  Hydrology Research | 46.5 | 2015

# Application of satellite-derived rainfall for hydrological modelling in the data-scarce Black Volta trans-boundary basin

Kwaku Amaning Adjei, Liliang Ren, Emmanuel Kwame Appiah-Adjei and Samuel Nii Odai

## ABSTRACT

This study, conducted in the Black Volta basin of Ghana, determined how well TRMM Multi-Satellite Precipitation Analysis (TMPA) data compare with rain gauge measurements. The potential of using the TMPA data as inputs into a hydrological model for runoff simulation in a data-scarce basin was also assessed. Using a point-to-grid approach, accumulations of ground measured rainfall on daily, monthly and annual time scales were compared with accumulations derived from TMPA daily rainfall grids. The TMPA derived data together with other free global data were used as input into the soil and water assessment tool (SWAT) to set up a hydrological model for the basin. This model was calibrated and validated using streamflow data from a station located downstream of the basin. The study results showed a correlation from 0.85 to 0.92 for the monthly accumulated rainfall. Also, good Nash–Sutcliffe efficiencies of 0.94 and 0.67 were obtained for calibration and validation, respectively, on monthly scale. Moreover, simulation of streamflow was 'satisfactory' to 'very good' in terms of trends and residual variation. The study, therefore, shows that the use of satellite rainfall in the basin would be of great benefit considering the difficulties in accessing data across the basin.

**Key words** | Black Volta, data-scarce region, Prediction in Ungauged Basins (PUB), SWAT, TMPA

**Kwaku Amaning Adjei** (corresponding author)
**Liliang Ren**
State Key Laboratory of Hydrology – Water Resources and Hydraulic Engineering,
College of Hydrology and Water Resources,
Hohai University,
No.1 Xikang Road, Nanjing 210098,
China
E-mail: *nanakadjei@yahoo.com*

**Kwaku Amaning Adjei**
**Samuel Nii Odai**
Department of Civil Engineering,
Kwame Nkrumah University of Science and Technology,
PMB, Kumasi,
Ghana

**Emmanuel Kwame Appiah-Adjei**
Department of Geological Engineering,
Kwame Nkrumah University of Science and Technology,
PMB, Kumasi,
Ghana

## INTRODUCTION

Hydrological modelling in data-scarce basins across the globe has always been a challenge. This situation is compounded when dealing with large-scale and trans-boundary basins. Many experts in the field of hydrology believe most of the catchments in the world are either ungauged or poorly gauged (Kundzewicz 2007). Hence, there have been initiatives by the International Association of Hydrological Sciences to improve on this situation with the introduction of the Prediction in Ungauged Basins (PUB) initiative (Sivapalan *et al.* 2003; Hrachowitz *et al.* 2013). Even though this initiative has yielded remarkable results (Franks 2005; Raju 2006; Sivapalan 2006), a great deal more still needs to be done in the area of the discovery of methodologies and data sources that can help in hydrological predictions and studies in poorly gauged basins.

The Volta basin is an important trans-boundary catchment in West Africa with very significant resources for its riparian countries. Even though this basin is important, adequate ground measurements of both meteorological and hydrological parameters are not available (GLOWA Volta Project (GVP) 2007; Taylor *et al.* 2007; Schuol *et al.* 2008), and in cases where there are measurements, they are often plagued with a great deal of missing data (Adjei *et al.* 2012). Also, access to data collected by institutions, researchers and other stakeholders from riparian countries is very difficult. Even though there have been some initiatives like the Global Change in the Hydrological Cycle (GLOWA) Volta project (Rodgers *et al.* 2007), the Volta Basin Authority and others to help bridge this gap, there still are challenges in respect of access to basin-wide hydrological and

doi: 10.2166/nh.2014.111

meteorological measurements. In recent times, various global data sets with durations significant for hydrological analysis, especially in poorly gauged catchments, have been made available. Current sources of some of these data sets are the Rainfall Estimates (USGS 2012), Global Land Data Assimilation System (Fang *et al.* 2009) and the TRMM Multi-Satellite Precipitation Analysis (TMPA) (Adler *et al.* 2007).

Satellite rainfall has been validated (Islam & Uyeda 2007; Adjei *et al.* 2012; Li *et al.* 2012; Yong *et al.* 2012) at different scales across the globe and has been used alone or as a complementary data source, satisfactorily, in some hydrological modelling studies (Tobin & Bennett 2009; Li *et al.* 2012, 2013). However, not much has been done in the use of satellite rainfall for hydrological studies in the Black Volta basin. Earlier studies in the basin obtained satisfactory results for the validation of the TMPA version 6 data set for the basin (Adjei *et al.* 2012). However, in this study, version 7 of the TMPA data set is used to set up a semi-distributed hydrological model for the Black Volta basin. Satellite-based rainfall studies are especially good for data-scarce basins because of their frequency of update (Li *et al.* 2013). Research carried out by Li *et al.* (2013) in the southern African region and Jung *et al.* (2012) in the Volta basin underscored the increasing use of large-scale hydrological models as the main assessment tool for global/regional water resources and the challenges associated with basins in Africa, which usually have sparse data and are poorly gauged. The availability of high quality and consistent data is usually a challenge making it difficult to set up regional models for water resources assessment. Thus, the aim of the research was: (1) to validate satellite-derived rainfall estimates by comparing them with ground-measured rainfall; and (2) to assess the suitability of using the satellite rainfall with other free global data sets as inputs into soil and water assessment tool (SWAT) to set up a hydrological model for runoff simulation in the Black Volta basin, which has been described as data scarce (Van de Giesen *et al.* 2001).

The SWAT model used in this study was generally developed for predicting the impact of land management practices on water, sediment and agricultural chemical yields in large complex watersheds with varying soils, land use and management conditions over long periods of time. It is a continuous time model, semi-distributed and physically based (Arnold *et al.* 1998; Neitsch *et al.* 2011). Due to lack of data in some parts of the basin and the sparse distribution of rain gauges, rainfall estimates provided by the USGS were used to model streamflow in the Black Volta basin.

TMPA rainfall has been used for the hydrological model in a few reported studies including a study by Mango *et al.* (2011) in the trans-boundary Mara River basin, which is shared between Kenya and Tanzania in East Africa. In their study, Mango *et al.* (2011) demonstrated that a semi-distributed hydrological model such as SWAT can be set up and calibrated in a poorly gauged rural African catchment to yield useful results from satellite-based rainfall estimates, which can be used to study catchment changes and management scenarios useful for decision-making (Mango *et al.* 2011). Another study by Schuol & Abbaspour (2006), using data from CRU, demonstrated that global data sources could be used to calibrate and validate large-scale hydrological models in data-scarce regions like West Africa. The two studies above (i.e., Schuol & Abbaspour 2006; Mango *et al.* 2011) present an opportunity to further explore how well global data sources can be used for large-scale hydrological modelling and, even, in small-scale basins where data are scarce. Also, these two studies show the possibility of using global data sets for large-scale hydrological modelling at a finer spatial resolution. However, other critical inputs to the SWAT model such as land use/cover, digital elevation model (DEM) and others could also be improved with fine resolution data sets like the SRTM (US Geological Survey 2012), Advanced Spaceborne Thermal Emission and Reflection Radiometer (US Geological Survey 2012) and GlobCover 2009 (Bontemps *et al.* 2011) data sets. Thus, other methods and approaches could be applied to understand the best way for using these global data sets for hydrological modelling in data-scarce regions.

