dfrac{\bar{P}_{IOK}}{\bar{P}_{E2O}}\right) = B, & \text{if } |A - P_{E2O}| \geq |B - P_{E2O}| \end{cases} \quad (14)$$

in which $P_{E2O,new}$ is the new precipitation value, P_{E2O} is the former precipitation value of the E2O forcing dataset, $\bar{P}_{IOK}$ and $\bar{P}_{E2O}$ are the mean precipitation value of that particular cell of the indicated dataset.

Additionally, two regular ratio bias corrections to the E2O forcing were performed, one with 25% and one with 40% more precipitation input.

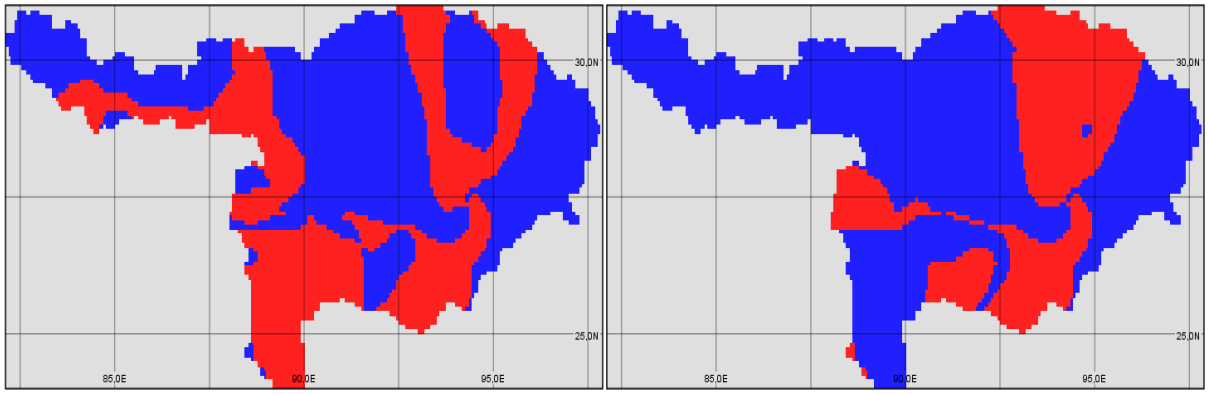


Figure 9. Two examples of the boolean condition map where red is 1 and blue is 0. These images are of the dates 26 and 27 September, 2011. If the statement is true (1), then the value of the additive bias correction was used, else the ratio bias correction was used.

3.3.2. Forcing Data Combinations

A different approach to adjust the input forcing data is to acknowledge that the region consists of two distinct physiographic and climatological regions (or even three [Immerzeel et al., 2008]). One being the upper part of the basin and one being the lower part of the Brahmaputra River Basin. Thus, a crude separation was made at latitude 28.50. The study was limited in creating straight separations along latitude and longitude lines. This separation is based on the fact that the southern section is represented by many stations contrary to the northern part (i). This makes for easy comparison between interpolated in-situ forcing data and E2O forcing data input. The temperature distribution is characterised by a clear separation along the Himalayan Mountains (ii) (Section 4.1.1). Mean difference in daily precipitation can be regarded as two separate regions, based on the ratio (iii) and the linear difference (iv) (Section 4.1.2). The precipitation in the upper region is regarded to be generally overestimated using the IOK dataset. Making the cut arguably slightly too high decreases the influence of the IOK precipitation in the northern part (v). The modelled snow cover is left to the upper part (vi). This approach results in four additional model runs: IOK precipitation data applied in the northern part and E2O precipitation data in the southern part of the basin or vice versa, and E2O or IOK temperature data applied over the whole basin.

Additionally, interpolated in-situ and E2O forcing data were combined by using local precipitation and global temperature, and vice versa. Furthermore, the difference in using Hamon or Penman-Monteith (P-M) approach has been investigated. For this purpose, the E2O evapotranspiration (ET) data has been applied in combination with both the interpolated in-situ forcing datasets and the E2O forcing data.

3.4. Verification Strategy

Model runs were performed with the E2O and interpolated in-situ forcing datasets. Temperature and precipitation forcing analysis was done by comparing spatial patterns of time-averaged maps and field-averaged daily values resulting in timeseries. The timing and amplitudes of minima and maxima in timeseries were analysed. As well as directly comparing input forcing data with observed station data at selected locations. Additionally, E2O forcing was compared with observed data at known locations. Similarly, this approach was applied to the resultant evapotranspiration of the model runs.

Furthermore, model estimated discharge was compared with observed discharge at the Bahadurabad gauging station. The estimated discharge patterns were analysed for timing and discharge values on daily, monthly and mean-monthly time scales. Additionally, the discharge was evaluated using statistical verification metrics as indication of the performance of the model-estimated discharge compared with the observed discharge (Helsel & Hirsch, 2002):

- The Nash–Sutcliffe model efficiency coefficient (NSE) (Moriasi et al., 2007; Nash & Sutcliffe, 1970) is used to assess the predictive power of hydrological models. The efficiency can range from $-\infty$ to 1. An efficiency of 1 corresponds to a perfect match of observed data and modelled discharge. An efficiency of zero means that the model results are as accurate as the mean of the observed data. An efficiency of less than zero occurs when the observed mean is a better predictor than the model.

- Pearson's correlation (r) is a measure of degree of linear correlation between measured and estimated discharge values. A value of 1 indicates a perfect correlation while 0 means no correlation, a value of -1 indicates a negative correlation.
- The mean absolute error (MAE) indicates how close the predicted and observed values are. Relative small values indicates that the predicated and observed values are similar.
- The root-mean-squared error (RMSE) is used as a measure of accuracy between the predicted and observed value. A relative large value indicates a low accuracy, whereas a relative smaller value indicates a higher accuracy.
- The mean values of observed and estimated discharge were calculated.
- Additionally, the bias is the difference between the model estimated value and the observed value.

4. Results

4.1. Model Input Analysis

4.1.1. Temperature

Figure 10 displays the mean temperature over the whole study period. There is a large difference in mean temperature over the basin. The upper part is colder than the lower part of the basin. ISpline performs poorly in the areas over which it needs to extrapolate: it calculated mean temperature values of -30°C in the North-Western area and +80°C in the South-Eastern area. Thus, the use of ISpline was discontinued at an earlier stage. The results using IOK and IIDW forcing data are better and display a similar pattern as the E2O forcing data model run.

The temperature data of the E2O forcing dataset in the northern part of the basin is significantly colder than the IOK and IIDW datasets. The mean temperature of E2O is around 3.25°C, while the interpolated in-situ datasets have a mean value of 6.11°C. In the lower part of the basin, the E2O dataset is only slightly warmer than the interpolated in-situ datasets.

Figure 11 shows the mean temperature of the whole basin of every timestep. A significant bias can be observed between the different forcing datasets. The difference between IOK and E2O is 3.17 ± 0.70°C and the difference between IIDW and E2O is 3.33 ± 0.75°C. However, as can be seen from Figure 12, the distribution of mean temperature difference is not uniform over the whole basin. In the lower part of the basin the temperature difference is between +1.0°C and -3.0°C, with the exception of the area at approximately (27.0, 89.0): here the difference is up to -12.0°C. In the upper part of the basin, the mean temperature difference is between +2.4°C and +14.5°C. There seems to be a large mean temperature difference in the North and a smaller mean temperature difference in the South of the basin between E2O forcing data and the interpolated in-situ forcing datasets.

Two stations were chosen for a more detailed analysis: one in the upper basin: Lhasa (29.65, 91.14) (Figure 13), and one in the lower part of the basin: Rangpur (25.76, 89.28) (Figure 14). The mean temperature difference is generally present over the whole study period for both locations: the temperature difference at the Rangpur station is small and E2O forcing values have a generally higher temperature value than the in-situ datasets. However, the temperature difference at Lhasa is large and E2O forcing values are significantly lower compared with interpolated in-situ forcing data.

Considering the temperature spatial differences and distribution, it differs significantly over the basin, however a distinction can be made between the upper and lower part of the basin.

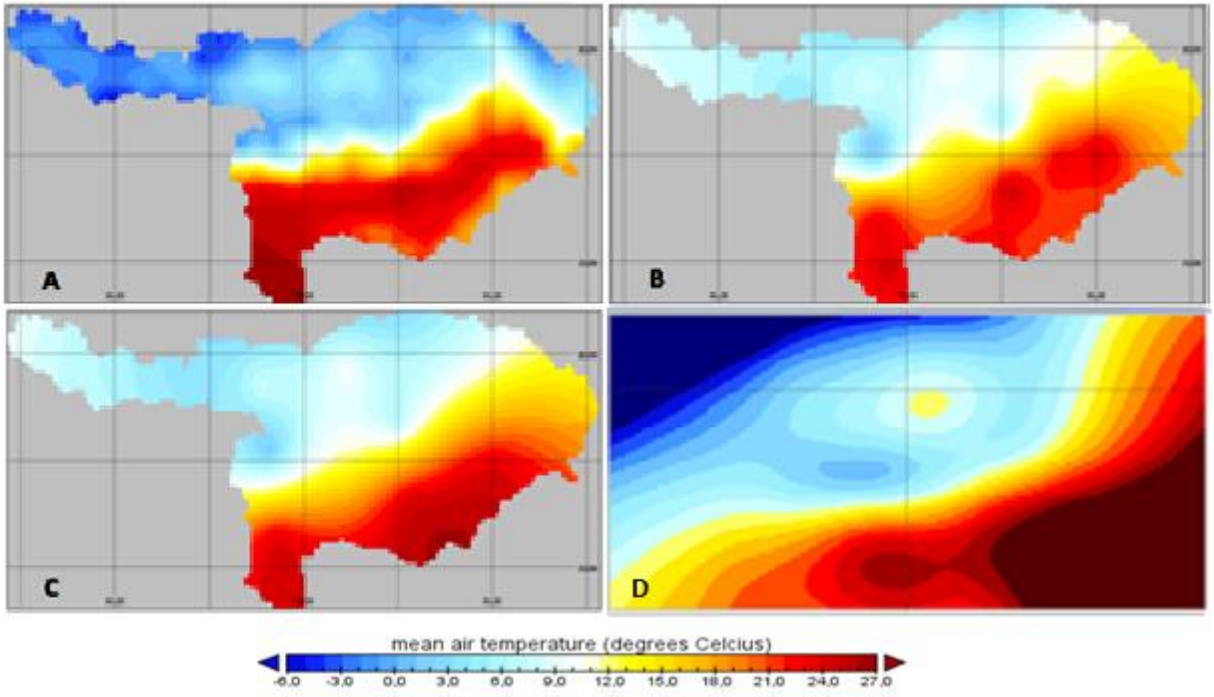


Figure 10. Mean temperature value distribution over the Brahmaputra River Basin over the whole study period 2009-2012. A – map based on the E2O forcing data input. B – map based on the interpolated station data using IDW. C – map based on interpolated station data using OK. D – map based on interpolated station data using Spline (it was discontinued in an earlier stage of the research).

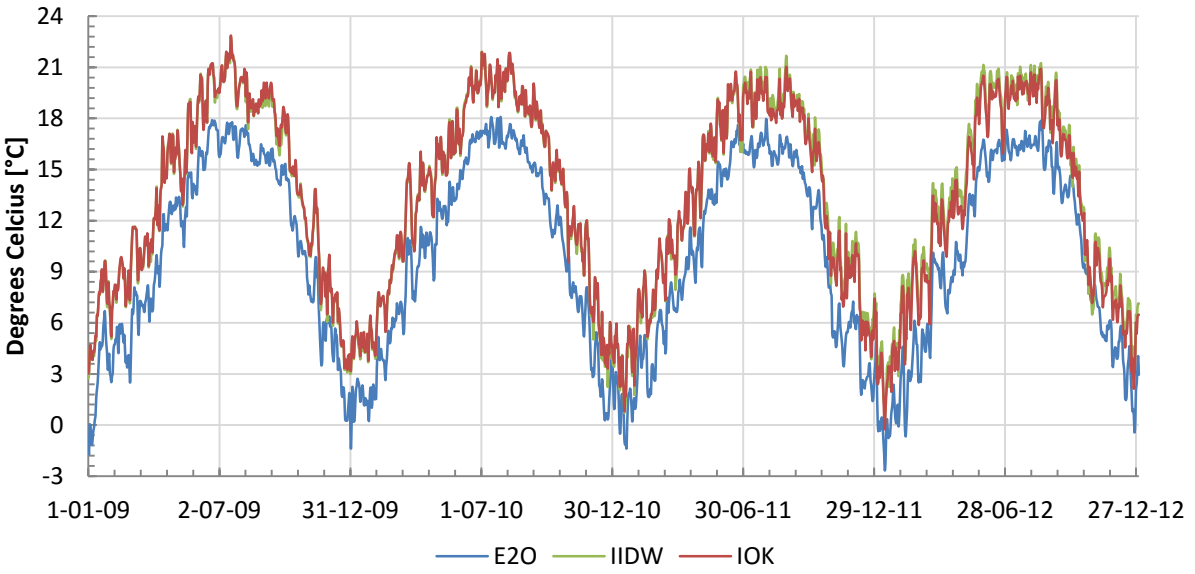


Figure 11. Basin-mean temperature timeseries. The basin mean temperature was calculated per timestep. The interpolated in-situ IOK and IIDW datasets are generally warmer than the E2O forcing dataset.

