Flood forecasting in Niger-Benue basin using satellite and quantitative precipitation forecast data

Alemseged Tamiru Haile a,∗, Fekadu Teshome Tefera b, Tom Rientjes c

a International Water Management Institute, P. O. Box 5689, Addis Ababa, Ethiopia
b Arba Minch University, P.O. Box 21, Arba Minch, Ethiopia
c Department of Water Resources, Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Hengelosestraat 99, Enschede 7514 AE, The Netherlands

Article history:
Received 22 March 2016
Received in revised form 20 June 2016
Accepted 29 June 2016

Keywords:
Benue river
HEC-HMS
Flood forecasting
Satellite rainfall product
Precipitation forecasts
Flood forecast skill

Abstract
Availability of reliable, timely and accurate rainfall data is constraining the establishment of flood forecasting and early warning systems in many parts of Africa. We evaluated the potential of satellite and weather forecast data as input to a parsimonious flood forecasting model to provide information for flood early warning in the central part of Nigeria. We calibrated the HEC-HMS rainfall-runoff model using rainfall data from post real time Tropical Rainfall Measuring Mission (TRMM) Multi satellite Precipitation Analysis product (TMPA). Real time TMPA satellite rainfall estimates and European Centre for Medium-Range Weather Forecasts (ECMWFW) rainfall products were tested for flood forecasting. The implication of removing the systematic errors of the satellite rainfall estimates (SREs) was explored. Performance of the rainfall-runoff model was assessed using visual inspection of simulated and observed hydrographs and a set of performance indicators. The forecast skill was assessed for 1–6 days lead time using categorical verification statistics such as Probability Of Detection (POD), Frequency Of Hit (FOH) and Frequency Of Miss (FOM). The model performance satisfactorily reproduced the pattern and volume of the observed stream flow hydrograph of Benue River. Overall, our results show that SREs and rainfall forecasts from weather models have great potential to serve as model inputs for real-time flood forecasting in data scarce areas. For these data to receive application in African transboundary basins, we suggest (i) removing their systematic error to further improve flood forecast skill; (ii) improving rainfall forecasts; and (iii) improving data sharing between riparian countries.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

The magnitude and frequency of heavy rainfall events have significantly increased in some parts of Nigeria during the period 1980–2010 (Babatolu et al., 2014). The recent flood event in 2012 was one of the most extreme events in Nigeria in terms of its magnitude and impact on lives and properties. This event was mainly attributed to unusually heavy and prolonged rainfall (Ologunorisa and Adeyemo, 2005) while impacts were exacerbated by upstream countries’ reluctance to share timely and accurate reservoir release data. It caused a widespread devastating flood disaster that hit about 14 states bordering the Niger and Benue River. In Edo state only, the 2012 flood has affected twenty communities with a population of over 500,000 persons, destroyed infrastructure, housing and agricultural production (Agbonkhese et al., 2014; Aderoju et al., 2014). The flood also submerged most drinking water sources and destroyed aquatic life (Mmom and Aifeshi, 2013).

Unwarranted floods cannot be prevented but their impacts can be reduced for instance by implementing preventive measures. Construction of large structures (e.g. reservoirs or dykes) may be considered effective to mitigate and reduce flood impacts. However, these structures enhance false sense of security encouraging economic development in flood prone areas. Preventive measures at household level include constructing flood proofing buildings and elevating house floors (e.g. Haile et al., 2013a). These measures can reduce flood damage to buildings and their contents by only 25–55% (Bubeck et al., 2012; Kreibich et al., 2005; Kreibich and Thieken, 2009).

When preventive measures are not sufficient, flood damage can still be reduced through raised preparedness (e.g. Thielen et al., 2009). This requires issuing flood forecasts and translating these forecasts into valuable early warning information. Few hours of lag
time between occurrence of rainfall and flood peaks can be translated to substantial forecast lead time which renders an important preparation time to protect life and property. Hence, flood forecasting is one of the most effective flood risk management measures (UNISDR, 2004). There is likely a substantial monetary benefit in transboundary flood early warning system (Pappenberger et al., 2015). Despite its obvious importance, establishment of flood early warning system (FEWS) has not received the attention it deserves in many African countries.

Barriers of flood forecasting in Africa include lack of adequate data, low modeling capacity and lack of political agreement to share hydrologic data (Sobowale and Oyedepo, 2013). In particular, hydro-meteorological observation networks have poor density and coverage while the data from these networks often is unreliable, inconsistent and incomplete. There is also considerable lag time between the actual measurement and transfer time of these data to the respective meteorological offices limiting the data application for flood forecasting.