## STUDY AREA AND DATA

The Black Volta basin, with a total area of about 155,000 $km^2$ (Figure 1) at the hydrological station in Bamboi, is shared by four West African countries. These are Burkina Faso, Cote d'Ivoire, Ghana and Mali. It is also

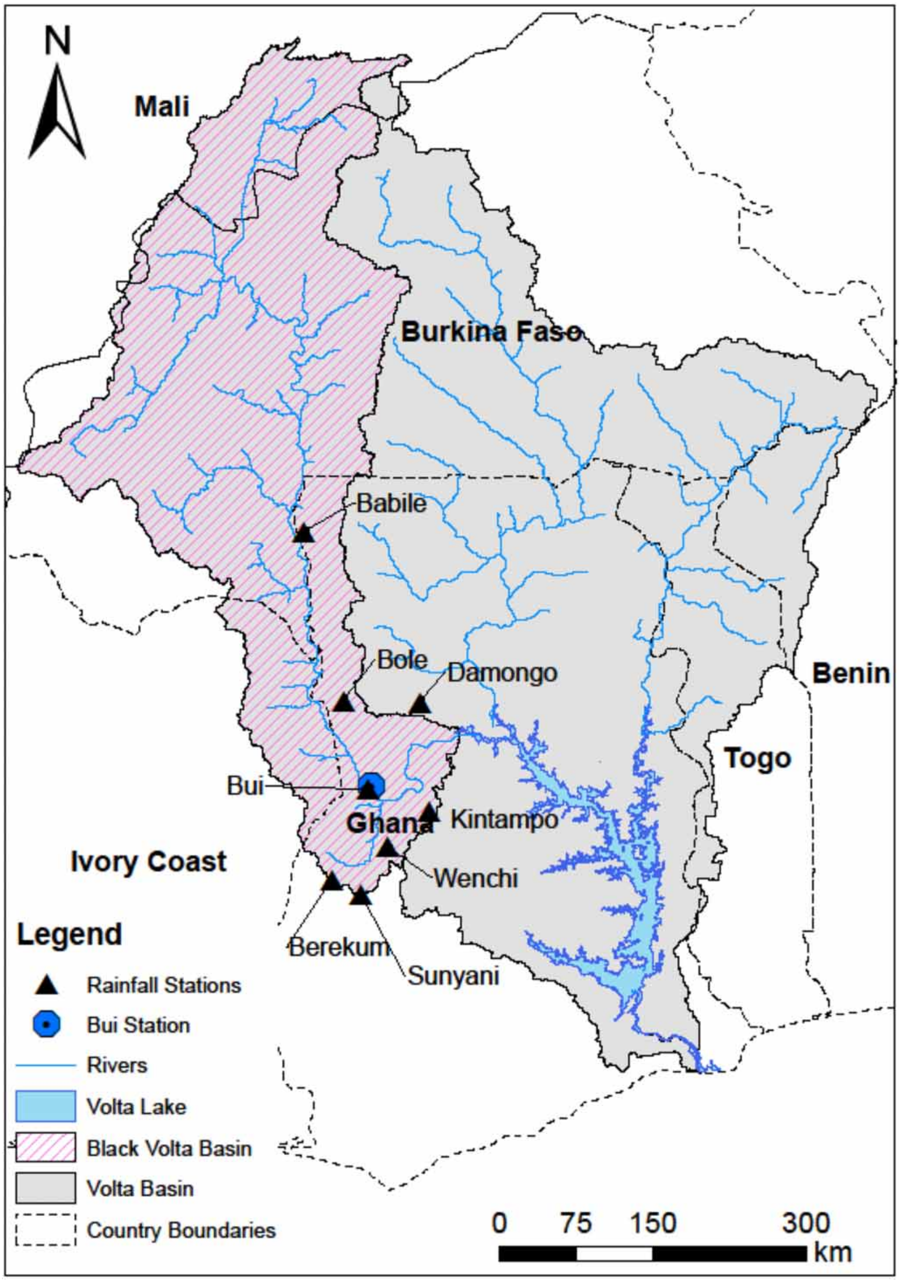


**Figure 1** | Map of the study area showing the riparian countries and the Volta basin.

one of the three main sub-basins of the Volta basin. The average runoff from the Black Volta basin is about 243 $m^3/s$ and it contributes about 18% of the total flows to Lake Volta. The specific suspended sediment yield in this basin ranges from 8.0 to 12.0 tonnes/yr/$km^2$. Current surface water use is estimated to be only about 0.03 $m^3/s$ for domestic water supply (FAO 1997; Barry *et al*. 2005).

The climate of the Black Volta basin is semi-arid to sub-humid with reference evaporation exceeding precipitation. There are two distinct seasons in the basin, namely the dry and rainy seasons, which are largely influenced by the West African monsoon (Global Water Partnersip 2012). The annual rainfall in the basin ranges from 600 mm to 1,348 mm while the reference evaporation ranges from

1,450 mm to 1,800 mm. The basin has a mean annual runoff of $7,673 \times 10^6$ m$^3$ and a length of about 1,363 km (Barry *et al.* 2005).

The major land use in the basin is agriculture with food crop cultivation under extensive bush fallow. The major food crops include yam, cassava, maize, sorghum, millet, groundnuts and beans. Animal grazing on free range is a significant activity. In the northwest of the basin, lands are highly degraded both in terms of physical status and fertility levels, and can barely support meaningful crop cultivation (Barry *et al.* 2005).

To set up a SWAT model, the DEM, digital land cover/land use map, digital soil map with physical properties, meteorological data (precipitation, dew point or relative humidity, maximum and minimum temperature, potential evapotranspiration, if available on daily time scale) and stream flow measurements are required. Some of these required data can be simulated if a custom weather generator has been created for the basin. Hence, the available limited data obtained from the Ghana Meteorological Agency (GMA) were used to create a custom weather generator for the basin. In all, three meteorological stations, Bole, Sunyani and Wenchi (see Figure 1), which had measurements for the parameters required for the creation of the weather generator, were used. The other stations (Berekum, Bui, Damongo and Kintampo) were added to the three and used for the validation of the satellite-derived rainfall. To set up the model, free global data from various sources listed in Table 1 were used.

**Table 1** | Data and source

| Data type | Source | Remarks |
|---|---|---|
| Digital elevation model | NASA Space Shuttle Radar Topography Mission (SRTM) | 90 m × 90 m |
| Land cover data | European Space Agency (ESA)-GlobCover 2009 project | 300 m × 300 m |
| Soil data | FAO-UNESCO Soil Map of the World (Batjes 1997) | 1:500,0000 scale |
| Measured (rainfall, temperature, wind speed, relative humidity) | Ghana Meteorological Service | 1998–2010[a] |
| Satellite rainfall (TMPA) | USGS Data Portal | 1998–2010 |
| Discharge data | Ghana Hydrological Service | 1998–2010[a] |

[a]With missing values.

## METHODOLOGY AND TOOLS

The data were obtained from the TRMM Online Visualization and Analysis System. The downloaded TMPA data were first compared with measured rainfall from some selected rain gauge (RNG) stations within the basin with minimal missing values before pre-processing them for use in ARCSWAT. A point-to-grid comparison was used to extract the daily time series from the data sets. Thus, the value for the grid within which a meteorological station is located was extracted and used for analyses. Using this approach, the time series data were extracted for the selected representative stations across the basin and compared with their respective RNG measurements.

The TMPA data set is produced as daily grids and has a spatial resolution of $0.25^\circ \times 0.25^\circ$. SWAT, however, does not accept gridded meteorological data but uses the stations closest to the centroid of a sub-basin as the meteorological station for that sub-basin. Thus, to use the TMPA data in SWAT, the centroids of the grids were calculated and used as virtual rainfall stations for the distributed nature of the satellite rainfall to be effectively utilized. SWAT requires other data sets in gridded formats; hence, the other required gridded inputs were obtained from freely available global sources (Table 1) and used for setting up the model.

The model calibration and validation were carried out using the sequential uncertainty fitting algorithm (SUFI-2) SWAT calibration and uncertainty procedures (SWAT-CUP) (Abbaspour *et al.* 2004, 2007; Schuol & Abbaspour 2006) with the aid of streamflow data from the Bui hydrological station. The data used for the modelling were from 1998 to 2010 (Table 1). The calibration of the model parameters was carried out using two approaches; namely, the value and relative approaches.

A code was placed beside each parameter to be calibrated to indicate the type of change to be applied to the parameter. In the case of the parameters whose initial values were obtained at the hydrologic response unit (HRU) level based on soil, land use or slope map, the relative (*r*) change method was used to preserve the variability in the values across the different sub-basins. The value (*v*) change method was, however, used for those parameters having a single value for the basin. The '*v*' change method means the existing parameter value is to be replaced by the given value while the '*r*' change method means the existing parameter value is to be multiplied by (1+ a given value).