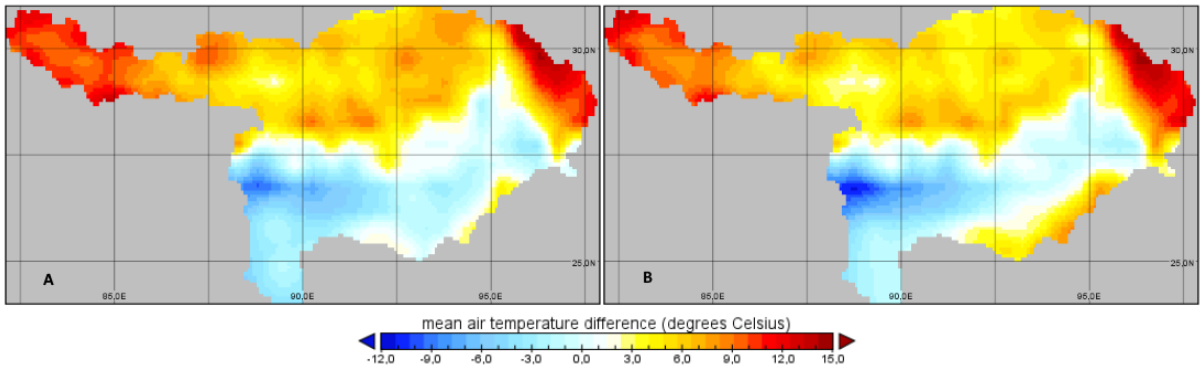


Figure 12. Mean temperature differences between the E2O and IIDW and IOK forcing datasets over the whole study period 2009-2012. A – map displaying IIDW dataset minus E2O forcing values. B – map displaying IOK dataset minus E2O forcing values. A negative number thus means that the E2O dataset predicts a warmer mean temperature.

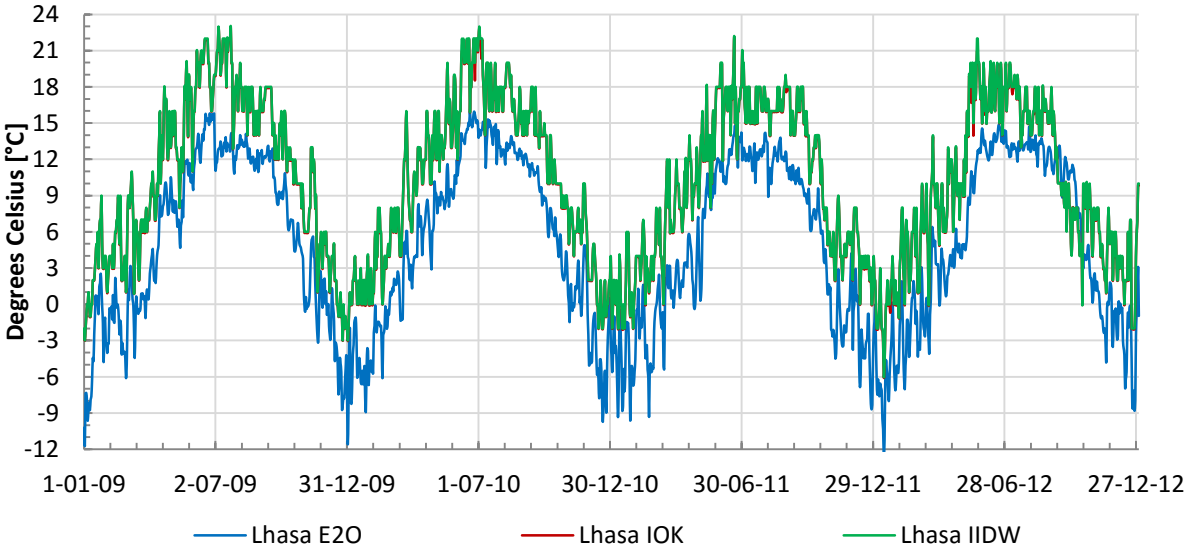


Figure 13. Temperature at Lhasa (29.65,91.14) in the upper part of the basin according to the different datasets. The forcing values differ significantly between the E2O forcing dataset and the interpolated-in-situ dataset.

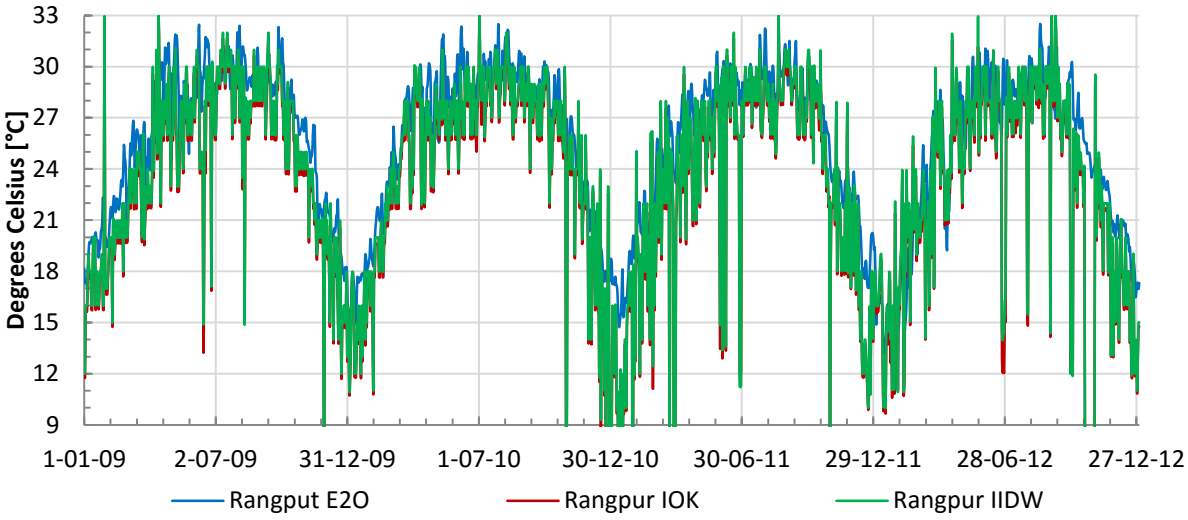


Figure 14. Temperature at Rangpur (25.76,89.28) in the lower part of the basin according to the different datasets. The forcing values are roughly the same.

4.1.2. Precipitation

Figure 15 displays the mean precipitation value per cell over the whole study period. There seems to be large differences in the distribution and magnitude of precipitation with the considered datasets.

The interpolated in-situ forcing datasets and E2O forcing dataset agree that less precipitation occurs in the upper part than in the lower part of the basin. The distribution of precipitation differ between the datasets, notably the location of hotspots of precipitation.

The magnitude of mean precipitation differ as well, as is shown in Figure 16A – B. The ratio of interpolated in-situ forcing data divided by the E2O forcing data shows that the IOK and IIDW mean precipitation values in the upper part of the basin is significantly more than the mean precipitation values of the E2O dataset since the values are 2.0 and higher. In the lower part of the basin, the difference is smaller: here the ratio is mostly in-between 0.7 to 1.1. This is better observed in a scale-adjusted version (Figure 16C – D).

Considering the spatial ratio distribution in the upper part of the basin, the interpolated values cause a higher ratio value when compared with the three knowns points in this section. However, the linear difference between the E2O forcing dataset and the IOK and IIDW datasets (Figure 16E – F) shows that the observed record generally consists of higher values than the E2O forcing data in this region. This is different in the lower part of the basin. Here estimated mean precipitation of the E2O forcing data is generally higher than the IOK and IIDW values.

The temporal pattern of mean precipitation linear difference shows that the difference mostly occurs during the monsoon months (J–J–A–S) (Figure 17). Even though, the ratio of mean precipitation difference in the upper part is generally much higher than in the lower part of the basin, the linear difference between E2O forcing data and the station in Lhasa is smaller than in the lower part of the basin (Rangpur). Considering the general higher ratio in the upper part of the basin, this section experiences a large input difference using E2O forcing data or the interpolated forcing datasets.

This has consequences for the snow accumulation and snowmelt in the basin. As is shown in Figure 18, the pattern of snow cover is considerably different between the datasets. The E2O forcing data results in more snow accumulation than using the IOK and IIDW forcing datasets and this difference is most notable in the North-Western part of the basin. Using the E2O forcing data, average snow cover per cell is up to 50 mm SWE. The SWE of average snow cover predicted by the IOK and IIDW datasets is up to 10 mm. Interestingly, even though the amount of precipitation according to E2O forcing data in the North-Western area is up to 3 times less than the interpolated in-situ datasets estimate, the mean snow cover is up to 5 times more. This can be explained with the fact that predicted temperature by IOK and IIDW in the upper part is higher than zero degrees Celsius earlier in a year (Figure 13), resulting in earlier snowmelt.

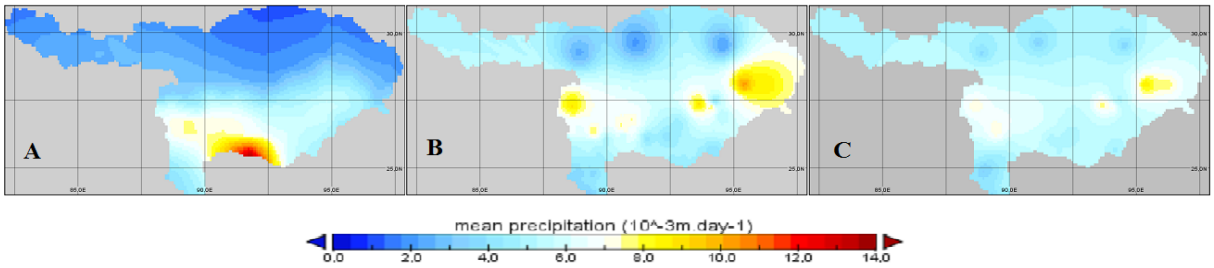


Figure 15. Mean precipitation values (in mm/day) per cell of the whole study period 2009-2012. A – E2O forcing data. B – IIDW forcing data. C – IOK forcing data.

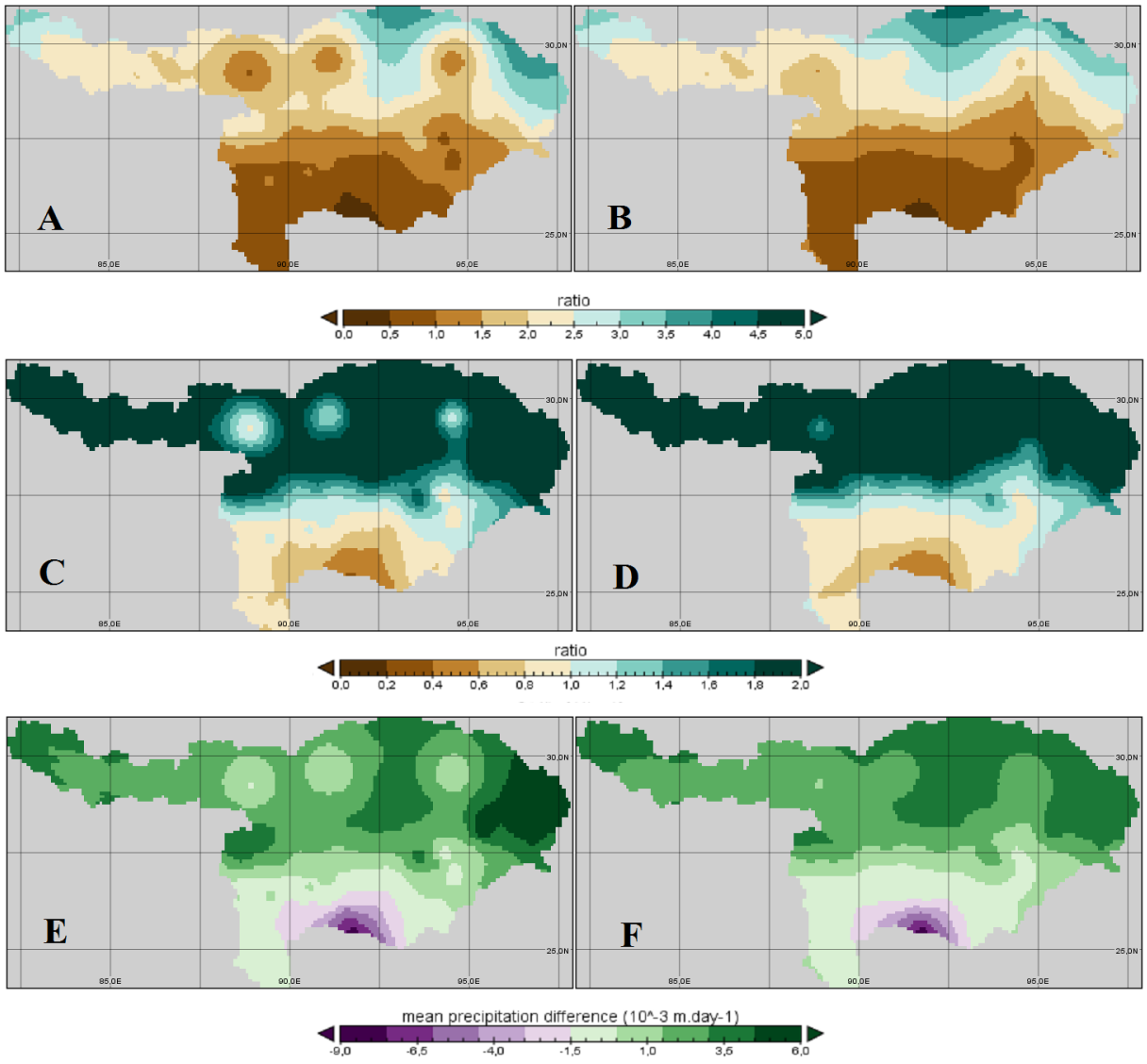


Figure 16. Mean precipitation values (P) for the ratio and linear differences over the whole study period 2009-2012. A and C – mean P of the IIDW forcing dataset divided by E2O forcing dataset with different scales. B and D – mean P of the IOK forcing datasets divided by E2O forcing dataset mean P data with different scales. E – mean P linear difference between IIDW and E2O forcing datasets. F – mean P linear difference between IOK and E2O forcing datasets. A negative value means that E2O dataset estimates a higher P value.