Satellite Rainfall Estimates (SREs) provide easy access to global data including data in transboundary river basins; potential for increased spatio-temporal frequency of sampling; and uninterrupted supply of rainfall data during catastrophic situations (Harris et al., 2007). However, SREs are often associated with large systematic and random errors (Haile et al., 2013b; Habib et al., 2012). Particularly, the biases (systematic errors) can show a complex dependency, in terms of magnitude and sign, on topography and latitude. Hence, a bias adjustment of SREs is crucial before these estimates receive application in flood forecasting (Habib et al., 2014). Main limitation is that SREs only offer real time or near real time data while future rainfall amounts are not forecasted.

Medium range weather forecasts produce rainfall forecasts up to 15 days in advance. A number of studies evaluated effectiveness of these weather forecasts for flood warning in hydrological basins of western countries with less attention to African basins. Studies reported that the weather forecasts better capture precipitation occurrence than magnitude and location of peaks (McBridge and Ebert, 2000); forecasts are most reliable for long lead time (5–6 days) for temperature and short lead time (2–3 days) for precipitation (Buzzi et al., 1999); forecast performance deteriorates with increasing lead time (e.g. Renner et al., 2009; Bennett et al., 2014); the lead time can be increased by using ensemble predictions of precipitation and mixing forecasts from different models or centers (e.g. Candille, 2009). These forecasts have poor spatial resolution hindering their application in comparatively small catchments which raise the need to apply disaggregating techniques to represent spatial precipitation variability (He et al., 2009).

Evaluation of the European Centre for Medium-Range Weather Forecasts (ECMWF) ensemble forecasts for flood forecasting outside Africa indicates systematic deviations from observed river flows (e.g. Pappenberger et al., 2005; Bartholmes and Todini, 2005). In this study, we evaluated flood forecast skill using SREs (post real time and real time data) and ECMWF rainfall forecasts as a model input in a data scarce area of the Benue basin. The forecast skill is evaluated for historic disastrous floods in central Nigeria for 1–6 days forecast window (lead time). For this purpose, the HEC-HMS rainfall-runoff model was set-up and calibrated for historical river flow. Input to the rainfall-runoff model was obtained from SREs and rainfall forecast from a numerical weather prediction model. HEC-HMS was also evaluated in flood forecasting mode. Three major components are identified in our approach: evaluate and bias correct rainfall products; set-up and evaluate flood forecasting and suggest the way forward. We believe that our study can contribute to efforts towards increasing preparedness of people, reduction of flood impacts and relief. Findings will directly serve Kogi state but also Benue and Edo which border the Benue River.

2. Study area

This study focuses on the Benue sub-basin which is one of the major tributaries of Niger River. The Benue sub-basin is located between 6° 10′ 0" to 13° 0′ 15" N latitudes and 9° 46′ 41" to 16° 0′ E longitudes. Its surface area is estimated as 918,872 km2 partially or fully covering about 8 of the 36 states of Nigeria (Fig. 1). With its headwaters in the Adamawa Plateau of the Northern Cameroon, the river first drains along the west direction in Cameroon until it crosses the Nigerian border. Then it drains to the south-west in Nigeria and then changes direction towards west until it joins the major Niger River at Lokoja in east-central part of Nigeria.

The rainfall, potential evapotranspiration (PET) and runoff of the sub-basin show considerable intra-annual variations. Analysis of SREs (TRMM 3B42) over 1998–2012 time period shows that Benue’s mean monthly rainfall peaks in August (257 mm) and attains smallest magnitude in December (1.24 mm). From 1998–2012, the highest mean monthly PET of the sub-basin is 167 mm (March) with 109 mm (August) as lowest PET. And also the monthly stream flow at Makurdi (averaged over the same time period) ranges between 96.1 mm in October and 2.5 mm in April. This shows the peakflow is attained two months after the rainfall peak.


3. Data sets and methods

3.1. Data sets

Tropical Rainfall Measuring Mission (TRMM) is a joint USA/Japan satellite mission designed to survey the rain structure, rate and distribution in tropical and subtropical regions (latitude range ±50°). In this study, we used the TRMM Multi satellite Precipitation Analysis product (TMPA-3B42—Huffman et al., 2007). This based on preliminary assessment of literature and data latency (time gap between data acquisition and delivery) which is a crucial criterion for flood forecasting. The post real time TRMM-3B42 product of TMPA does not have much application in flood forecasting as it is made available to users several weeks (37 days) after the satellite observations are acquired. This substantial latency is as a result of the need to incorporate gauge observations. However, the real time product of TMPA (3B42RT) does not require significant post processing. As a result, it is made available to users within 6–9 h of data acquisition.