In the model validation process, the final parameter set of the simulation that gave the best fit and, thus, had the highest Nash–Sutcliffe efficiency (NSE) was used to re-run the SWAT model without adjusting the parameters. The performance of the model in both calibration and validation processes were evaluated using the recommended statistics by Moriasi *et al.* (2007). These recommended statistics are the NSE in Equation (1), percent bias (PBIAS) in Equation (2) and the root mean square error observations standard deviation ratio (RSR) in Equation (3). Their study (Moriasi *et al.* 2007) recommended ranges of these statistics (Table 2) that can be used to assess the performance of a calibrated model based on streamflow.

$$NSE = 1 - \left( \frac{\sum_{i=1}^{n} (O_i - S_i)^2}{\sum_{i=1}^{n} (O_i - \overline{O})^2} \right) \quad (1)$$

$$RSR = \frac{RMSE}{STDEV_O} = \left( \frac{\sqrt{\sum_{i=1}^{n} (O_i - S_i)}}{\sqrt{\sum_{i=1}^{n} (O_i - \overline{O})^2}} \right)^2 \quad (2)$$

$$PBIAS = \left( \frac{\sum_{i=1}^{n} (O_i - S_i)}{\sum_{i=1}^{n} (O_i)} \times 100 \right) \quad (3)$$

## RESULTS AND DISCUSSION

### Comparison of daily rainfall

The results from the comparison of the daily rainfall from the TMPA data set and the rain gauge measurements show very weak correlations, ranging from 0.25 to 0.40, between the two data sets (Table 3). This poor correlation is as a result of the differences in the time of accumulation for the two different data sets. The daily accumulated TMPA begins at 00Z and ends at 21Z, which is equivalent to the UTC time zone. However, the ground measured accumulations are from 09:00am to 09:00am UTC. This difference in the accumulation periods is one of the main reasons for the very poor correlation between the data sets. This was further corroborated when the correlation coefficients between the two data sets improved when a 2-day running average was carried out on the two data sets (Table 3). Another possible source of error for the low daily correlations and underestimations may be due to the nature of rain that falls within the Volta basin, which is rainfall controlled by the Inter-tropical Convergence Zone. This kind of rain is mainly convective in nature and has high intensity, but of short duration, usually 20–60 min (Salako 2007; Van de Giesen *et al.* 2011). The low daily correlation difficulties observed are not really peculiar to this study but were also observed by Islam & Uyeda (2007), who found that the estimation patterns of satellite rainfall showed good similarity with the RNG values, but the numbers of deviations were very high. However, they attributed the deviations to the

**Table 2** | General performance ratings of recommended statistics for a monthly time step (Moriasi *et al.* 2007)

| Performance rating | RSR | NSE | PBIAS (%) |
|---|---|---|---|
| Very good | 0.00 < RSR <0.50 | 0.75 < NSE <1.00 | PBIAS < ±10 |
| Good | 0.50 < RSR <0.60 | 0.65 < NSE <0.75 | ± 10 < PBIAS < ±15 |
| Satisfactory | 0.60 < RSR <0.70 | 0.50 < NSE <0.65 | ± 15 < PBIAS < ±25 |
| Unsatisfactory | RSR >0.70 | NSE <0.50 | PBIAS > ±25 |

**Table 3** | Pearson correlation between rain gauge and TMPA satellite measurements

| | Pearson correlation (R) | | |
|---|---|---|---|
| Station | Original data set | 2-day running average | Mean annual difference (mm)[a] |
| Berekum | 0.27 | 0.45 | 32 |
| Bole | 0.34 | 0.48 | −39 |
| Bui | 0.25 | 0.44 | −34 |
| Damongo | 0.34 | 0.51 | 20 |
| Kintampo | 0.32 | 0.48 | −72 |
| Sunyani | 0.36 | 0.52 | 16 |
| Wenchi | 0.31 | 0.48 | −32 |

[a]Mean annual difference = TMPA − RNG.

short life of the cloud cells and the frequency of satellite passes during the 24-hour period.

## Comparison of monthly rainfall

The correlation coefficients for the monthly accumulations, which were obtained from the daily data set, ranged from 0.85 to 0.92 with a mean monthly rainfall of all the stations having a correlation coefficient of 0.97 (Table 4). These were tremendous improvements in comparison to the daily accumulation correlations in the previous section because the time of accumulation was not very significant on a monthly basis. This shows that whenever a TMPA data set is to be compared with rain gauge data, the time of accumulation (24-hour meteorological day definition) needs to be chosen appropriately to avoid arriving at any wrong conclusions in the two data sets.

**Table 4** | Correlation between monthly rain gauge and TMPA satellite accumulations

| Station | Pearson correlation (R) |
|---|---|
| Berekum | 0.87 |
| Bole | 0.90 |
| Bui | 0.85 |
| Damongo | 0.90 |
| Kintampo | 0.86 |
| Sunyani | 0.92 |
| Wenchi | 0.87 |
| Average | 0.97 |

Also, in the monthly accumulations, the overestimations (positive) and underestimations (negative) cancel out to give an improved correlation. The stations at the middle part of the catchment (i.e., Bole and Damongo) showed relatively higher correlations in comparison to the stations at the lower belt, except for Sunyani station.

The long-term monthly mean data sets of both the TMPA and RNG follow the same trends as those observed in Figure 2. The TMPA overestimated, slightly, by 0.42 to 19.0 mm/month in seven months at a mean overestimation of 6.52 mm/month. This result indicates that the TMPA can be used to describe the long-term rainfall characteristics of the basin.

Both underestimations and overestimations were recorded at all the stations (Figure 3). However, the TMPA is more likely to overestimate in the basin. The number of months in which TMPA overestimated was higher than 50% at all the stations except at Kintampo station, which was slightly lower than 50% (Table 5). The bulk of underestimations or overestimations at most stations was below 30 mm with just a few stations above 30 mm.

## Comparison of annual rainfall

The results of the annual rainfall calculation for each station are shown in Tables 6 and 7. The result indicates both underestimation (negative bias) and overestimation (positive bias) of the annual RNG rainfall by the TMPA annual accumulations.

The overestimations and underestimations were mostly below 20% of the RNG measured rainfall with just a few exceeding 20% at some stations. These results show that the TMPA data can be used to describe the long-term climate of a basin when gauged data are not available since the biases were relatively small in most of the years. The correlation between the two data sets can be improved through error modelling in the catchment based on a dense network or wide distribution of meteorological stations.

## Selection of pre-calibration parameters

The selection of pre-calibration parameters was done through a review of the literature on the application of SWAT in different catchments across the globe and, especially, in catchments, which have characteristics similar