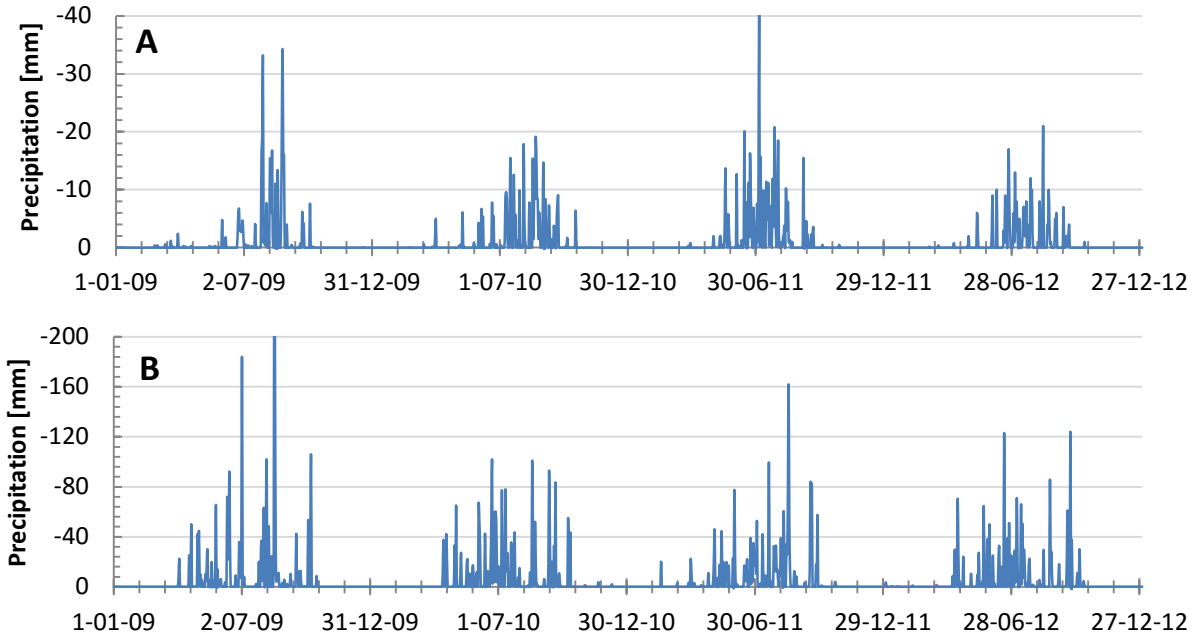


Figure 17. Total daily precipitation difference between E2O forcing data and observed data at Lhasa (A) (upper part of the basin) (29.65, 91.14) and Rangpur (B) (lower part of the basin) (25.76, 89.28) from 2009 to 2012. At Lhasa the difference ranges between 0.03 to -40.4 mm, and at Rangpur the difference ranges from 0.09 to -313 mm. A negative value means that E2O predicts a lower value than was observed. Dates with missing values in the station records have been left out of the analysis.

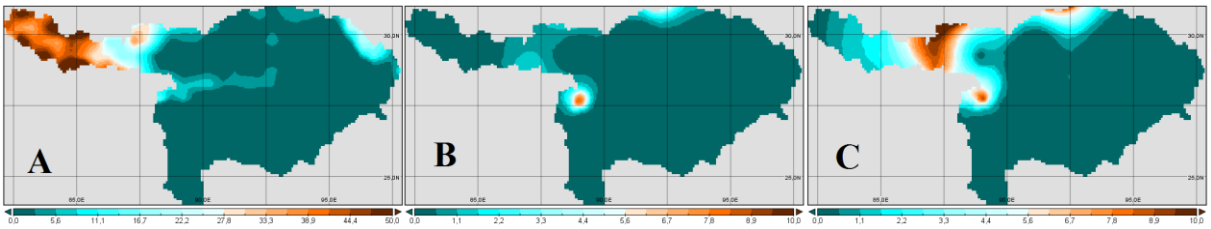


Figure 18. Average snow cover per cell during the study period 2009-2012. A – the snow cover calculated with the E2O dataset. The scale bar is from 0 to 50 mm in snow water equivalent (SWE). B – the snow cover calculated with the IIDW dataset. The scale bar is from 0 to 10 mm SWE. C – the snow cover calculated with the IOK dataset. The scale bar is from 0 to 10 mm SWE for improved visualization.

4.2. Evapotranspiration

The large differences in temperature and precipitation forcing have major consequences on the calculation of the potential and actual evapotranspiration (ET) with PCR-GLOBWB. As can be seen from Figure 19 the potential and actual ET using IOK data is higher than using E2O forcing data. Additionally, the ratio of actual over potential ET is generally larger with IOK than using E2O forcing input. This is caused by the general higher temperature input of the interpolated in-situ datasets.

As shown in Figure 20 and Figure 21, a large distinction can be made between the upper and lower part of the basin. In the lower part of the basin, the estimated potential ET using the E2O forcing data is slightly higher than when using IOK and IIDW data. In the upper part of the basin, the estimated potential ET using the E2O forcing data is substantially lower than using the IOK and IIDW data. Logically, this is similar to the comparison of the temperature using different datasets.

In the upper part of the basin, the potential evaporation was never reached. Meaning that the available precipitation is the limiting factor, otherwise known as water limited. In the lower part of the basin, the potential is reached during the monsoon season (J– J–A–S) and notably not reached in the pre- and post-monsoon season. During the pre- and post-monsoon periods the mean temperature is high while the precipitation is relatively low. During the monsoon season, the ET is energy limited.

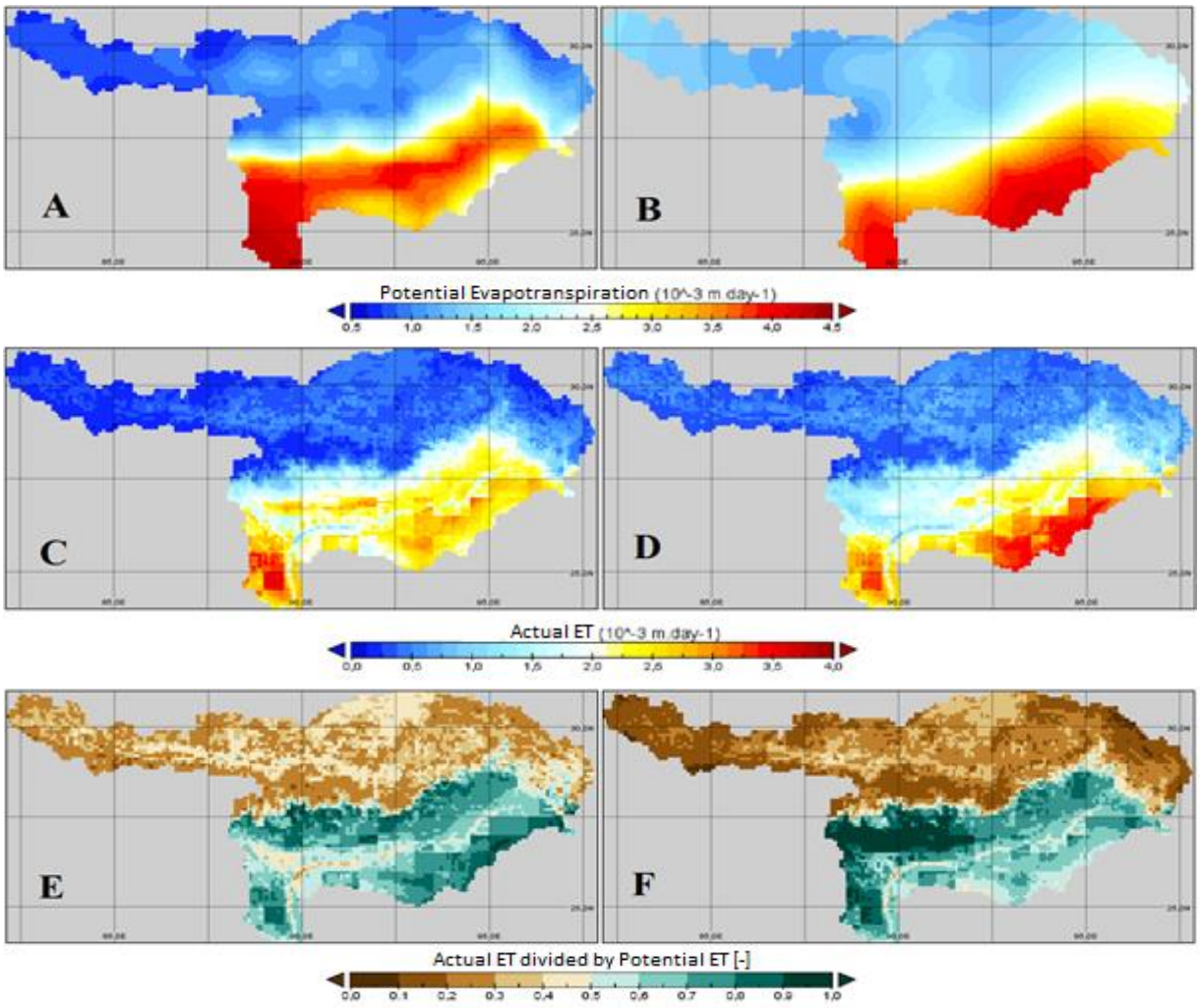


Figure 19. A and B – estimated mean potential ET per cell using E2O (A) and IOK (B) forcing data over the whole study period 2009-2012. C and D – estimated mean actual ET per cell using E2O (C) and IOK (D) forcing data. E and F – ratio of actual to potential ET using E2O (E) and IOK (F) forcing data. Images of the IIDW dataset has been left out due to its similarity to the IOK results.

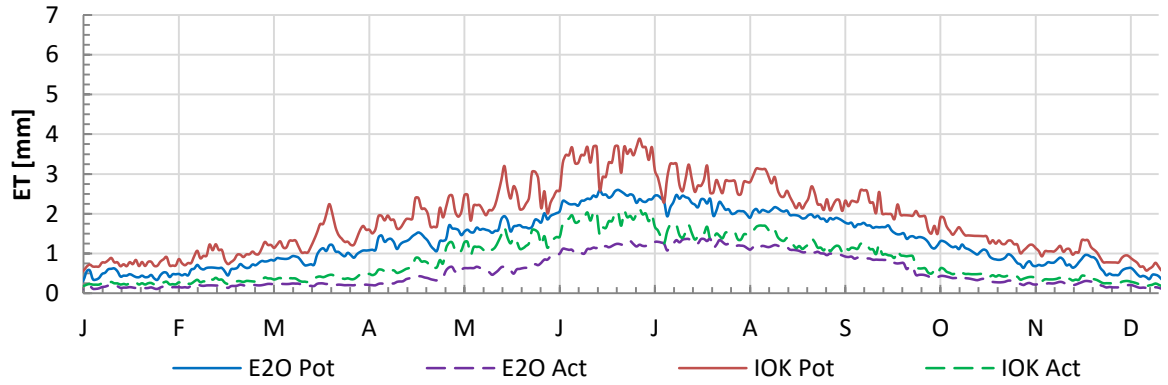


Figure 20. The calculated daily-mean potential and actual ET at Lhasa (29.65, 91.14) during the study period 2009-2012.

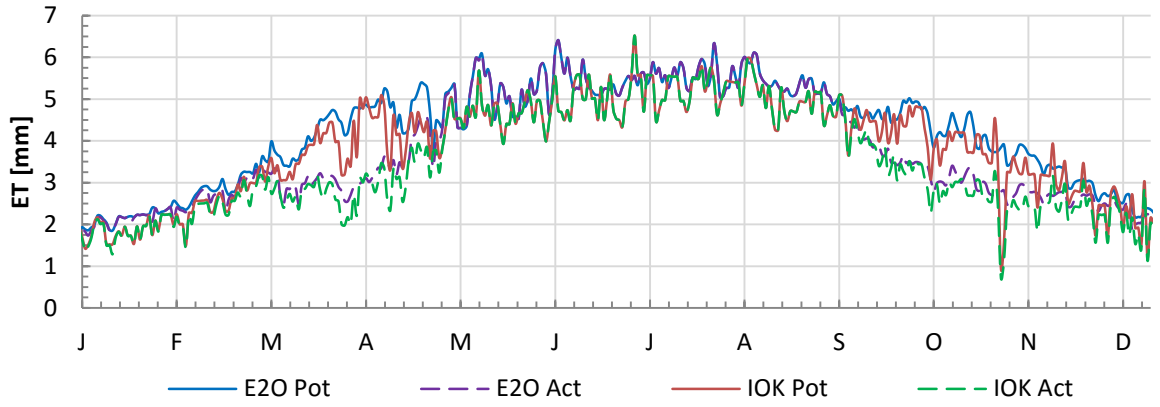


Figure 21. The calculated daily-mean potential and actual ET at Rangpur (B) (25.76, 89.28) during the study period 2009-2012.

4.3. Simulated Discharge

4.3.1. Initial Model Run Discharge Results

Figure 22 displays the daily discharge values at the Bahadurabad gauging station estimated by the PCR-GLOBWB model using the E2O, IOK and IIDW forcing datasets. The daily values are compared to the daily measured values observed at the Bahadurabad station. From these graphs, one can see that the three model runs agree with the observed record during the low flow period in timing and magnitude. Secondly, the estimated values using the E2O forcing data underestimates discharge. The mean discharge using the E2O forcing data is 16,746 m^3/s, while the mean discharge estimated using the IOK and IIDW forcing are 26,573 m^3/s and 26,493 m^3/s, respectively. Additionally, the model runs are more erratic than the observed discharge record. This is observable in the amplitude of the peaks in the hydrographs. However, generally the timing of peaks and lows is similar in the model runs and the observed record. Finally, the estimated discharge using the IOK and IIDW forcing seems to overestimate discharge most notably in the pre-monsoon period.