The post real time SRE was acquired at daily time step for the period since 1998 from: ftp://disc2.nascom.nasa.gov/data/TRMM/Grided/Derived_Products/3B42_V6/Daily The daily real time SRE was downloaded from: http://gdata1.sci.gsfc.nasa.gov/daacbin/G3/gui.cgi?instance_id=TRMM-3B42RT-Daily. The real time data was made available since 2000. Both data sets of TRMM have 0.25°X0.25° spatial resolution.

We also evaluated the European Center of Medium-range Weather Forecasting (ECMWF) data. The daily precipitation forecast of ECMWF was obtained from the THORPEX Interactive Global Grand Ensemble (TIGGE) data archive at: http://apps.ecmwf.int/datasets/data/tigge/levtype=sfc/type=cf in GRIB data format. We acquired this data for the period between 1st of August and 12th of September 2008 and 2012 since flooding intensified over the study area during this period. The rainfall forecast data for 1–6 days lead
time and a spatial resolution of 0.25° X 0.25° (to be consistent with the satellite data) was also used as input to HEC-HMS model in forecasting mode.

Data for estimating potential evapotranspiration (PET) was obtained from the National Climate Data Center (NCDC). These data include monthly and minimum temperature (24 stations) and wind speed (21 stations). Thiessen polygon method was used to convert the station PET values to sub-basin values.

We obtained stream flow data for some gauging stations along Benue River from Nigeria Hydrological Services Agency (NIHSA). However, only one of these gauging stations (Makurdi) was used in this study since records of the other stations contain several missing data. For the analysis period (1998–2012), the time series of flow data at Makurdi contains missing data in most days of the years 2001, 2002, 2010 and 2011.

Digital elevation model (DEM) was obtained from the Shuttle Radar Topography Mission (SRTM). It has a spatial resolution of 90 m by 90 m and can be acquired from http://www.cgiar-csi.org/data/srtm-90m-digital-elevation-database-v4-1

The land cover data was obtained from the Food and Agriculture Organization of the United Nations (UN-FAO) land use dataset: http://www.fao.org/news/story/en/item/216144/icode/. The land cover data was produced on 02 January 2009. The FAO HWSD (Harmonized world soil data) data was downloaded from http://www.fao.org/geonetwork. Both the land cover and soil data have equal spatial resolution of 30 arc seconds (approximately 1 km by km). The characteristics of Lagdo reservoir were obtained from literature.

3.2. Bias correction of satellite rainfall data

In this study, the bias correction factors (BF) were estimated as a ratio of satellite and rain gauge data over each of the rainy months (June to September) for the period 2000–2011. This approach was adopted after Habib et al. (2014). About 11 rainfall stations, which are distributed over the study area, were used to estimate BFIs separately for the real time and post real time SREs. The BFIs values were then converted to sub-basin values using Thiessen polygon method. Bias correction was applied by multiplying the post real time and real time SREs of each sub-basin by the corresponding BF. Note that bias correction was not done for dry season rainfall amount due to its small magnitude and floods occur only in the rainy season.

3.3. Set-up and evaluate flood forecasting model

HEC-GeoHMS software which is an extension of ArcGIS software was used to derive the river network of Benue basin, to delineate watershed boundaries and to extract other watershed characteristics. The basin is divided into 15 sub-basins based on location of gauging stations, Lagdo dam, and sub-basin size (Fig. 2). The largest sub-basin has a drainage area of 39,848 km² (which is sub basin W590) and the smallest sub-basin is 2145 km² (which is sub basin W1000). However, most of the sub-basins have comparable surface area as shown in Fig. 2.

We used the post real time SREs as input for HEC-HMS model calibration and validation. We compared calibration and validation results with and without bias correction of the SREs. The charac-
teristics of the rainfall product are presented in Section 3.1 of this manuscript.