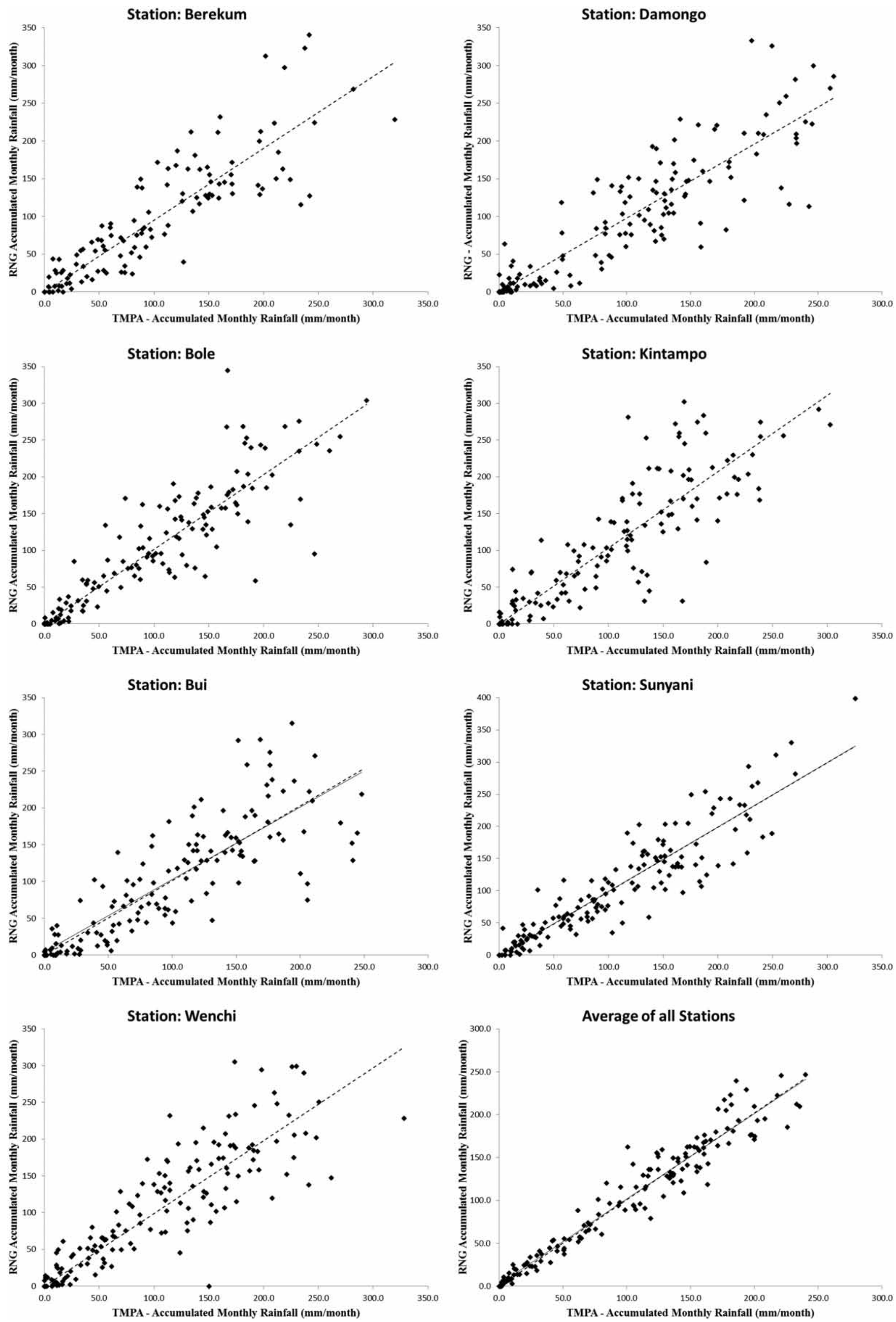


**Figure 2** | Plot of accumulated monthly rainfall of TMPA versus RNG.

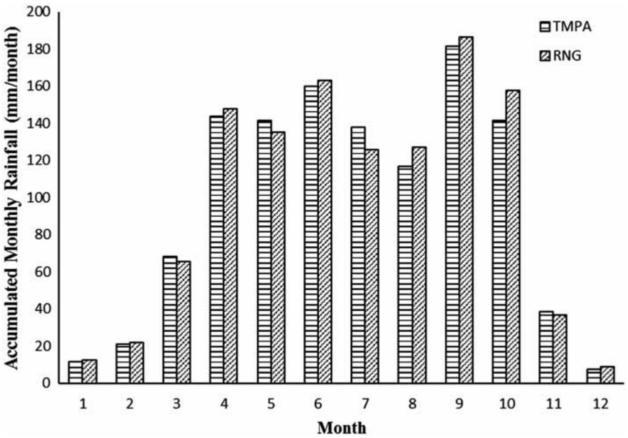


**Figure 3** | Average monthly rainfall of the TMPA versus RNG for all the stations from 1998 to 2010.

to the Black Volta basin (Schuol *et al.* 2008; Rouhani *et al.* 2009; Fadil *et al.* 2011; Mango *et al.* 2011; Nie *et al.* 2011; Obiero *et al.* 2011; Güngör & Göncü 2013). Further to the selection of the parameters, the SUFI-2 of the SWAT-CUP tool which uses the Latin Hypercube (McKay *et al.* 1979) sampling method for parameter calibration was used to narrow down the number of parameters for model calibration. Many of the parameters used to calibrate SWAT, like curve number (CN2), available water capacity (AWC) or the bulk densities (BD) are closely related to different land use classes and soil textures; hence, their parameter values change accordingly. In all, a total of 28 parameters (13 global parameters, one parameter with a separate value for each of the six dominant land use types and three parameters with a separate value for each of the dominant soil textures, i.e., AWC, K and Z for sand, sandy loam and sandy clay loam soils (Table 9)) were selected for the calibration of the model after the sensitivity approach.

## Sensitivity analysis of parameters

The sensitivity of the parameters used in the modelling are summarized in Table 8. The ranking of the parameters are based on t-stat from the sensitivity analysis of all the parameters for the whole basin. Parameter sensitivities were determined by calculation, using the multiple regression system, which regresses the Latin Hypercube generated parameters against the objective function values. The t-stat was used as a measure of sensitivity (larger in absolute values are more sensitive) and the p-values determined the significance of the sensitivity (values close to zero have more significance) (Abbaspour 2011). Analyses of the sensitivity results (Table 8) show the most sensitive parameters were the

**Table 5** | Comparison of TMPA and RNG performances for selected stations from 1998 to 2010

| | | Number of months within bias range for station | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Bias range (mm)/month | Berekum | Bole | Bui | Damongo | Wenchi | Sunyani | Kintampo | Average |
| Overestimation | > 30 | 20 | 13 | 27 | 23 | 22 | 21 | 26 | 4 |
| | 30–20 | 18 | 8 | 11 | 13 | 6 | 16 | 10 | 10 |
| | 20–10 | 12 | 18 | 16 | 18 | 16 | 17 | 14 | 15 |
| | 0–10 | 17 | 50 | 28 | 36 | 30 | 27 | 19 | 47 |
| Total number of months (count) | | 67 | 89 | 82 | 90 | 74 | 81 | 69 | 76 |
| Total number of months (%) | | 55.8 | 57.1 | 52.6 | 57.7 | 51.4 | 51.9 | 44.2 | 48.7 |
| Total number of months (count) | Equal monthly rain | 3 | 2 | 5 | 7 | 2 | 6 | 5 | 0 |
| Total number of months (%) | | 2.5 | 1.3 | 3.2 | 4.5 | 1.4 | 3.8 | 3.2 | 0.0 |
| Underestimation | 0–10 | 10 | 19 | 16 | 19 | 13 | 27 | 20 | 48 |
| | 20–10 | 13 | 15 | 13 | 6 | 18 | 8 | 20 | 14 |
| | 30–20 | 8 | 6 | 7 | 10 | 4 | 11 | 9 | 7 |
| | > 30 | 19 | 25 | 33 | 24 | 33 | 23 | 33 | 11 |
| Total number of months (count) | | 50 | 65 | 69 | 59 | 68 | 69 | 82 | 80 |
| Total number of months (%) | | 41.7 | 41.7 | 44.2 | 37.8 | 47.2 | 44.2 | 52.6 | 51.3 |