The mean-monthly hydrograph of the study period 2009-2012 is shown in Figure 23. All model runs show the same trend: low flow in the beginning and end of a year, increasing rise of discharge with peak discharge in July. The mean-monthly curve estimated with the E2O forcing underestimates discharge at all instances. The IOK and IIDW peak discharge is close to the observed mean-monthly flow in July. However, the mean-monthly discharge calculated with the IOK and IIDW forcing data is overestimated in the months April, March and June. There seems to be a substantial 'jump' in discharge from March to April using the IOK and IIDW forcing data. This 'jump' is not present (or slightly present) in the observed record and the model run using E2O forcing data.

The modelled discharge results have been compared on a daily time scale using statistical indicators (Table 3). The NSE of the daily E2O discharge values of 2009 to 2012 is 0.591, the bias is -7615 m^3/s which means that the discharge is underestimated compared to the observed record. For comparison, the NSE of the E2O dataset over the period 1985 to 2012 is 0.669 and the bias -4217 m^3/s. The NSE values of the IOK and IIDW runs are 0.666 and 0.534, respectively. The correlation of the estimated and observed values are similar, however the IOK performs best with 0.876. The MAE and RMSE are smallest with the IOK model results. The mean modelled discharge value calculated with the IOK and IIDW (26,493 m^3/s and 26,573 m^3/s) model results is higher than the mean observed discharge value (23,305 m^3/s). The mean modelled discharge value using E2O model run results (15,746 m^3/s) is lower than the mean observed discharge value. The bias of the IOK and IIDW runs are similar and closer to zero than the E2O model result.

On a mean-monthly time scale, the NSE values of the IOK and IIDW model results are 0.913 and 0.881, which is better than the E2O model result: 0.730. The correlation of the model results with the observed record is high. The MAE, RMSE and bias are closer to zero compared with the daily time scale results and the IOK and IIDW forcing data come are nearer to zero than the E2O model result.

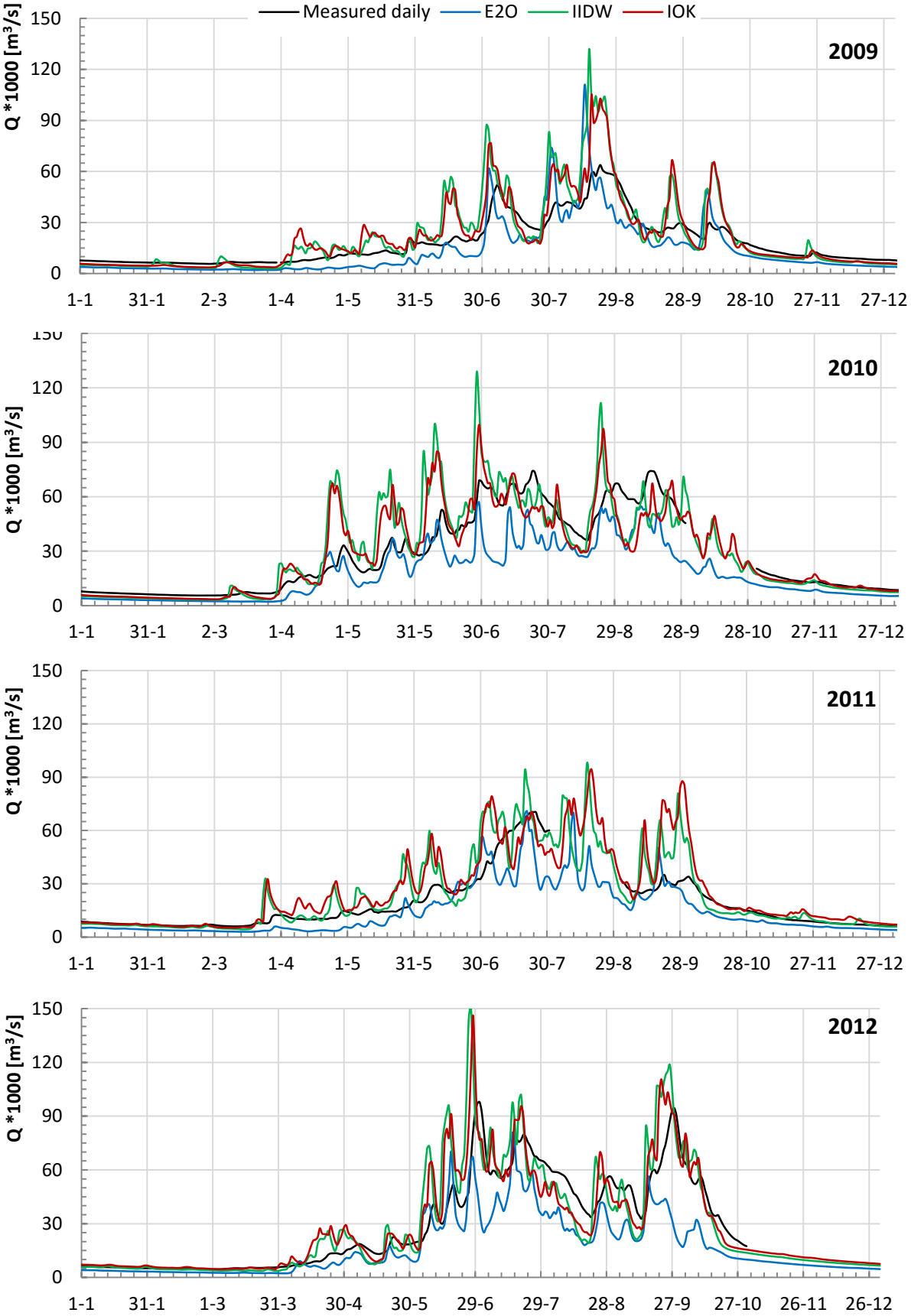


Figure 22. Daily discharge estimates using the E2O, IIDW and IOK forcing data. The discharge values measured at the Bahadurabad gauging station has been put in as reference. The in-situ datasets seem to simulate the observed discharge data better than the E2O dataset.

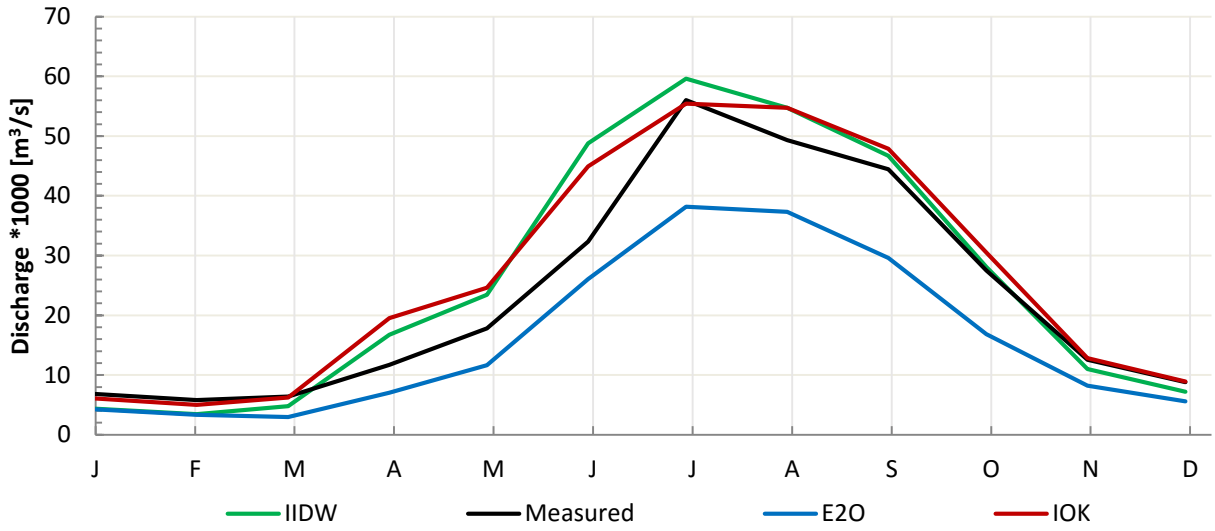


Figure 23. Mean-monthly (of the study period 2009-2012) estimated discharge using E2O, IOK and IIDW forcing data compared with mean-monthly observed data from the Bahadurabad gauging station.

Table 3. Verification metrics, with various performance indicators (PI), of the discharge of the different combined and modified datasets. Nash–Sutcliffe model efficiency coefficient (NSE), Pearson's correlation (r), mean absolute error (MAE), root mean squared error (RMSE), the mean modelled values (MMV), mean observed values (MOV) and the bias, are shown for the indicated forcing datasets for three different time scales.

Time Scale	Performance Indicator (PI)	Forcing Dataset		
		E2O	**IIDW**	**IOK**
Daily	NSE	0.591	0.534	0.666
	r	0.861	0.857	0.876
	MAE	8858	7923	7002
	RMSE	12904	13779	11671
	Mean modelled value	15746	26573	26493
	Mean observed value	23305	23305	23305
	Bias	-7615	3299	2959
Monthly	NSE	0.700	0.810	0.854
	r	0.952	0.943	0.944
	MAE	7579	5372	5016
	RMSE	10436	8303	7287
	Mean modelled value	15671	26463	26385
	Mean observed value	23183	23183	23183
	Bias	-7579	3295	2963
Mean-monthly	NSE	0.730	0.881	0.913
	r	0.993	0.974	0.978
	MAE	7626	4072	3480
	RMSE	9057	6026	5145
	Mean modelled value	15671	26463	26385
	Mean observed value	23297	23297	23297
	Bias	-7626	3166	3089

4.3.2. Combined Dataset Model Runs

The model run using IOK forcing data seems to perform slightly better than the IIDW dataset. On a daily scale, the NSE value is higher, the correlation is higher, the MAE is smaller, the RMSE is smaller, the difference of mean modelled value to mean observed value is slightly smaller, and the bias is smaller. For further comparisons and combinations of datasets, the IOK forcing dataset was used to represent interpolated local forcing data.

The difference between using the Hamon method or the Penman-Monteith approach (P-M) is small (Figure 24). The P-M model runs have a higher mean-monthly discharge. The relative mean-monthly difference ranges from 0.40% in the low flow period up to 3.13% during peak flow. The NSE value of IOK forcing combined with P-M ET calculation is slightly higher, while E2O forcing combined with P-M ET calculation resulted in a slightly lower NSE value. The other performance indicators show the same: a minor improvement for the E2O forcing and slightly worse with the IOK forcing (Table 4).

The results shown in Figure 25 clearly indicate that the precipitation input governs the discharge output. The run with temperature input of the IOK forcing dataset and precipitation input of the E2O forcing dataset (tIOKpE2O) resulted in a discharge curve nearly identical to the discharge curve estimated by using only E2O forcing data. Similarly, the run with temperature input of the E2O forcing dataset and precipitation input of the IOK forcing dataset (tE2OpIOK) is nearly identical to the discharge curve estimated by using only IOK forcing data. According to the statistical indicators, tE2OpIOK performs worse than the IOK dataset, and tIOKpE2O performs worse than the E2O dataset: the bias, MAE and RMSE are larger, and the NSE is smaller (Table 4).

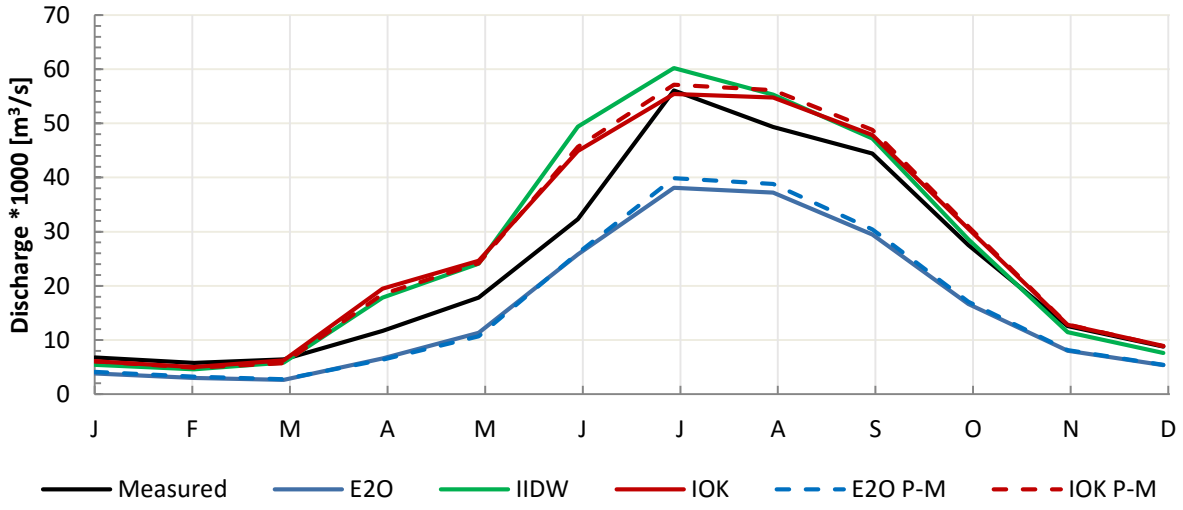


Figure 24. Mean-monthly (of the study period 2009-2012) estimated discharge using E2O, IOK and IIDW forcing data using the Hamon method and Penman-Monteith approach (P-M). This is compared with mean-monthly observed data from the Bahadurabad gauging station.