The inflow hydrograph of a reservoir is translated and attenuated as it leaves the reservoir. This affects the magnitude and timing of flood peaks over floodplains further downstream of the reservoir. As a result, the characteristics of a reservoir and its release should be carefully specified in HEC-HMS. In this study, the effect of Lagdo reservoir was modeled by specifying the time series data of reservoir release which was estimated based on simple spreadsheet model for reservoir water balance. The reservoir release is estimated as a function of reservoir inflows, outflows other than release, and reservoir characteristics affecting the relationship between reservoir storage and stage (actual water level). The inputs to the spreadsheet model are reservoir characteristics (storage capacity, reservoir stage-volume curve, reservoir full level, dead storage level, spillway characteristics), environmental releases, reservoir inflow, rainfall and evapotranspiration.

3.4. HEC-HMS model calibration and validation

We applied the deficit and constant loss method of HEC-HMS to simulate soil moisture dynamics. The method was selected as it is found easy to conceptualize and is not data demanding. It assumes that the soil has a fixed water holding capacity, typically based on the active rooting depth of vegetation. Simplifying assumptions are made regarding soil dynamics so that infiltration (approximated by saturated hydraulic conductivity) only occurs when the soil is saturated. Direct runoff was simulated using the Soil Conservation Service (SCS) unit hydrograph method which is based on a relationship between time of concentration (Tc) and lag time (Tlag). Tc can be estimated based on sub-basin characteristics including topography and reach length whereas Tlag is estimated as fraction of concentration time. Base flow was simulated using the recession method. An exponential decay of base flow is assumed in this method. Stream flow through channel reaches was routed using the Muskingum routing method which has two parameters: the K parameter simulates a delay in stream flow (in hours) as it moves through the channel and X is the parameter to weigh influence of inflow and outflow hydrograph. Attenuation is maximum and zero for X value of 0.0 and 0.5, respectively.

Model initialization is necessary to bring the antecedent moisture conditions in the model closer to the actual conditions in the study area. We initialized (warm up) HEC-HMS model of the study area using two years (1998–1999) rainfall and PET data. Then, the initialized model was calibrated for a period of 8 years (2000–2007) in which normal, wet and dry conditions prevailed. Model calibration is aimed to optimize model parameter values until specific aspects of the simulated hydrograph acceptably match with counterparts from the observed hydrograph.

We evaluated HEC-HMS model performance through visual inspection of the simulated and observed hydrographs and a set of objective functions that measure the goodness-of-fit between various aspects of the two hydrographs. The objective functions which were used in this study are Nash-Sutcliffe model efficiency (NSE) coefficient and Relative Volumetric Error (RVE) as shown below.

\[ NSE = 1 - \frac{\sum_{i=1}^{n}(Q_{obs,i} - Q_{sim,i})^2}{\sum_{i=1}^{n}(Q_{obs,i} - \bar{Q}_{obs})^2} \]  
\[ RVE = \frac{\sum_{i=1}^{n}(Q_{sim,i} - Q_{obs,i})}{\sum_{i=1}^{n}Q_{obs,i}} \times 100 \]  

where, \( Q_{obs,i} \) is the observed stream flow at time step \( i \), \( \bar{Q}_{obs} \) is the mean of the observed stream flow, \( Q_{sim,i} \) is the simulated stream flow at the time step \( i \) and \( n \) is the number of observations.

NSE is used to evaluate the model ability to reproduce the pattern of the observed hydrograph. NSE ranges between \(-\infty \) and 1 (1 inclusive), with NSE = 1 being the target value. Values between 0.6 and 1.0 are generally viewed as acceptable level of performance, whereas NSE = 0 indicates that the mean observed value is a better predictor than the simulated value, which indicates unacceptable model performance. The values of RVE range between \(-\infty \) to \(+\infty \). The model performance is good for RVE values between \(-5\% \) to \(5\% \), while values between \(-10\% \) to \(-5\% \) and \(5\% \) to \(10\% \) suggest satisfactory performance with regard to volumetric error.

3.5. Evaluation of flood forecasting skill

For forecasting floods, initial conditions in the HEC-HMS model were specified based on model simulations using bias corrected
satellite rainfall and observed evapotranspiration data. The post real time SREs serves as input for model initialization. In the absence of data due to data latency, the real time data (TRMM 3B42RT) serves as input to initialize the model in forecasting mode.

A flood level threshold serves as a reference point to issue a flood warning. In this study, the flood threshold was fixed using the Weibull plotting formula for flow magnitudes with 5%, 1% and 0.5% exceedance probabilities (Table 1). These flow thresholds were fixed to correspond to various flood severity levels (medium, severe and extreme severe, respectively) at Makurdi gauging station. The threshold is estimated separately for the observed and simulated stream flows. We note that use of simulated stream flows to set flood warning thresholds is found more effective than observed flow data (e.g. Thielen et al., 2009). In this study we present results for the medium floods.