**Table 6** | Comparison of TMPA and RNG performances in the estimation of annual rainfall for some stations from 1998 to 2010

| | Berekum | | | Bole | | | Bui | | | Damongo | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Year | TMPA | RNG | %Bias | TMPA | RNG | %Bias | TMPA | RNG | %Bias | TMPA | RNG | %Bias |
| 1998 | 1065 | 1024 | 4.0 | 885 | 950 | – 6.8 | 965 | 833 | 15.9 | 874 | 716 | 22.0 |
| 1999 | 1303 | 1102 | 18.3 | 1273 | 1163 | 9.4 | 1223 | 1219 | 0.3 | 1309 | 1275 | 2.7 |
| 2000 | 1084 | 1056 | 2.6 | 1178 | 1226 | – 4.0 | 1051 | 1150 | – 8.6 | 1100 | 1117 | – 1.5 |
| 2001 | 1171 | 1125 | 4.1 | 915 | 904 | 1.2 | 1084 | 1052 | 3.0 | 889 | 747 | 19.0 |
| 2002 | 1142 | 1363 | – 16.2 | 1073 | 1153 | – 6.9 | 1014 | 1303 | – 22.2 | 1015 | 1010 | 0.5 |
| 2003 | 1321 | 952 | 38.8 | 1161 | 1251 | – 7.2 | 1130 | 1110 | 1.8 | 1234 | 1127 | 9.5 |
| 2004 | 1270 | 1333 | – 4.7 | 1110 | 1245 | – 10.8 | 1225 | 981 | 24.9 | 1165 | 1085 | 7.4 |
| 2005 | 1235 | 1183 | 4.4 | 1159 | 1187 | – 2.3 | 1180 | 1328 | – 11.1 | 1100 | 1073 | 2.5 |
| 2006 | 1209 | 1319 | – 8.3 | 1069 | 1108 | – 3.5 | 1119 | 1048 | 6.9 | 990 | 973 | 1.7 |
| 2007 | 1237 | 1256 | – 1.5 | 1057 | 949 | 11.4 | 893 | 955 | – 6.5 | 1229 | 996 | 23.3 |
| 2008 | 1230 | | | 1086 | 1089 | – 0.3 | 1077 | 1261 | – 14.6 | 1141 | 1378 | – 17.2 |
| 2009 | 1088 | | | 1132 | 1135 | – 0.3 | 1104 | 1121 | – 1.5 | 989 | 1143 | – 13.5 |
| 2010 | 1351 | | | 1151 | 1393 | – 17.3 | 1134 | 1285 | – 11.7 | 1160 | 1290 | – 10.1 |

**Table 7** | Comparison of TMPA and RNG performances in the estimation of annual rainfall for some stations from 1998 to 2010

| | Kintampo | | | Sunyani | | | Wenchi | | | Average | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Year | TMPA | RNG | %Bias | TMPA | RNG | %Bias | TMPA | RNG | %Bias | TMPA | RNG | %Bias |
| 1998 | 1093 | 1290 | – 15.2 | 1091 | 914 | 19.4 | 1101 | 1053 | 4.6 | 1011 | 969 | 4.4 |
| 1999 | 1359 | 1650 | – 17.6 | 1296 | 1178 | 10.0 | 1367 | 1289 | 6.1 | 1304 | 1268 | 2.9 |
| 2000 | 1272 | 1249 | 1.9 | 1079 | 1018 | 6.0 | 1097 | 1187 | – 7.5 | 1123 | 1143 | – 1.8 |
| 2001 | 954 | 869 | 9.8 | 1223 | 1238 | – 1.3 | 1179 | 987 | 19.4 | 1059 | 989 | 7.1 |
| 2002 | 1342 | 1635 | – 17.9 | 1239 | 1224 | 1.2 | 1284 | 1413 | – 9.1 | 1158 | 1300 | – 10.9 |
| 2003 | 1308 | 1226 | 6.7 | 1344 | 1326 | 1.4 | 1306 | 1396 | – 6.5 | 1258 | 1198 | 4.9 |
| 2004 | 1403 | 1531 | – 8.4 | 1270 | 1286 | – 1.2 | 1264 | 1350 | – 6.3 | 1244 | 1259 | – 1.2 |
| 2005 | 1350 | 1294 | 4.3 | 1203 | 1084 | 11.0 | 1201 | 1330 | – 9.7 | 1204 | 1211 | – 0.6 |
| 2006 | 1064 | 1235 | – 13.8 | 1201 | 1156 | 3.9 | 1219 | 1196 | 1.9 | 1125 | 1148 | – 2.0 |
| 2007 | 1196 | 1358 | – 11.9 | 1307 | 1459 | – 10.5 | 1356 | 1216 | 11.5 | 1182 | 1170 | 1.0 |
| 2008 | 1345 | 1112 | 20.9 | 1281 | 1327 | – 3.4 | 1228 | 1311 | – 6.3 | 1193 | 1246 | – 4.3 |
| 2009 | 988 | 1088 | – 9.2 | 1111 | 1307 | – 15.0 | 1052 | 1201 | – 12.4 | 1063 | 1166 | – 8.9 |
| 2010 | 1286 | | | 1478 | 1404 | 5.3 | 1404 | 1549 | – 9.3 | 1266 | 1384 | – 8.6 |

SCS curve number, soil available water storage capacity and the soil evaporation compensation factor. These parameters go to show that more emphasis should be placed on getting very good land use/land cover and soil properties data for the basin since they have a high impact on the model performance.

## Calibration and validation

Using the NSE as the main evaluation criteria, the model was calibrated using monthly observed flows at the Bui discharge station. Although inputs in the model were on a daily time scale the model was calibrated and validated

**Table 8** | Results of sensitive analysis of calibration parameters with their sensitivity ranking

| Parameter type | Definition | Sensitivity rank |
|---|---|---|
| **Global** | | |
| SURLAG[a] | Surface runoff lag coefficient [days] | 15 |
| CH_K2[a] | Effective hydraulic conductivity in the main channel [mm/h] 10 | 4 |
| CH_N2[a] | Manning's n value for the main channel | 12 |
| ESCO[a] | Soil evaporation compensation factor [–] | 3 |
| SOL_BDr | Moist soil BD [g/cm$^3$] | 7 |
| ALPHA_BF[a] | Baseflow alpha factor [day] | 8 |
| GW_DELAY[a] | Groundwater delay time: lag between the time that water exits the soil profile and enters the shallow aquifer [days] | 14 |
| GW_REVAP[a] | Groundwater 'revap' coefficient [–] | 9 |
| GWQMN[a] | Threshold depth of water in the shallow aquifer required for return flow to occur [mm $H_2O$] | 5 |
| RCHRG_DP[a] | Deep aquifer percolation fraction [–] | 13 |
| REVAPMN[a] | Threshold depth of water in the shallow aquifer for 'revap' or percolation to the deep aquifer [mm $H_2O$] | 11 |
| **Land use/cover based** | | |
| CN2[b] | SCS runoff curve number [–] | 1 |
| **Soil type based** | | |
| SOL_AWC[b] | Soil available water storage capacity [mm $H_2O$/mm soil] | 2 |
| SOL_Kr | Soil conductivity [mm/h] | 6 |
| SOL_Z[b] | Soil depth [mm] | 10 |

[a]Existing parameter value is to be replaced by the given value.
[b]Existing parameter value is multiplied by (1+ a given value).

using the monthly outputs. Also, the quality of the final parameter ranges and the fitted values for the last iteration of SUFI-2 are shown in Table 9. The values obtained are consistent with results obtained from SWAT modelling in some major basins in West Africa (Schuol *et al.* 2008).

The performance of the model (Table 10) according to the main efficiency criteria used can be described as *very good* (NSE of 0.94). The other performance indicators also show the model to be performing very well. The calibrated model showed marginal (PBIAS 3.34%) underestimation of the observed flows.

The performance of the model validation, according to the criteria used, can be described as *good* since the NSE is 0.67. However, the RSR, PBIAS and the $R^2$ obtained for the model puts its performance as *satisfactory*. Since the main criterion used in the calibration of the model was the NSE, the performance of the model, in general, can be described as *good*. The model's ability to simulate very high flows during the validation period was not very good since it underestimated the very high peaks for the two events recorded (Figure 4). This observation is not unique to this study but has been observed in similar studies where the TMPA data set was used. For example, a study by Li *et al.* (2012) found that a model with TMPA data set input is unable to simulate the peak flows very well like the model with rain gauge measured rainfall as input. The model underestimated the mean annual flow by 10% of the 292 m$^3$/s observed flow.

The simulated flows showed a narrower standard deviation of 389 m$^3$/s compared to 462 m$^3$/s for the observed flows. This variation is due to the model's inability to simulate very well the highs flows, hence, the narrower standard deviation.