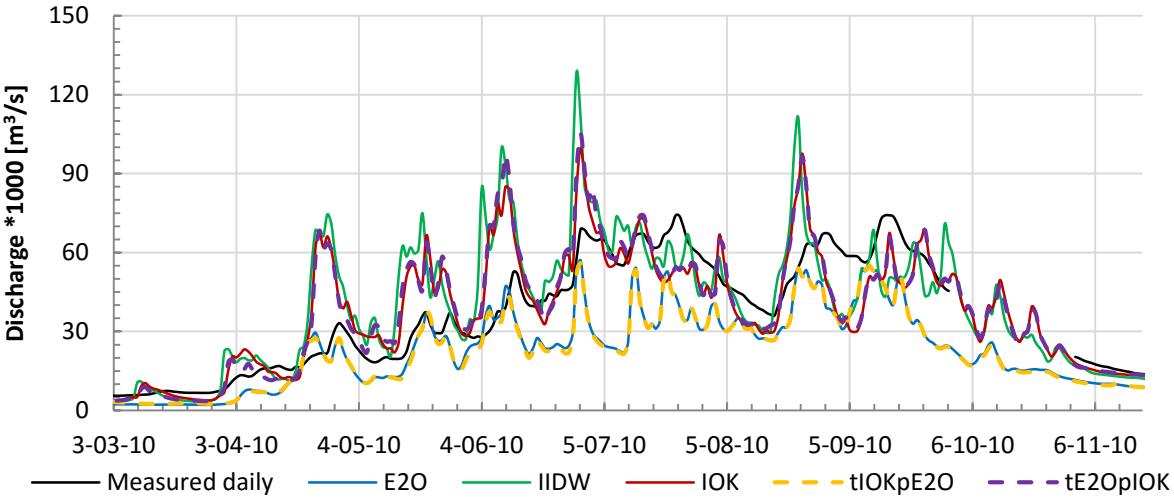


Figure 25. Daily estimated discharge in 2010, using E2O, IOK and IIDW forcing data. As well as combined data input: IOK temperature and E2O precipitation forcing data (tIOKpE2O), and E2O temperature and IOK precipitation forcing data (tE2OpIOK). This is compared with daily observed data from the Bahadurabad gauging station. From this graph one can see that the precipitation input governs the discharge values as opposed to temperature forcing.

Table 4. Verification metrics, with various performance indicators (PI), of the discharge of the different combined and modified datasets. Nash–Sutcliffe model efficiency coefficient (NSE), Pearson's correlation (r), mean absolute error (MAE), root mean squared error (RMSE), the mean modelled values (MMV), mean observed values (MOV) and the bias, are shown for the indicated forcing datasets for three different time scales.

Time Scale	PI	Forcing Dataset							
		E2O	**IOK**	**E2O P-M**	**IOK P-M**	**tIOK-pE2O**	**tE2O-pIOK**	**Bias Corr. +tE2O**	**Bias Corr. +tIOK**
Daily	NSE	0.591	0.666	0.607	0.648	0.578	0.632	0.510	0.548
	r	0.861	0.876	0.858	0.879	0.862	0.875	0.876	0.874
	MAE	8858	7002	8650	7035	9078	7216	7642	7255
	RMSE	12904	11671	12661	11981	13118	12243	14133	13575
	MMV	15746	26493	16156	26756	15382	26919	26925	26851
	MOV	23305	23305	23305	23305	23305	23305	23305	23305
	Bias	-7615	2959	-7206	3211	-7978	3398	3640	3572
Monthly	NSE	0.7	0.854	0.726	0.846	0.681	0.819	0.720	0.766
	r	0.952	0.944	0.951	0.947	0.955	0.939	0.944	0.947
	MAE	7579	5016	7273	5045	7942	5317	6007	5373
	RMSE	10436	7287	9965	7480	10749	8092	10078	9202
	MMV	15671	26385	16077	26645	15305	26806	26785	26716
	MOV	23183	23183	23183	23183	23183	23183	23183	23183
	Bias	-7579	2963	-7172	3210	-7942	3394	3603	3546
Mean-monthly	NSE	0.73	0.913	0.766	0.906	0.707	0.875	0.811	0.854
	r	0.993	0.978	0.993	0.981	0.995	0.972	0.991	0.992
	MAE	7626	3480	7220	3745	7991	3976	4907	4259
	RMSE	9057	5145	8440	5355	9446	6176	7578	6666
	MMV	15671	26385	16077	26645	15305	26806	26785	26716
	MOV	23297	23297	23297	23297	23297	23297	23297	23297
	Bias	-7626	3089	-7220	3348	-7991	3509	3488	3419

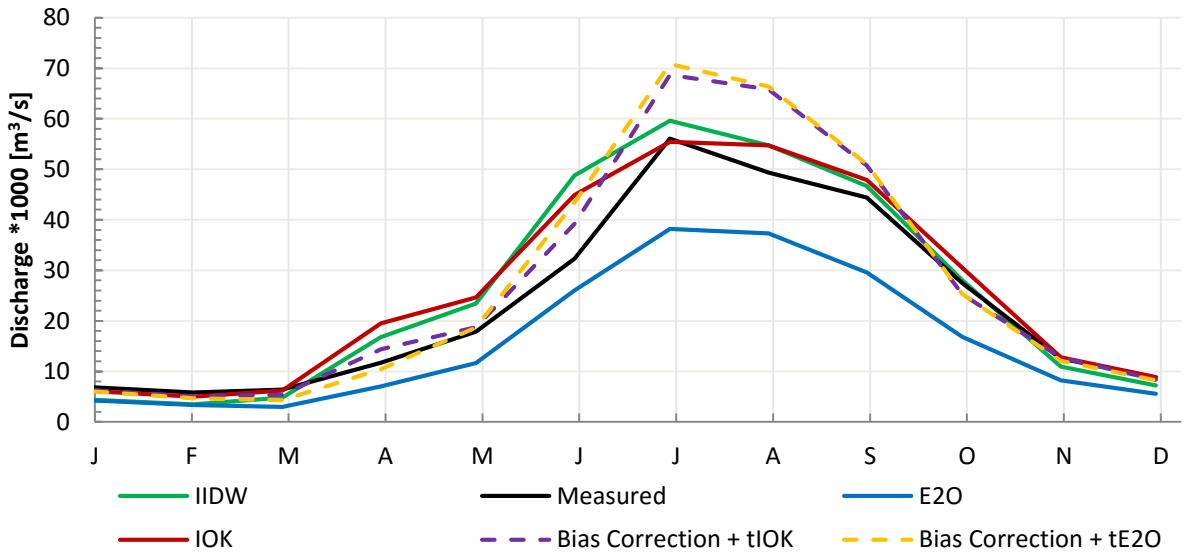


Figure 26. Mean-monthly discharge of the Bias Correction Method compared with model runs using E2O, IOK and IIDW forcing data, and the observed record from the Bahadurabad gauging station. The adjusted precipitation input created by the Bias Correction Method was paired with temperature data from either the IOK dataset (Bias Corr. +tIOK) or the E2O dataset (Bias Corr. +tE2O).

The Bias-Correction Method forcing model runs performed worse than the E2O and IOK model runs (Figure 26). The general mean-monthly discharge trend is present; however, the discharge during the monsoon period is overestimated. The Bias Corr.+tIOK model run estimates discharge in April to be more than calculated with the Bias Corr.+tE2O model run. The precipitation input is the same; therefore, the temperature input difference causes the different mean-monthly discharge values in the months April through July. The performance indicators (Table 4) show that these model runs are worse than the E2O and IOK runs on all time scales: the NSE is worse, the bias and RMSE are larger. Only the MAE is smaller.

The results of the model runs with E2O precipitation input modified with 25% (E2Omod[f1.25]) and 40% (E2Omod[f1.40]) are shown in Figure 27 and Figure 28. The peak flow in July fits better than the Bias-Correction Method and the E2O model run. On a mean-monthly scale, the trend is an improvement compared with the E2O forcing data model run. However, the statistical performance indicates that the NSE is slightly worse than the model run using IOK input (0.648 against 0.666) (Table 5). The bias is better than model runs using E2O and IOK input. The bias of E2Omod[f1.25] and E2Omod[f1.40] are -1935 and 1485, respectively. In addition, the MAE and RMSE are smaller.

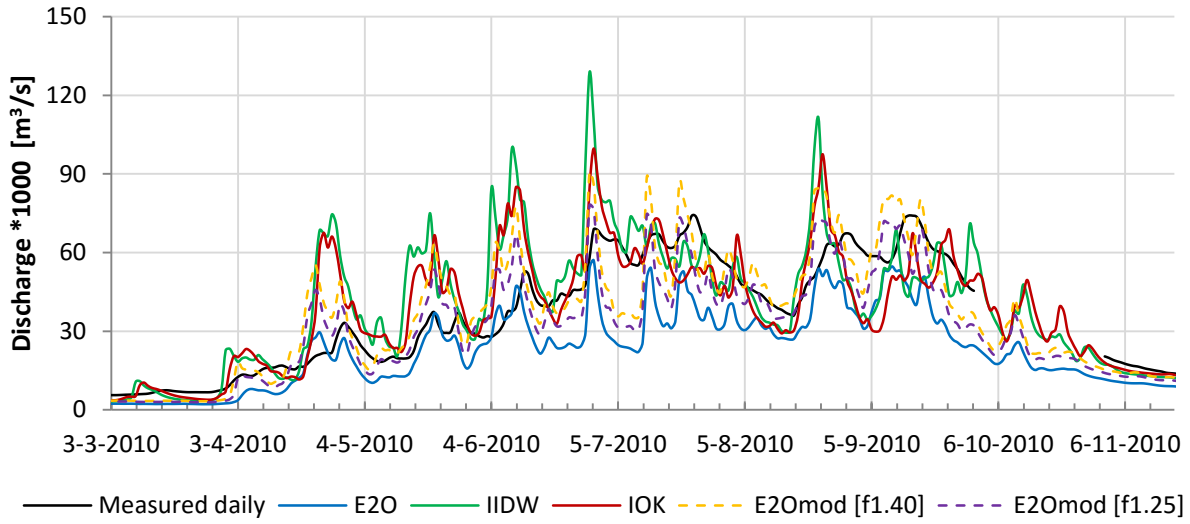


Figure 27. Daily estimated discharge in 2010, using E2O, IOK and IIDW forcing data. As well as modified E2O precipitation forcing data. The E2Omod[f1.25] indicates an increase in E2O precipitation data by 25%. The E2Omod[f1.40] indicates an increase in E2O precipitation data by 40%. For these two model runs the E2O temperature forcing data was used. The model runs are compared with the measured discharge record from the Bahadurabad gauging station.

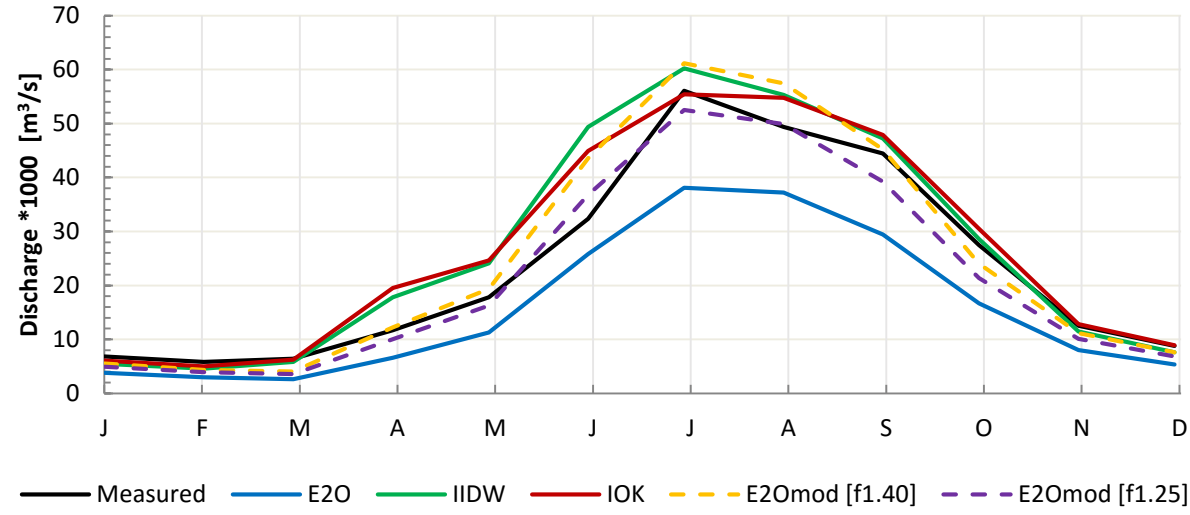


Figure 28. Mean-monthly estimated discharge in 2010, using E2O, IOK and IIDW forcing data. As well as modified E2O precipitation forcing data. The E2Omod[f1.25] indicates an increase in E2O precipitation data by 25%. The E2Omod[f1.40] indicates an increase in E2O precipitation data by 40%. For these two model runs the E2O temperature forcing data was used. The model runs are compared with the measured discharge record from the Bahadurabad gauging station.

The model runs using datasets combinations by adjusting for the physiographic and climatological regions in the basin are shown in Figure 29. The runs using IOK precipitation data in the North and E2O precipitation data in the South (nIOKsE2O+tE2O and nIOKsE2O+tIOK) are similar to the model run using only IOK input data. The runs using E2O precipitation data in the North and IOK precipitation data in the South (nE2OsIOK+tE2O and nE2OsIOK+tIOK) are similar to the model run using only E2O input data. The different temperature input used in these model runs created differences in the months April through July. The nIOKsE2O+tIOK model run performs best in terms of statistical performance (Table 5). The NSE on the daily time scale is best with 0.751. The MAE, RMSE and bias are smallest of all model runs: 6020, 10065 and 1405, respectively. On a monthly scale this dataset is best as well, with a NSE value of 0.890.