The forecast skill can be evaluated using the contingency table as shown in Table 2 (Haile et al., 2010). The possible outcomes in the contingency table are: Hit (HIT), Miss (MISS), False alarm (FA) and Non-Event (NE). If both the observed and forecasted stream flows of a particular day exceed the specified threshold, then the forecast is considered as hit. If the forecasted flow is smaller than its threshold while the observed flow exceeds its threshold, then it is called as missed event. If the forecasted flow exceeds the threshold while the observed flow is below the threshold, then the forecast is considered as false alarm.

The forecast skill was evaluated using categorical verification statistics. The categorical verification statistics include Probability Of Detection (POD), Frequency Of Hit (FOH) and Frequency Of Miss (FOM). These statistics are estimated based on the contingency table which is shown in Table 2 and read as follows:

\[
POD = \frac{HIT}{HIT + MISS} \text{ range } [0, 1] \text{, best } : 1
\]

\[
FOH = \frac{HIT}{HIT + FA} \text{ range } [0, 1] \text{, best } : 1
\]

\[
FOM = \frac{MISS}{HIT + MISS} \text{ range } [0, 1] \text{, best } : 0
\]

POD is used as a measure of how accurately the model detects actual flood events. FOH is the ratio of the number of correctly forecasted flood events and the total number of forecasted flood events. It measures the accuracy of the forecasted flood events. FOM measures the observed flood events that are missed by the forecasts. All three categorical statistics range between 0 and 1.0. The POD and FOH values of 1.0 and FOM = 0 indicate perfect forecast skill.

In this study, the stream flow data at Makurdi station was made available at daily time-step. Therefore, the usefulness of the satellite and weather forecast data for flood forecasting at this station was evaluated for 1–6 days lead time.

### 4. Results

#### 4.1. Bias of satellite rainfall product

Both the post real time and real time version of TRMM-3B42 rainfall products contain systematic errors (biases) in the study area (Fig. 3). Bias of the real time SREs (3B42RT) vary spatially with magnitudes within 10% at most stations but reaches up to 40% at few stations. Bias of the post real time SREs (TRMM-3B42) is only slightly smaller than the real time estimates and exceeds that of the real time estimates at some stations. This suggests that the post processing to produce post real time estimates as introduced in the SREs did not lead to significant bias reduction in the study area. Though the performance of TRMM-3B42 is not as poor as reported for other basins (e.g. Haile et al., 2013b), we still needed to adjust the SREs bias before the estimates receive application for flood forecasting in the Benue sub-basin.

#### 4.2. HEC-HMS model calibration and validation

We first calibrated HEC-HMS model of the study area by iteratively changing the value of model parameters until a reasonable match is attained between simulated and observed dry season and annual stream flow volume. Next, parameter values were adjusted individually by matching pattern and peaks of simulated and observed hydrographs. Table 3 shows the parameters values which were obtained after manual calibration of HEC-HMS model. Calibrated values of all model parameters are within their allowable range. Values of some of the calibrated parameters were kept spatially uniform; i.e., their values did not vary for the sub-basins. For instance, the value of X (Muskingum routing parameter to weigh influence of inflow and outflow hydrograph) is fixed at 0.2, constant rate is 0.42 mm per hour, ratio-to peak is 0.12 m²/s, maximum deficit is 265 mm and initial discharge is 0.02 m³/km² which is estimated as a ratio of minimum dry season flow and drainage area at Makurdi station. Values of the remaining parameters (e.g. initial deficit, impervious surface, lag time and K) were allowed to spatially vary since these are related to catchment characteristics such as land cover and terrain characteristics.

Fig. 4 shows the observed and simulated stream flow hydrographs for the calibration period. The model satisfactorily reproduced the observed hydrograph pattern. There is also a good match between the rising and recession limbs of the simulated and observed stream flow hydrographs. The timing and magnitude of peak floods is well reproduced in 2004 and 2005. However, there is slight underestimation (2000, 2003 and 2007) and overestimation (2006).