**Table 9** | The 28 SWAT model parameters included in the final calibration, their fitted value and final ranges

| | | | Final parameter range | |
|---|---|---|---|---|
| | **Parameter name** | **Fitted value** | **Min. value** | **Max. value** |
| 1 | CN2______AGRL[a] | – 0.18 | – 0.31 | 0.08 |
| 2 | CN2______CRGR[a] | – 0.67 | – 0.71 | −0.47 |
| 3 | CN2______SHRB[a] | – 0.03 | – 0.20 | 0.03 |
| 4 | CN2______SAVA[a] | 0.10 | – 0.11 | 0.37 |
| 5 | CN2______FODB[a] | – 0.53 | – 0.79 | −0.41 |
| 6 | CN2______GRAS[a] | 0.28 | 0.04 | 0.32 |
| 7 | ALPHA_BF[b] | 0.52 | 0.40 | 0.62 |
| 8 | GW_DELAY[b] | 48 | 43 | 105 |
| 9 | GWQMN[b] | 234 | 0 | 400 |
| 10 | GW_REVAP[b] | 0.13 | 0.12 | 0.17 |
| 11 | RCHRG_DP[b] | 0.82 | 0.33 | 0.87 |
| 12 | REVAPMN[b] | 305 | 303 | 381 |
| 13 | ESCO[b] | 0.45 | 0.12 | 0.45 |
| 14 | CH_K2[b] | 114 | 79 | 121 |
| 15 | SOL_Z____SANDY_CLAY_LOAM[a] | – 0.05 | – 0.07 | 0.06 |
| 16 | SOL_Z____SANDY_LOAM[a] | – 0.17 | – 0.20 | −0.03 |
| 17 | SOL_Z____SAND[a] | – 0.31 | – 0.46 | −0.30 |
| 18 | SOL_AWC____SAND[a] | – 0.20 | – 0.25 | 0.00 |
| 19 | SOL_AWC____SANDY_LOAM[a] | 0.10 | – 0.10 | 0.17 |
| 20 | SOL_AWC____SANDY_CLAY_LOAM[a] | – 0.12 | – 0.19 | −0.10 |
| 21 | SOL_K____SAND[a] | 0.28 | 0.25 | 0.71 |
| 22 | SOL_K____SANDY_LOAM[a] | – 0.11 | – 0.21 | −0.04 |
| 23 | SOL_K____SANDY_CLAY_LOAM[a] | 0.02 | – 0.16 | 0.13 |
| 24 | SURLAG[b] | 2.0 | 0.09 | 2.21 |
| 25 | HRU_SLP[a] | 0.20 | 0.13 | 0.40 |
| 26 | OV_N[a] | – 0.32 | – 0.35 | −0.07 |
| 27 | SLSUBBSN[a] | – 0.37 | – 0.39 | −0.21 |
| 28 | SOL_BD[a] | 0.39 | 0.13 | 0.51 |

[a]Existing parameter value is multiplied by (1+ a given value).
[b]Existing parameter value is to be replaced by the given value.

**Table 10** | Summary results for SWAT model calibration and validation using TMPA data

| **Criteria** | **Calibration** | **Validation** |
|---|---|---|
| $R^2$ | 0.94 | 0.69 |
| NSE | 0.94 | 0.67 |
| RSR | 0.25 | 0.57 |
| PBIAS (%) | 3.34 | 17.87 |

The overall performance of the model in the estimation of the flows above 100 m$^3$/s during the validation can be described as very good when compared with the low flows (Figures 5 and 6). This observation is as a result of the objective function used, which is more inclined towards high flows (Legates & McCabe 1999; Moriasi et al. 2007). However, the impact of the contribution of

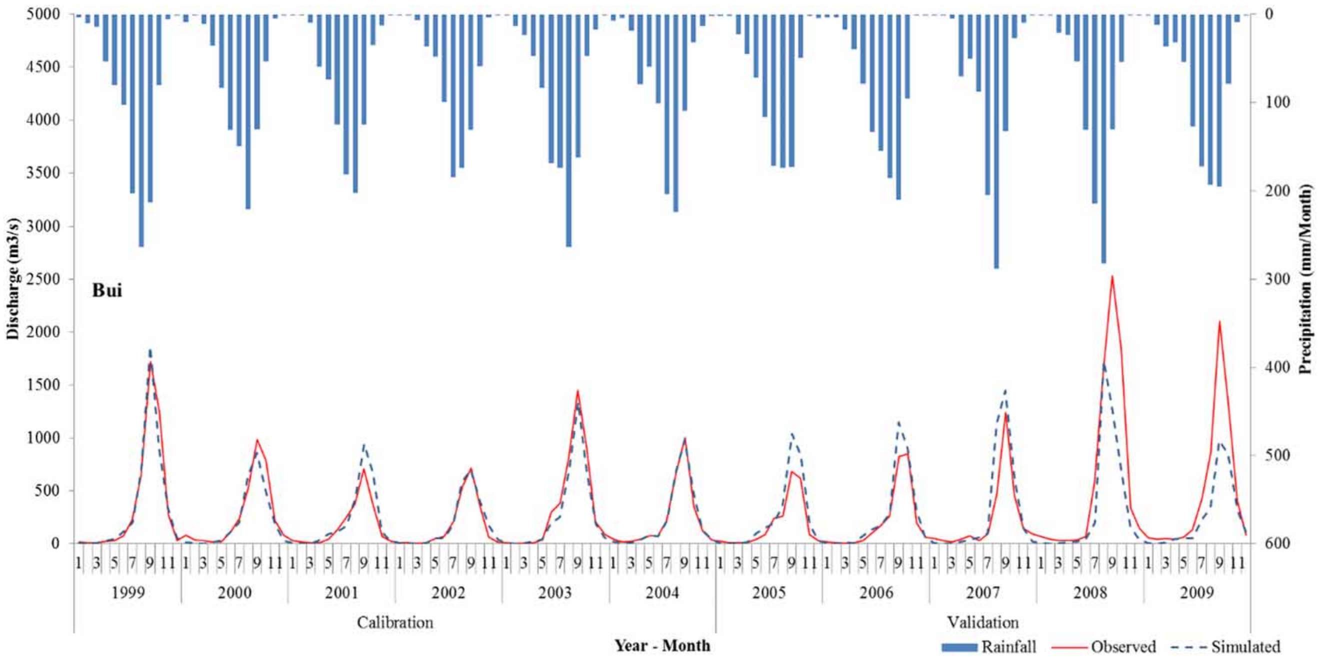


**Figure 4** | A plot of monthly calibration and validation results for the Bui station.

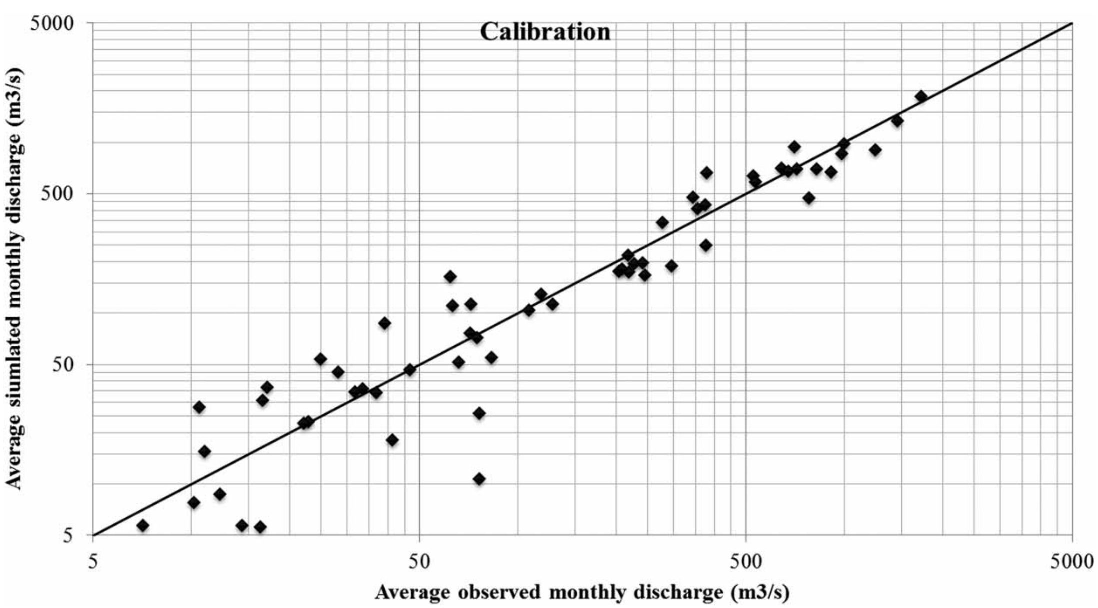


**Figure 5** | Scatter plot of average observed monthly discharge against average simulated monthly discharge for the Bui station during calibration.

the low flows to the total streamflow compared to that of the high flows is less significant. Thus the use of NSE as the main evaluation criterion is still appropriate. The representation of groundwater flow in SWAT is simplified; hence, the appropriate calibration of low flows is, sometimes, a challenge for basins, which have significant periods where flows are mainly as a result of groundwater contribution.