Secondly, using IOK or E2O temperature data in the upper part of the basin results in different mean-monthly discharge quantities during the months A–M–J. This coincides with the period in which snowmelt occurs. IOK temperature is generally warmer and thus earlier above zero degrees Celsius, resulting a higher discharge in the early stage of this period.

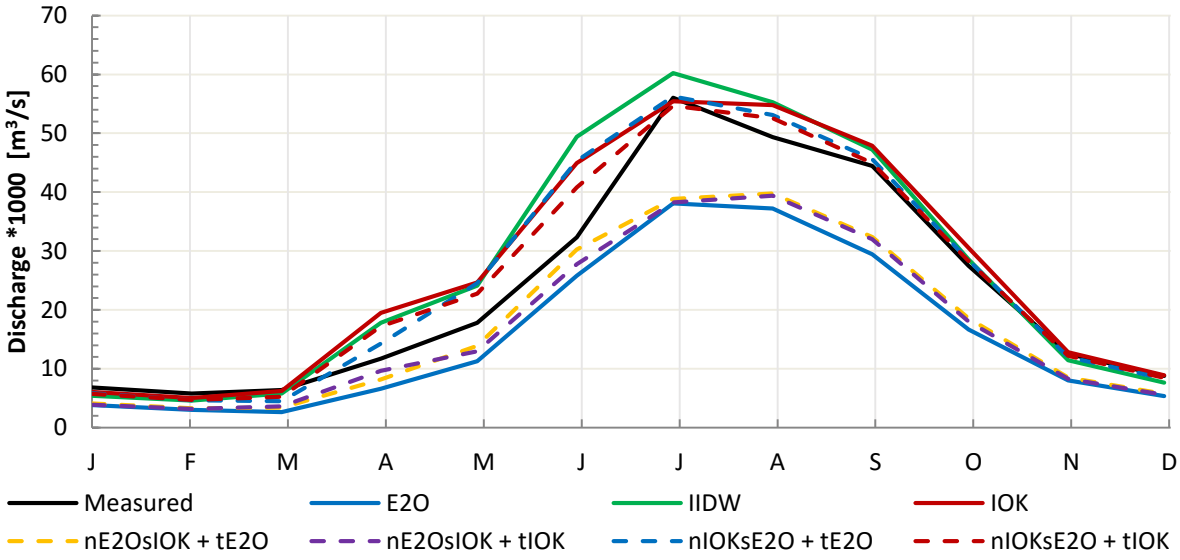


Figure 29. Mean-monthly discharge results of the E2O, IOK and IIDW datasets. As well as the datasets created to adjust for the physiographic and climatological regions in the basin. nE2OsIOK+tE2O – E2O precipitation (P) data in the North and IOK P data in the South with E2O temperature (T) data over the whole basin. nE2OsIOK+tIOK – E2O P data in the North and IOK P data in the South with IOK T data over the whole basin. nIOKsE2O+tE2O – IOK P data in the North and E2O P data in the South with E2O T data over the whole basin. nIOKsE2O+tIOK – IOK P data in the North and E2O P data in the South with IOK T data over the whole basin. These discharge results are compared with the mean-monthly observed record at the Bahadurabad gauging station.

Table 5. Verification metrics, with various performance indicators (PI), of the discharge of the different combined and modified datasets. Nash–Sutcliffe model efficiency coefficient (NSE), Pearson's correlation (r), mean absolute error (MAE), root mean squared error (RMSE), the mean modelled values (MMV), mean observed values (MOV) and the bias, are shown for the indicated forcing datasets for three different time scales.

Time Scale	PI	Forcing Datasets							
		E2O	**IOK**	**E2Omod f1.40**	**E2Omod f1.25**	**nE2OsIOK +tIOK**	**nE2OsIOK +tE2O**	**nIOKsE2O +tIOK**	**nIOKsE2O +tE2O**
Daily	NSE	0.591	0.666	0.531	0.648	0.665	0.674	0.751	0.722
	r	0.861	0.876	0.835	0.840	0.877	0.876	0.890	0.886
	MAE	8858	7002	7345	6967	7981	7854	6020	6339
	RMSE	12904	11671	13820	11986	11687	11522	10065	10637
	MMV	15746	26493	24755	21378	16969	17320	24906	25344
	MOV	23305	23305	23305	23305	23305	23305	23305	23305
	Bias	-7615	2959	1485	-1935	-6434	-6084	1405	1865
Monthly	NSE	0.700	0.854	0.828	0.877	0.759	0.772	0.890	0.865
	r	0.952	0.944	0.939	0.943	0.957	0.952	0.949	0.944
	MAE	7579	5016	5303	4697	6476	6288	4548	4814
	RMSE	10436	7287	7897	6681	9352	9086	6329	6989
	MMV	15671	26385	24640	21277	16893	17243	24798	25233
	MOV	23183	23183	23183	23183	23183	23183	23183	23183
	Bias	-7579	2963	1475	-1927	-6397	-6049	1408	1863
Mean-monthly	NSE	0.730	0.913	0.933	0.965	0.792	0.810	0.961	0.933
	r	0.993	0.978	0.987	0.989	0.989	0.986	0.986	0.979
	MAE	7626	3480	3198	2853	6404	6054	2400	2780
	RMSE	9057	5145	4528	3269	7957	7600	3462	4508
	MMV	15671	26385	24640	21277	16893	17243	24798	25233
	MOV	23297	23297	23297	23297	23297	23297	23297	23297
	Bias	-7626	3089	1343	-2020	-6404	-6054	1502	1937

5. Discussion and Conclusions

The current analysis has focussed on discharge results of the uncalibrated PCR-GLOBWB model in the Brahmaputra River Basin using the WFDEI (E2O) global forcing dataset of the *eartH2Observe* Project.

Due to the fact that measured meteorological data is not readily available in this part of the world, an assumption was made considering the gathered local data. Data received from the Institute for Water Modelling (IWM, personal communication, 2015) were scientific and reliable. Other local station data was downloaded from a meteorological website (Weather Underground, 2015), which clearly state that the data is not scientific. Nevertheless, this data was applied. To compensate for this statement, the most complete records from the best locations (airports and governmental buildings) were chosen. The station data is not well distributed and especially in the upper part of the basin, the Himalayan Plateau, reliable station data is lacking. The shortage of stations is evident from Figure 6. Local station data was interpolated using three techniques, which resulted in three datasets distinguished as In-situ OK (IOK), In-situ Spline (ISpline), and In-situ IDW (IIDW). The ISpline dataset was found to be unfit for this study.

The interpolation techniques were applied neglecting elevation considerations. However, this basin is dominated by two distinct regions based on elevation. From the temperature and precipitation distribution one can observe the influence of elevation on meteorological phenomena in this basin. Select interpolation techniques were chosen to reflect this, however future projects may consider including elevation difference more completely to account for its influence.

Temperature forcing data analysis showed that the basin mean temperature difference between the interpolated local data and the global reanalysis data is approximately 3.17 ± 0.70°C to 3.33 ± 0.75°C. This difference is not spatially uniform (Figures 11 through 14): in the upper part of the basin the interpolated local mean temperature data is warmer, while in the lower part it is slightly colder. The difference is amplified during the monsoon period. Considering the spatial distribution, a distinction can be made between the upper and lower part of the basin.

Likewise, precipitation distribution can be considered as two distinct regions as well. The interpolated local data and global reanalysis data agree that less rain precipitates in the upper part compared to the lower part of the basin. However, quantities differ substantially. According to the interpolated local data, mean precipitation in the upper part is more than two times the rainfall in the global reanalysis data. This is evident from the quantities in the station records. These are consistently higher than the global reanalysis data (Figure 16), and especially during the monsoon (Figure 17). In the lower part of the basin the ratio is mostly around 1, however the absolute quantities associated with the small ratio difference is considerably higher.

Evapotranspiration (ET) analysis emphasises the distinction between the upper and lower part. The upper part of the basin is water limited while in the lower part the system is energy limited during the monsoon period (Figures 20 and 21). The interpolated local temperature data is generally warmer thus the ratio actual to potential ET is higher than is the case with global reanalysis data (Figure19).

The differences in precipitation and temperature has influence on the snow cover and snowmelt. Even though the amount of precipitation according to the global reanalysis forcing data in the North-Western area is up to 3 times less than the interpolated local forcing estimates, the mean snow cover is up to 5 times more. This is caused by the higher estimated temperatures of the interpolated local forcing data (Figure 18).

In total 15 model runs were considered. The model runs using local forcing data interpolated with Ordinary Kriging (IOK) estimates discharge better than the model run using WFDEI global reanalysis data (E2O): on a daily time scale, the bias, mean absolute error (MAE) and root-mean-squared error (RMSE) are smaller and the Nash–Sutcliffe model efficiency coefficient (NSE) is higher. The results of the interpolated local forcing data model runs (IOK and IIDW) agree with earlier studies: average discharge of the Brahmaputra River is measured to be approximately 20,000 m^3/s (Datta & Singh, 2004; Immerzeel, 2008) and is estimated slightly higher. Simulated discharge using E2O forcing data is underestimated significantly: approximately 15,746 m^3/s. Nevertheless, these three runs when compared with local observed discharge at the Bahadurabad gauging station capture the trend and timing of peaks similarly. The mean monthly discharge is highest in July and lowest in February. E2O forcing data obviously lack in precipitation quantity.

Using the Hamon or Penman-Monteith approach to calculate ET did not result in large differences (Figure 24). The combinations of forcing data showed that precipitation input governs discharge output (Figure 25).

The Bias-Correction Method did not perform as expected (Figure 26). The expectation was that the mean-monthly curve would end up near the observed discharge. During the monsoon period (months J–J–A–S) the resulting discharge and therefore precipitation is too much. The author attributes this to the fact that the method was applied to an already interpolated field. Vila et al. (2009) applied this method on station data before interpolating (Figure 26).

Applying a fixed percentage to the E2O precipitation data resulted in an improved mean-monthly discharge curve. However, the timing of peak flow on a daily scale is not changed and the IOK and IIDW forcing dataset capture this better. This is reflected by a reduced correlation between the observed and estimated discharge. The NSE, bias, MAE and RMSE were better the model run using unmodified E2O forcing data (Figures 27 and 28). This emphasises that the input precipitation provided by the E2O forcing data is lacking, and the need to adjust for this deficit.

The best forcing dataset in terms of verification metrics and mean-monthly discharge characteristics, is the run where interpolated local precipitation input was used in the upper part of the basin and global reanalysis precipitation data in the lower part, and interpolated local temperature data over the whole basin (nIOKsE2O+tIOK). The NSE on a daily scale is highest with 0.751. The bias, MAE and RMSE are smallest with 1405 m^3/s, 6020 m^3/s and 10,065 m^3/s. The correlation is highest as well: 0.890. On a monthly scale this combined dataset performs best as well (Table 5).

From the four combinations in Figure 29, it can be deduced that the precipitation input in the upper part of the basin predominately determines discharge output at the Bahadurabad gauging station. To elaborate: nE2OsIOK+tE2O and nE2OsIOK+tIOK forcing resulted in a mean-monthly discharge curve very similar to the E2O result and a verification metrics profile similar to the IOK run, but with a substantially larger bias also found in the E2O result (i). The nIOKsE2O+tE2O and nIOKsE2O+tIOK forcing data model run results are more similar to the IOK model run (ii). The absolute precipitation differences between IOK and E2O is largest in the lower part of the basin, however the ratio is much larger in the upper part of the basin (iii).

Thus, the IOK precipitation input in the upper part of the basin causes an improved discharge estimation as opposed to using E2O forcing data. To emphasise the consequences of this finding, the precipitation in the upper part is based on only three station records, which indeed have higher precipitation values than the global reanalysis data indicate at those locations, however the interpolated values in the surrounding area are estimated to be higher, which may or may not be the case. These estimated values are higher due to the fact that other stations, which are located in the lower part of the basin, have higher values and are taken into account during the interpolation process.

This is similarly the case with interpolating the temperature. The temperature gradient over the basin with the E2O forcing dataset is much stronger than with the IOK and IIDW datasets (Figure 10). Thus, the temperature input using IOK and IIDW is most likely too warm. This is also evident from the snow cover built up differences (Figure 18). Additionally, the relative ‘jump’ in discharge in the month March is not present in the observed record. It is present in all cases when using interpolated local temperature data in the northern part of the basin, and it is not present when using global reanalysis data in the northern part,

Thus, the best data combination would be: (i) interpolated local data in the upper part of the basin to appreciate the precipitation deficit in this section, (ii) global reanalysis precipitation data in the lower part of the basin since this is similar to the interpolated local data input, and (iii) global reanalysis temperature data in the whole basin since that does more right to snow cover and melt timing. This is the nIOKsE2O+tE2O combined dataset.

The general conclusion is that the WFDEI forcing dataset applied to the PCR-GLOBWB model underestimates discharge in the Brahmaputra Basin. This is caused by the underestimation of input precipitation, especially during the monsoon (months J–J–A–S). The WFDEI temperature input was apparently sufficient. To improve water resource modelling the WFDEI forcing dataset needs to be combined with local station data. Combined datasets with local precipitation data in the upper part and global WFDEI precipitation data in the lower part of the basin produced the best discharge results on daily, monthly and mean-monthly time scales as well as peak and low flow estimation. Nevertheless, the verification metrics of the WFDEI forcing data model run proves that the WFDEI forcing is not unusable when taken into account the precipitation deficit.