As can be observed in Figures 4 and 6, the model is not good at simulating very high flows. Therefore, flows above

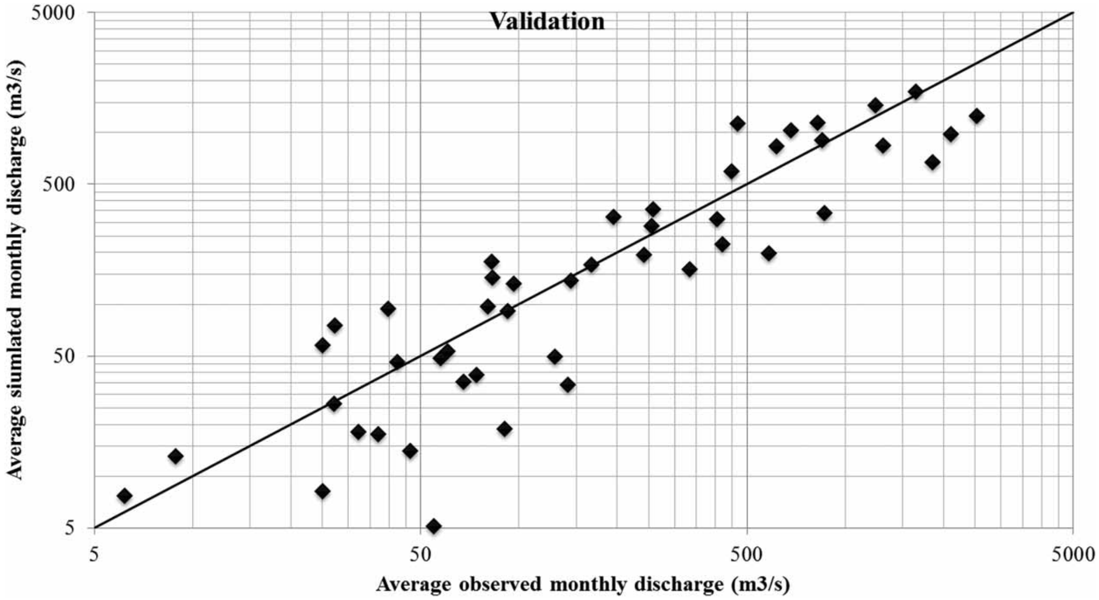


**Figure 6** | Scatter plot of average observed monthly discharge against average simulated monthly discharge for the Bui station during validation.

2,000 $m^3$/s, which were two points out of the 60 points used for the validation, were excluded in the computation of the performance indicators to obtain the following: $R^2$ (0.71), RSR (0.55), PBIAS (7.69) and NSE (0.70). Marginal improvements were observed in all the performance indicators except the PBIAS, which showed significant improvement from 17.87 to 7.69%. The changes observed in the indicators show that the model performance in the simulation of the overall flow is *good*. Hence it can be used for assessment of catchment response in cases where interest is not in the simulation of very high flows.

## CONCLUSION

This study evaluated how well satellite-derived rainfall estimates compare with ground-measured rainfall. It also assessed the potential of using the satellite-derived rainfall together with other free global data sets as inputs into SWAT for runoff simulation in the Black Volta basin, which is a large-scale trans-boundary basin in a data-scarce region.

Analysis of TMPA data set revealed unsatisfactory daily correlation between the rain gauge measurements and the satellite-derived rain estimates. However, there was an appreciable improvement in the correlations when the monthly accumulations were used. The correlations could be described as 'very good', with monthly correlation coefficients ranging from 0.85 to 0.92. On the other hand, analysis of the annual rainfall for the stations showed that the TMPA both overestimated and underestimated the rain gauge measurements at all stations in the various years with no clear trend observed.

The use of the satellite-derived rainfall estimates as input into the SWAT model for the basin indicated that a model set-up using the data would not be good for flood analysis since its peak discharge estimates were not very good. However, the model showed indications of 'very good' to 'satisfactory' performance according to the PBIAS and the RSR at the downstream station. Also, the model simulation of the overall streamflow was 'good' to 'very good' in terms of trends and magnitude (R and NSE). Therefore, the use of satellite data in the study basin for water resources assessment would be of great benefit considering the difficulties in accessing data both locally and from the other riparian countries. The methods applied in this study can also be replicated in other data-scarce basins across the globe for hydrological studies.

This study had some limitations, which included our inability to access ground meteorological measurements across the whole basin. Also, some of the discharge data obtained had some reliability issues during verification and cross-validation and, thus, only one station which was found to be reliable was used for the model calibration and validation. These limitations made it difficult to set up a model solely based on ground measurements for comparison with that of the satellite-based model.

One of the ways which can greatly help in the improvement of satellite-derived rainfall estimates, to enable its direct application in hydrological modelling, is to encourage a great deal more localized validation of the rainfall estimates. This will help in the assessment of the performance of the estimates in different climatic zones and regions. This will also lead to the development of more refined algorithms incorporating more localized data and conditions leading to better estimates.

## ACKNOWLEDGEMENTS

This work was supported by the Innovative Research Team Project under Grant No. 2009585412 by Basic Research Funds for National University at State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering and also by the Ministry of Education and State Administration of Foreign Experts Affairs, P. R. China through the 111 Project under Grant No. B08048. The images and data used in this study were acquired using the Goddard Earth Sciences (GES)-Data and Information Services Center (DISC) Interactive On-line Visualization and Analysis Infra-structure (Giovanni) as part of NASA's GES DISC. We are also grateful to the GMA for providing us with the rainfall data and the European Space Agency TIGER II Capacity Building Facility for Africa for supporting this project through training and data facilitation.

## REFERENCES

Abbaspour, K. C. 2011 *SWAT-CUP4: SWAT Calibration and Uncertainty Programs – A User Manual*. Swiss Federal Institute of Aquatic Science and Technology, Eawag, Duebendorf, Switzerland.

Abbaspour, K. C., Johnson, C. A. & van Genuchten, M. T. 2004 Estimating uncertain flow and transport parameters using a sequential uncertainty fitting procedure. *Vadose Zone J.* **3**, 1340.

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

Adjei, K. A., Ren, L., Appiah-Adjei, E. K., Kankam-Yeboah, K. & Agyapong, A. A. 2012 Validation of TRMM data in the Black Volta Basin of Ghana. *J. Hydrol. Eng.* **17**, 647–654.

Adler, R., Braun, S., Stocker, E. & Marius, J. 2007 *Tropical Rainfall Measuring Mission, TRMM, Senior Review Proposal*. Laboratory for Atmospheres, NASA Goddard Space Flight Centre, Greenbelt, MD.

Arnold, J. G., Srinivasan, R., Muttiah, R. S. & Williams, J. R. 1998 Large area hydrologic modeling and assessment part I: model development 1. *J. Am. Water Resour. Ass.* **34**, 73–89.

Barry, B., Obuobie, E., Andreini, M., Andah, W. & Pluqet, M. 2005 *The Volta River Basin: Comprehensive Assessment of Water Management in Agriculture Comparative Study on River Basin Development and Management*. International Water Management Institute (IWMI), Accra.

Batjes, N. H. 1997 A world dataset of derived soil properties by FAO–UNESCO soil unit for global modelling. *Soil Use and Management* **13**, 9–16.

Bontemps, S., Defourny, P., Bogaert, E., Van, Arino, O., Kalogirou, V. & Perez, J. R. 2011 GLOBCOVER 2009 Products Report Description and Validation Report.

Fadil, A., Hassan, R., Kaoukaya, A., Kharchaf, Y. & Bachir, O. A. 2011 Hydrologic modeling of the Bouregreg watershed (Morocco) using GIS and SWAT model. *J. Geogr. Inform. Syst.* **3**, 279–289.

Fang, H., Beaudoing, H. K., Rodell, M., Teng, W. L. & Vollmer, B. E. 2009 Global Land Data Assimilation System (GLDAS) Products, Services and Application from NASA Hydrology Data and Information Services Center (HDISC). In: *ASPRS 2009 Annual Conference. American Society of Photogrammetry and Remote Sensing*, Baltimore, MD.

FAO 1997 FAO Land and Water Bulletin, 4th edn. *Irrigation potential in Africa: A basin approach.* Food and Agriculture Organization of the United Nations. Rome.

Franks, S. W. 2005 *Predictions in Ungauged Basins: International Perspectives on the State of the Art and Pathways Forward*. IAHS Publication Series. IAHS Press, Wallingford.

Global Water Partnership 2012 Volta basin. http://www.gwp.org/en/WACDEP/IMPLEMENTATION/Where/Volta/ (accessed 4.23.12).

GLOWA Volta Project (GVP) 2007 *Partners' Capacity Needs Assessment Workshop*. Accra.

Güngör, Ö. & Göncü, S. 2013 Application of the soil and water assessment tool model on the Lower Porsuk Stream Watershed. *Hydrol. Process.* **27**, 453–466.

Hrachowitz, M., Savenije, H. H. G., Blöschl, G., McDonnell, J. J., Sivapalan, M., Pomeroy, J. W., Arheimer, B., Blume, T., Clark, M. P., Ehret, U., Fenicia, F., Freer, J. E., Gelfan, A., Gupta, H. V., Hughes, D. A., Hut, R. W., Montanari, A., Pande, S., Tetzlaff, D., Troch, P. A., Uhlenbrook, S., Wagener, T., Winsemius, H. C., Woods, R. A., Zehe, E. & Cudennec, C. 2013 A decade of predictions in ungauged basins (PUB) – a review. *Hydrol. Sci. J.* **58**, 1198–1255.

Islam, M. & Uyeda, H. 2007 Use of TRMM in determining the climatic characteristics of rainfall over Bangladesh. *Remote Sens. Environ.* **108**, 264–276.

Jung, G., Wagner, S. & Kunstmann, H. 2012 Joint climate–hydrology modeling: an impact study for the data-sparse environment of the Volta Basin in West Africa. *Hydrology Res.* **43** (3), 231.

Kundzewicz, Z. W. 2007 Prediction in ungauged basins – a systemic perspective. In: *Predictions in Ungauged Basins*. IAHS Publications, Wallingford, 309, pp. 38–44.

Legates, D. R. & McCabe, G. J. 1999 Evaluating the use of 'goodness-of-fit' measures in hydrologic and hydroclimatic model validation. *Water Resour. Res.* **35**, 233–241.

Li, L., Ngongondo, C. S., Xu, C. & Gong, L. 2013 Comparison of the global TRMM and WFD precipitation datasets in driving a large-scale hydrological model in southern Africa. *Hydrol. Res.* **44**, 770.

Li, X., Zhang, Q. & Xu, C. 2012 Suitability of the TRMM satellite rainfalls in driving a distributed hydrological model for water balance computations in Xinjiang catchment, Poyang lake basin. *J. Hydrol.* **426–427**, 28–38.

Mango, L. M., Melesse, A. M., McClain, M. E., Gann, D. & Setegn, S. G. 2011 Land use and climate change impacts on the hydrology of the upper Mara River Basin, Kenya: results of a modeling study to support better resource management. *Hydrol. Earth Syst. Sci.* **15**, 2245–2258.

McKay, M. D., Beckman, R. J. & Conover, W. J. 1979 Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* **21**, 239–245.

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. *Trans. ASABE* **50**, 885–900.

Neitsch, S. L., Arnold, J. G., Kiniry, J. R. & Williams, J. R. 2011 *Soil & Water Assessment Tool Theoretical Documentation Version 2009*. Texas Water Resources Institute Technical Report No. 406. College Station, TX.

Nie, W., Yuan, Y., Kepner, W., Nash, M. S., Jackson, M. & Erickson, C. 2011 Assessing impacts of landuse and landcover changes on hydrology for the upper San Pedro watershed. *J. Hydrol.* **407**, 105–114.

Obiero, J. P. O., Hassan, M. A. & Gumbe, L. O. M. 2011 Modelling of streamflow of a catchment in Kenya. *J. Water Resour. Protect.* **3** (9), 667–677.

Raju, K. S. 2006 *Predictions In Ungauged Basins For Sustainable Water Resources Planning And Management*. Jain Brothers, New Delhi.

Rodgers, C., Vlek, P. L. G., Eguavoen, I. & Amtz, C. (eds) 2007 *GLOWA Volta Phase II* Completion Report. GLOWA Volta, Bonn.

Rouhani, H., Willems, P. & Feyen, J. 2009 Effect of watershed delineation and areal rainfall distribution on runoff prediction using the SWAT model. *Hydrol. Res.* **40**, 505–519.

Salako, F. K. 2007 Temporal variation of rainfall erosivity in southern Nigeria. *ASSET* **7**, 190–202.

Schuol, J. & Abbaspour, K. C. 2006 Calibration and uncertainty issues of a hydrological model (SWAT) applied to West Africa. *Adv. Geosci.* **9**, 137–143.

Schuol, J., Abbaspour, K. C., Srinivasan, R. & Yang, H. 2008 Estimation of freshwater availability in the West African sub-continent using the SWAT hydrologic model. *J. Hydrol.* **352**, 30–49.

Sivapalan, M. (ed.) 2006 *Predictions in Ungauged Basins: Promise and Progress*. IAHS Publication No. 303, IAHS Press, Wallingford.

Sivapalan, M., Takeuchi, K., Franks, S. W., Gupta, V. K., Karambiri, H., Lakshmi, V., Liang, X., Mcdonnell, J. J., Mendiondo, E. M. & Connell, P. E. O. 2003 IAHS Decade on Predictions in Ungauged Basins (PUB), 2003–2012: Shaping an exciting future for the hydrological sciences. *Hydrol. Sci. J.* **48**, 857–880.

Taylor, J. C., Van de Giesen, N. & Steenhuis, T. S. 2007 West Africa: Volta discharge data quality assessment and use. *J. Am. Water Resour. Ass.* **42**, 1113–1126.

Tobin, K. J. & Bennett, M. E. 2009 Using SWAT to model streamflow in two river basins with ground and satellite precipitation data. *J. Am. Water Resour. Ass.* **45**, 253–271.

US Geological Survey 2012 Earth Explorer. http://earthexplorer.usgs.gov/ (accessed 1.25.12).

USGS 2012 FEWS NET Africa Data Portal. http://earlywarning.usgs.gov/fews/africa/index.php (accessed 2.1.13).

Van de Giesen, N., Andreini, M., Van Edig, A. & Vlek, P. 2001 Competition for water resources of the Volta basin. In: *Regional Management of Water Resources*. IAHS Press, Wallingford, pp. 199–206.

Van de Giesen, N., Stomph, T.-J., Ajayi, A. E. & Bagayoko, F. 2011 Scale effects in Hortonian surface runoff on agricultural slopes in West Africa: Field data and models. *Agric. Ecosyst. Environ.* **142**, 95–101.

Yong, B., Hong, Y., Ren, L., Gourley, J. J., Huffman, G. J., Chen, X., Wang, W. & Khan, S. I. 2012 Assessment of evolving TRMM-based multisatellite real-time precipitation estimation methods and their impacts on hydrologic prediction in a high latitude basin. *J. Geophys. Res.* **117**, D09108.

First received 18 June 2013; accepted in revised form 16 July 2014. Available online 26 August 2014