For further study in this basin, it is recommended to incorporate elevation in the interpolation method for local data. Secondly, a similar study for comparing local forcing data with global reanalysis

forcing data in a different basin can benefit the evaluation of the WFDEI forcing dataset. Continued regional basin-wide cooperation in the Brahmaputra River Basin in terms of data sharing and availability would improve system modelling. Additionally, more meteorological stations on the Himalayan Plateau and in the Himalayan Mountains can significantly advance the usage of local and satellite data inter-comparison.

Acknowledgements

I am grateful to Abu Saleh Khan for providing excellent accommodation near the IWM complex during my stay in Dhaka, Bangladesh. Furthermore, I would like to thank the people of the Flood Management Group at IWM for discussing interpretations, helping with gathering local station data and showing me around in Bangladesh. I would like to express my sincere gratitude to the Department of Physical Geography of Utrecht University for the grant that allowed me to visit Bangladesh. Additionally, I would like to thank my supervisors dr. ir. Geert Sterk and ir. Patricia Lopez Lopez for the excellent guidance, suggestions, feedback and support during these months.

References

Adler, R. F., Huffman, G. J., Bolvin, D. T., Curtis, S., & Nelkin, E. J. (2000). Tropical rainfall distributions determined using TRMM combined with other satellite and rain gauge information. *Journal of Applied Meteorology*, 39(12), 2007-2023.

Ahmad, Q. K., & Ahmed, A. U. (2003). Regional cooperation in flood management in the Ganges-Brahmaputra-Meghna region: Bangladesh perspective. In *Flood Problem and Management in South Asia* (pp. 181-198). Springer Netherlands

Ahmed, A. U., & Mirza, M. M. Q. (2000). Review of Causes and Dimensions of Floods with Particular Reference to Flood '98: National Perspectives. In Ahmad, Q. K., Chowdhury, M. R., Azad, A. K., Imam, S. H., & Sarker, M. (Eds.), *Perspectives on Flood 1998* (67-84), University Press Limited, Dhaka.

Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). Crop evapotranspiration, FAO Irrigation and drainage paper 56, FAO, Rome.

Beniston, M. (2003). Climatic Change in mountain regions: a review of possible impacts. *Climatic Change*, 59, 5-31.

Bergström, S. (1995). The HBV Model. In Singh, V. P. (Ed), *Computer models of watershed hydrology* (443-476). Colorado, Highlands Ranch, Water Resources Publications.

Bierkens, M. F. P., & Van Beek, L. P. H. (2009). Seasonal predictability of European discharge: NAO and hydrological response time. *Journal of Hydrometeorology*, 10(4), 953-968.

BWDB (2015). Bangladesh Water and Power Development Board, Flood Forecasting and Warning Centre, Bangladesh. http://www.ffwc.gov.bd/. Last accessed November 2015.

Childs, C. (2004). *Interpolating surfaces in ArcGIS spatial analyst*. ArcUser, July-September, 3235.

Chow, V. T., Maidment, D. R., & Mays, L. W. (1988). *Applied Hydrology*, McGraw-Hill, New York.

Chowdhury, M. R. (2003). The El Nino-Southern Oscillation (ENSO) and seasonal flooding–Bangladesh. *Theoretical and applied climatology*, *76*(1-2), 105-124.

Datta, B., & Singh, V. P. (2004). Hydrology; In V. P. Singh, N. Sharma, C. Shekhar & P. Ojha (Eds.), *The Brahmaputra Basin Water Resources* (pp. 139–195). The Netherlands: Kluwer Academic Publishers

Davis, J. C., & Sampson, R. J. (1986). *Statistics and data analysis in geology*. New York Wiley.

De Graaf, I. E. M., van Beek, L. P. H., Wada, Y., & Bierkens, M. F. P. (2014). Dynamic attribution of global water demand to surface water and groundwater resources: effects of abstractions and return flows on river discharges. *Advances in Water Resources*, 64, 21-33.

FAO (1998). *Digital Soil map of the World*, Food and Agriculture Organization of the United Nations (FAO), Rome, Italy.

FAO, Aquastat. (2011). Ganges-Brahmaputra-Meghna Basin. Retrieved from http://www.fao.org/nr/water/aquastat/basins/gbm/index.stm. Accessed on 10/09/2015.

Gain, A. K., Immerzeel, W. W., Sperna Weiland, F. C., & Bierkens, M. F. P. (2011). Impact of climate change on the stream flow of the lower Brahmaputra: trends in high and low flows based on discharge-weighted ensemble modelling. *Hydrology and Earth System Sciences*, 15(5), 1537-1545.

Gain, A. K., & Giupponi, C. (2015). A dynamic assessment of water scarcity risk in the Lower Brahmaputra River Basin: An integrated approach. *Ecological Indicators*, *48*, 120-131.

Gain, A. K., & Wada, Y. (2014), Assessment of future water scarcity at different spatial and temporal scales of the Brahmaputra River Bain. *Water Resource Management*, 28, 999-1012.

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 Science*, 121(3), 637-657.

Gleeson, T., Wada, Y., Bierkens, M. F. P., & van Beek, L. P. H. (2012). Water balance of global aquifers revealed by groundwater footprint. *Nature*, 488(7410), 197-200.

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

Haddeland, I., et al. (2011), Multimodel estimate of global terrestrial water balance: Setup and first results, J. Hydrometeorology., 12, 869–884.

Harding, R. J. and Warnaars, T. A., 2011 Water and global change: The WATCH Project Outreach Report. Centre for Ecology and Hydrology, Wallingford, 40pp.

Harding, R., et al. (2011), WATCH: Current knowledge of the Terrestrial global water cycle, J. Hydrometeorology., 12, 1149–1156

Hagemann, S. (2002). *An improved land surface parameter dataset for global and regional climate models*, Max-Planck-Institute for Meteorology, Report 336, Hamburg.

Hagemann, S., & Gates, L. D. (2003). Improving a subgrid runoff parameterization scheme for climate models by the use of high-resolution data derived from satellite observations. *Climate Dynamics* 21, 349-359.

Hamon, R. W., Weiss, L. L., & Wilson, W. T. (1954). Insolation as an empirical function of daily sunshine duration. *Monthly Weather Review*, 82(6), 141-146.

Helsel, D. R., & Hirsch, R. M. (2002). *Statistical methods in water resources*. Reston, VA: US Geological survey.

Huffman, G. J., Adler, R. F., Arkin, P., Chang, A., Ferraro, R., Gruber, A., ... & Schneider, U. (1997).

The global precipitation climatology project (GPCP) combined precipitation dataset. *Bulletin of the American Meteorological Society*, 78(1), 5-20.

Immerzeel, W. W., Droogers, P., De Jong, S. M., & Bierkens, M. F. P. (2009). Large-scale monitoring of snow cover and runoff simulation in Himalayan river basins using remote sensing. *Remote sensing of Environment*, 113(1), 40-49.

Immerzeel, W. W., Van Beek, L. P. H., & Bierkens, M. F. P. (2010). Climate change will affect the Asian water towers. *Science*, 328, 1382-1385.

Immerzeel, W. (2008). Historical trends and future predictions of climate variability in the Brahmaputra basin. *International Journal of Climatology*, *28*(2), 243-254.

Immerzeel, W. W., & Bierkens, M. F. P. (2012). Asia's water balance. *Nature Geoscience*, *5*(12), 841-8.

IWM, personal communication at the Flood Management Group in September and October 2015.

Karmaker, T., & Dutta, S. (2010). Generation of synthetic seasonal hydrographs for a large river basin. *Journal of Hydrology*, *381*(3), 287-296.

Liang, X., Lettenmaier, D. P., Wood, E. F., & Burges, S. J. (1994). A simple hydrologically based model of land surface water and energy fluxes for general circulation models. *Journal of Geophysical Research*, 99, 14415-14428.

Milly, P. C. D., Dunne, K. A. & Vecchia, A. V. (2005). Global pattern of trends in streamflow and water availability in a changing climate. *Nature*, 438, 347-350.

Mirza, M. M. Q. (2003). Three Recent Extreme Floods in Bangladesh: A Hydro-meteorological Analysis. *Natural Hazards*, 28(1), 35-64.

Mirza, M. M. Q., Warrick, R. A., & Ericksen, N. J. (2003). The implications of climate change on floods of the Ganges, Brahmaputra and Meghna rivers in Bangladesh. *Climatic Change*, *57*(3), 287-318.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, 50(3), 885-900.

Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I—A discussion of principles. *Journal of hydrology*, 10(3), 282-290.

Nepal, S., & Shrestha, A. B. (2015). Impact of climate change on the hydrological regime of the Indus, Ganges and Brahmaputra river basins: a review of the literature. *International Journal of Water Resources Development*, *31*(2), 201-218.

Postel, S. L., Daily, G. C., & Ehrlich, P. R. (1996). Human appropriation of renewable fresh water. *Science*, 271, 785-788.

Sarker, M. H., Huque, I., Alam, M., & Koudstaal, R. (2003). Rivers, chars and char dwellers of Bangladesh. *International Journal of River Basin Management*, 1:1, 61-80.

Sarma, J. N. (2005). Fluvial process and morphology of the Brahmaputra River in Assam, India. *Geomorphology*, 70(3), 226-256.

Sloan, P. G., & Moore, I. D. (1984). Modeling subsurface stormflow on steeply sloping forested watersheds, *Water Resources Research*, 20, 1815-1822.

Sperna Weiland, F. C., Van Beek, L. P. H., Kwadijk, J. C. J., & Bierkens, M. F. P. (2010). The ability of a GCM-forced hydrological model to reproduce global discharge variability. *Hydrology and Earth System Sciences*, 14, 1595-1621.

Sperna Weiland, F. C., Van Beek, L. P. H., Kwadijk, J. C. J., & Bierkens, M. F. P. (2012). Global patterns of change in discharge regimes for 2100. *Hydrology and Earth System Sciences*, 16(4), 1047.

Sutanudjaja, E. H., Van Beek, L. P. H., De Jong, S. M., Van Geer, F. C., & Bierkens, M. F. P. (2011). Large-scale groundwater modelling using global datasets: a test case for the Rhine-Meuse basin. *Hydrology and Earth System Sciences*, 15, 2913-2935.

Taylor, R. (2009). Rethinking water scarcity: the role of storage. *Eos, Transactions American Geophysical Union*, 90(28), 237-238.

Uppala, S. M., et al. (2005), The ERA-40 re-analysis, Q. J. R. Meteorol. Soc., 131, 2961–3012.

USGS EROS Data Center (2002), *Global land cover characteristics data base version 2.0*, http://edcdaac.usgs.gov/glcc/globedoc2_0.html.

USGS EROS Data Center (2006). *Hydro1K Elevation Derivative Database*, LP DAAC: http://edcdaac.usgs.gov/gtopo30/hydro/.

Van Beek, L. P. H., Wada, Y., Bierkens, M. F. P. (2011). Global monthly water stress: 1. Water balance and water availability. *Water Resources* Research, 47, W07517.

Van Beek, L. P. H. (2008). Forcing PCR-GLOBWB with CRU meteorological data, report, Utrecht University, Utrecht, The Netherlands, http://vanbeek.geo.uu.nl/suppinfo/vanbeek2008.pfd. Accessed on 13/09/2015.

Van Beek, L. P. H., & Bierkens, M. F. P. (2009). *The Global Hydrological Model PCR-GLOBWB: Conceptualization, Parameterization and Verification*, Report Department of Physical Geography, Utrecht University, Utrecht, The Netherlands.

Vila, D. A., deGoncalves, L. G. G., Toll, D. L., & Rozante, J. R. (2009). Statistical evaluation of combined daily gauge observations and rainfall satellite estimates over continental South America. *Journal of Hydrometeorology*, 10, 533-543

WAPCOS (Water and Power Consultancy Services [India]) (1993). *Morphological Studies of the River*

Brahmaputra. North Eastern Council. Govt. of India, New Delhi, pp. I-1–IX-131.

Wada, Y., van Beek, L. P. H., & Bierkens, M. F. P. (2012). Nonsustainable groundwater sustaining irrigation: A global assessment. *Water Resources Research*, 48(6).

Wada, Y., van Beek, L. P. H., van Kempen, C. M., Reckman, J. W., Vasak, S., & Bierkens, M. F. P. (2010). Global depletion of groundwater resources. *Geophysical Research Letters*, 37(20).

Wada, Y., van Beek, L. P. H., Viviroli, D., Dürr, H. H., Weingartner, R., & Bierkens, M. F. P. (2011). Global monthly water stress: 2. Water demand and severity of water stress. *Water Resources Research*, 47, W07518.

Wada, Y., Wisser, D., & Bierkens, M. F. P. (2014). Global modeling of withdrawal, allocation and consumptive use of surface water and groundwater resources. *Earth System Dynamics Discussions*, 5(1), 15-40.

Wada, Y., Wisser, D., Eisner, S., Flörke, M., Gerten, D., Haddeland, I., ... & Schewe, J. (2013). Multi-model projections and uncertainties of irrigation water demand under climate change. *Geophysical research letters*, 40(17), 4626-4632.

Weather Underground (2015). https://www.wunderground.com/. Last accessed October 2015.

Weedon, G. P., Balsamo, G., Bellouin, N., Gomes, S., Best, M. J., & Viterbo, P. (2014). The WFDEI meteorological forcing data set: WATCH Forcing Data methodology applied to ERA- Interim reanalysis data. *Water Resources Research*, 50(9), 7505-7514.

Wesseling, C. G., Karssenberg, D. J., Burrough, P. A., & Deursen, W. (1996). Integrating dynamic environmental models in GIS: the development of a Dynamic Modelling language. *Transactions in GIS*, 1(1), 40-48.

Appendices

Appendix A: Creation of the Spatial Extent using PCRaster

A spatial extent which encompasses the Brahmaputra Basin is created using PCRaster functions. To accomplish this, first a catchment map was created with a global coverage. This was done using the 5 arc minute global LDD map. Subsequently, the appropriate catchment was identified. Due to the inherent process of the *catchment* function, the Brahmaputra River Basin was not separated from the Ganges River Basin using the global projection. The whole landmask of the Ganges-Brahmaputra-Meghna (GBM) Basin is shown in Figure A. A smaller projection was used using the *resample* and *map attribute* functions. The alignment for this procedure was the location of the Bahadurabad station (25°18'N, 89°66'E) (the crosshairs in show this location) and Figure 1. Then the LDD of this clone map was determined using the *lddmask* function in order to use the *catchment* function again. The correct catchment was identified and a landmask was created. The associated coordinates of the clone map are (82.0, 32.0) and (98.0, 24.0) for the left upper corner and lower right corner respectively (both in decimal notation). Figure compares the Brahmaputra River basin area with the entire GBM basin. The total Brahmaputra basin area was determined to be 520,763 km^2: a deviation of 1.7% compared to the basin area extent mentioned by Immerzeel (2008).

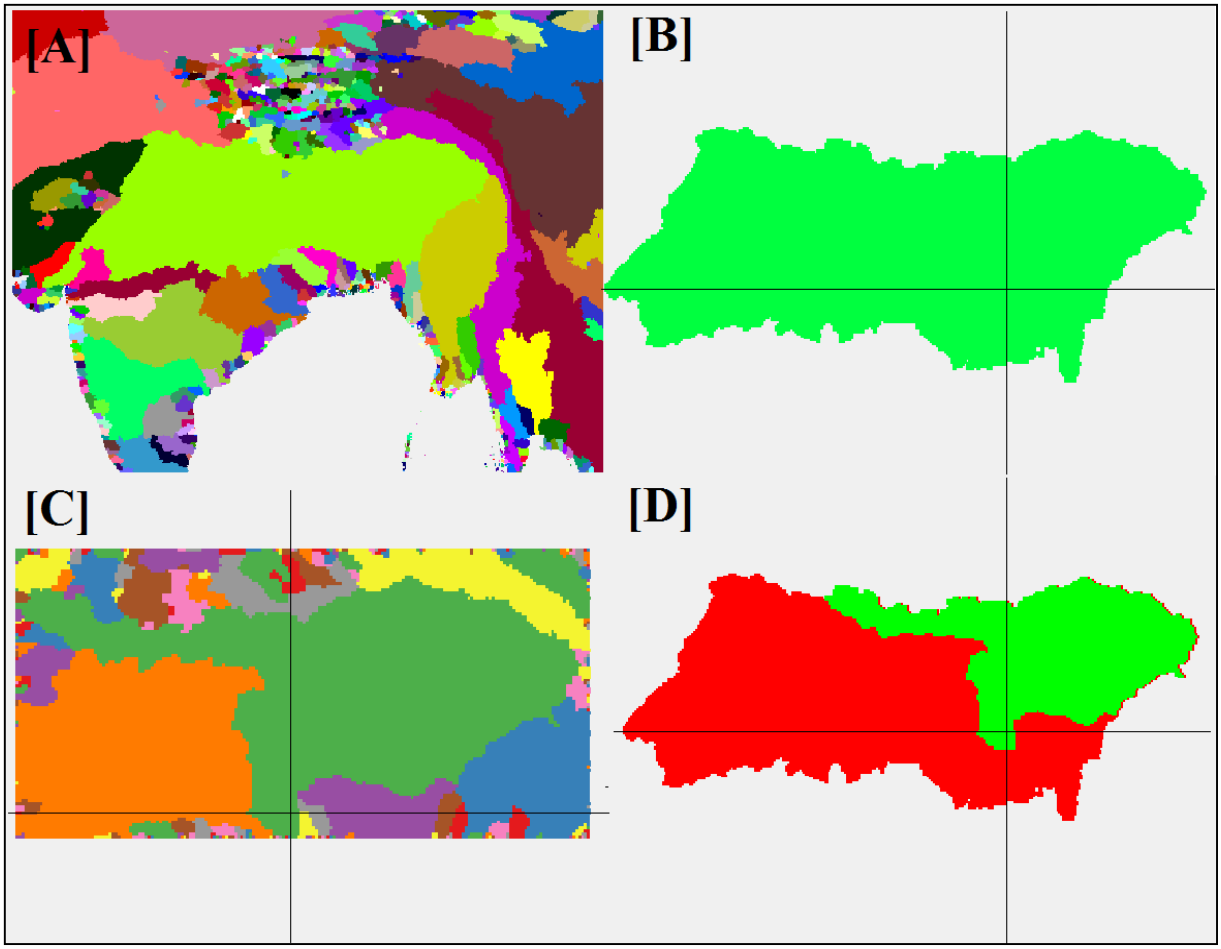


Figure A1. The PCRaster spatial extent creation in a view pictures. A: The calculated catchment map based on the 5 arc minute LDD map. The large yellow basin is the GBM basin (without coastal zone). B: The total GBM basin including the entire coastal zone. The crosshair points to the location of the Bahadurabad gauging station in northern Bangladesh (25°18'N, 89°66'E). C: After cutting out unnecessary parts, catchments were redefined. D: A comparison map showing the Brahmaputra River Basin compared to the total GBM basin.

Appendix B: In-situ Data

Table B1. Collected temperature data of local stations. Latitude and longitude coordinates are in decimal degrees. Sources: (*1*) IWM, personal communication (2015); (*2*) Weather Underground (2015).

Country	Station	Latitude	Longitude	Time Period	% MV	Source
Bangladesh	Bogra	24.83	89.46	2004-2015	5.41	2
Bangladesh	Sylhet	24.90	91.88	2004-2015	1.61	2
Bangladesh	Dinajpur	25.68	88.62	2011-2015	8.08	2
Bangladesh	Saidpur	25.76	88.91	2011-2015	8.91	2
Bangladesh	Rangpur	25.76	89.28	2004-2015	5.96	2
India	Tezpur	26.62	92.80	2004-2015	1.61	2
India	North-Lakhimpur	27.24	94.11	2005-2015	1.53	2
India	Dibrugarh	27.48	95.03	2004-2015	0.19	2
China, Tibet	Pagri	27.73	89.08	2004-2015	0.00	2
China, Tibet	Lhunze	28.42	92.47	2004-2015	0.00	2
China, Tibet	Tingri	28.63	87.08	2004-2015	0.00	2
China, Tibet	Xigaze	29.27	88.90	2004-2015	0.00	1,2
China, Tibet	Nyingchi	29.55	94.54	2004-2015	0.00	2
China, Tibet	Lhasa	29.65	91.14	2004-2015	0.00	1,2
China, Tibet	Xainza	30.95	88.63	2004-2015	0.00	2
China, Tibet	Baingoin	31.37	90.02	2004-2015	0.00	2
China, Tibet	Nagqu	31.48	92.07	2004-2015	0.00	2
China, Tibet	Sog Xian	31.88	93.78	2004-2015	0.00	2

Table B2. Collected precipitation data of local stations. Latitude and longitude coordinates are in decimal degrees. Sources: (1) IWM, personal communication (2015); (2) Weather Underground (2015).

Country	Station	Latitude	Longitude	Time Period	% MV	Source
Bangladesh	Munshigar	23.55	90.50	1970-2014	1.07	1
Bangladesh	Dhaka	23.70	90.48	1970-2014	2.43	1
Bangladesh	Pabna	23.98	89.30	1970-2014	0.02	1
Bangladesh	Joydevpur	24.01	90.42	1970-2011	0.01	1
Bangladesh	Rajshahi	24.38	88.49	1965-2014	1.21	1
Bangladesh	Serajganj	24.39	89.73	1970-2014	1.59	1
Bangladesh	Mymensingh	24.76	90.47	1970-2015	1.35	1
Bangladesh	Bogra	24.83	89.46	1970-2014	1.94	1
Bangladesh	Naogoan	24.89	88.91	1970-2012	1.54	1
Bangladesh	Jamulpur	24.89	89.94	1970-2014	2.74	1
Bangladesh	Joypurhat	25.07	89.02	1970-2013	0.55	1
Bangladesh	Gaibanda	25.18	89.47	1970-2014	3.19	1
Bangladesh	Chilmari	25.56	89.68	1965-2014	2.53	1
Bangladesh	Dinajpur	25.68	88.62	1965-2014	2.67	1
Bangladesh	Rangpur	25.76	89.28	1970-2014	1.87	1
Bangladesh	Thakurgaon	26.09	88.43	1965-2014	3.66	1
Bangladesh	Panchagarh	26.32	88.56	1965-2014	3.25	1
India	Dhollabazar	24.52	90.29	2009-2014	19.21	2
India	Matijuri	24.85	92.59	2009-2014	0.00	2
India	Shillong	25.57	91.89	2009-2014	4.06	2
India	Kohima	25.65	94.17	2009-2014	9.22	2
India	Dhubri	26.02	90.00	2009-2014	8.85	2
India	Guwahati	26.08	91.58	2009-2014	4.66	2
India	Dharamtul	26.17	92.35	2009-2014	23.82	2
India	Goalpara	26.19	90.63	2009-2014	19.03	2
India	Coochbihar	26.34	89.47	2009-2014	10.14	2
India	Beki Road Bridge	26.42	90.94	2009-2014	17.66	2
India	Manas N H Xing	26.45	90.75	2009-2014	21.25	2
India	AIE N H Xing	26.57	90.55	2009-2014	22.46	2
India	Tezpur	26.62	92.80	2009-2014	8.40	2
India	Beki Mathanguri	26.76	90.99	2009-2014	24.24	2
India	Jorhat	26.93	94.54	2009-2014	8.72	2
India	North-Lakhimpur	27.24	94.11	2009-2014	9.82	2
India	Gangtok	27.33	88.63	2009-2014	5.80	2
India	Itnagar	27.37	93.64	2009-2014	7.80	2
India	Badatighat	27.48	94.27	2009-2014	22.87	2
India	Dibrugarh	27.48	95.03	2009-2014	8.05	2
India	Pasighat	28.09	95.39	2009-2014	8.26	2
India	Tezu	28.10	96.27	2009-2014	24.24	2
China, Tibet	Xigaze	29.27	88.90	2009-2014	24.69	2
China, Tibet	Nyingchi	29.55	94.54	2009-2014	16.34	2
China, Tibet	Lhasa	29.65	91.14	2009-2014	18.26	2

Appendix C

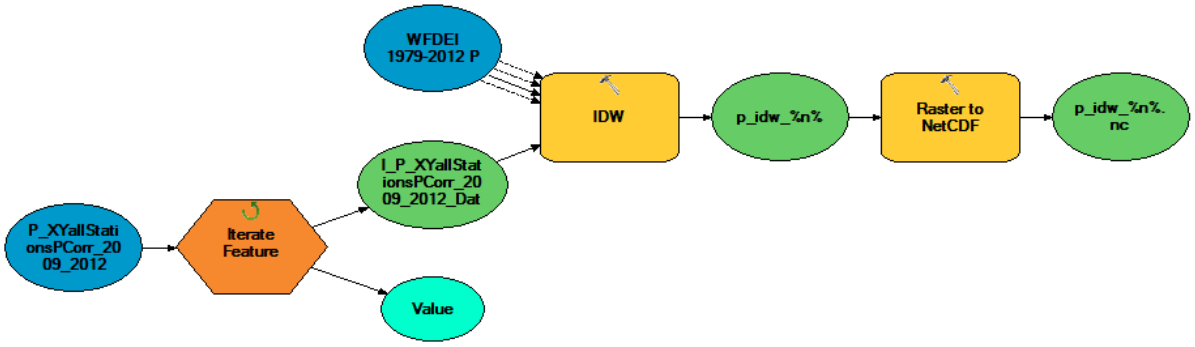


Figure C. The created ArcMAP model to create a NetCDF file per timestep. ArcMAP was limited in merging the created NetCDF files into one NetCDF for the PCR-GLOBWB model. An iteration tool was used to apply interpolation over all timesteps present in the point data shapefile. Per timestep the point data was interpolated with environment conditions set using the original WFDEI file. The created raster file was then converted to an NetCDF file.

Appendix D

Dataset 1 and 2 have an overlap period of half a year in 2012. The discharge presented by the datasets during this overlap period differ by only 0.25% ± 0.0048. The consistency in the deviation is due to the way the Rated Q is determined. To determine the discharge a rating curve is used. Water height of the river at this location is measured daily; however, the discharge is determined afterwards. The process involves the determination of the cross-sectional area of the river at the location of the water height measurement. The cross-sectional area is measured once or twice per month. The reason for this is the fact that the river at this location is nearly 11 km wide. The measurements of the cross-sectional area are used to adjust the rating curve over which the discharge is measured. In short, the values of the 'observed' discharge are directly related to the water height with an assumption made about the cross-sectional area. Dataset 1 and 2 use different rating curve equations, but are still similar.

Table 6. Discharge data from the Bahadurabad gauging station (25°18'N, 89°66'E) (BWDB, 2015). The dataset used in this study was created from two separate datasets. This is tied to the establishment of the rating curve over which the Rated Q is determined.

Range	Interval	Time of measurement	Description
1985-1986	Daily	18:00	Dataset 1
1986-1992	Daily	06:00	Dataset 1
1992-1998	Daily	00:00	Dataset 1
1998-1999	3-Hourly	00:00, 03:00, 06:00, 18:00, 21:00	Dataset 1
1999-2012	3-Hourly	06:00, 09:00, 12:00, 15:00, 18:00	Dataset 1
2012-2014	3-Hourly	06:00, 09:00, 12:00, 15:00, 18:00	Dataset 2