Evaluation of the model performance in terms of the objective functions also shows very good model performance in line with our visual inspection. The volumetric error is very small (RVE = -1.47%) suggesting very good model performance in terms of capturing observed stream flow volume. The model performance is fair in reproducing the pattern of the observed hydrograph (NSE = 0.5). Further, we note that flood discharges commonly are difficult to observe since discharges often are based on stage-discharge rating curves that generally become inaccurate for higher river water levels. We can see that the discharge during the 2006 extreme year and
the model did not capture it well. This could be caused by changes in the rainfall–runoff relationship, by discharge observation errors or an inaccurate rating curve during periods of overflow at river banks during a flood event. The values of the objective functions provide an aggregated assessment of model performance.

The calibrated model capability to reproduce stream flow hydrograph outside the calibration period is evaluated using observed meteorological inputs over the time period 2008–2012. However, stream flow data of Makurdi station has several missing data in 2010 and 2011 (Fig. 5) which makes the evaluation questionable. Overall, the calibrated model satisfactorily reproduced the pattern of the hydrograph which is observed outside the calibration period. It also captured the observed peak flow in 2008. However, it significantly underestimated the peaks in 2009 and 2012 (extreme flood year). We note that the result may be affected by unreliable discharge observations that commonly are less reliable during extreme events. River banks are commonly overtopped during extreme floods and as a result conversion of observed river stage to stream flow data (rating curve) can be unreliable.

The values of NSE and RVE were 0.5 and −8.89% respectively for the validation period. The bias correction did not bring significant improvement in model performance in terms of the two objective

![Fig. 3. Bias of the post real time and real time TRMM-3B42 rainfall product in the rainy season (JJAS).](image)

### Table 3
The calibrated parameters set of HEC-HMS model for the Benue basin at Makurdi station.

<table>
<thead>
<tr>
<th>Sub basin Name</th>
<th>Initial Deficit (mm)</th>
<th>Imp.serf(%)</th>
<th>Lag Time(hour)</th>
<th>Reach Name</th>
<th>K(hr)</th>
</tr>
</thead>
<tbody>
<tr>
<td>W510</td>
<td>2.91</td>
<td>8</td>
<td>317</td>
<td>R130</td>
<td>150</td>
</tr>
<tr>
<td>W560</td>
<td>2.85</td>
<td>10</td>
<td>293</td>
<td>R500</td>
<td>109</td>
</tr>
<tr>
<td>W580</td>
<td>2.31</td>
<td>11</td>
<td>396</td>
<td>R80</td>
<td>131</td>
</tr>
<tr>
<td>W590</td>
<td>3.56</td>
<td>11</td>
<td>287</td>
<td>R70</td>
<td>150</td>
</tr>
<tr>
<td>W1100</td>
<td>2.57</td>
<td>1</td>
<td>287</td>
<td>R180</td>
<td>150</td>
</tr>
<tr>
<td>W1050</td>
<td>2.97</td>
<td>7</td>
<td>373</td>
<td>R200</td>
<td>119</td>
</tr>
<tr>
<td>W680</td>
<td>3.45</td>
<td>9</td>
<td>229</td>
<td>R270</td>
<td>138</td>
</tr>
<tr>
<td>W690</td>
<td>2.65</td>
<td>13</td>
<td>313</td>
<td>R1170</td>
<td>60</td>
</tr>
<tr>
<td>W700</td>
<td>2.72</td>
<td>12</td>
<td>221</td>
<td>R330</td>
<td>150</td>
</tr>
<tr>
<td>W1000</td>
<td>3.16</td>
<td>0.3</td>
<td>382</td>
<td>R430</td>
<td>150</td>
</tr>
<tr>
<td>W880</td>
<td>3.52</td>
<td>7</td>
<td>330</td>
<td>R350</td>
<td>133</td>
</tr>
<tr>
<td>W1150</td>
<td>3.45</td>
<td>12</td>
<td>255</td>
<td>R1020</td>
<td>51</td>
</tr>
<tr>
<td>W920</td>
<td>2.8</td>
<td>13</td>
<td>243</td>
<td>R380</td>
<td>51</td>
</tr>
<tr>
<td>W980</td>
<td>4.52</td>
<td>14</td>
<td>334</td>
<td></td>
<td></td>
</tr>
<tr>
<td>W1010</td>
<td>2.81</td>
<td>0.4</td>
<td>263</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

![Fig. 4. Observed and simulated daily discharge hydrograph of Benue basin at Makurdi station for the calibration period (2000–2007) using bias corrected TRMM-3B42 data as rainfall input.](image)
functions. However, these objectives functions describe the model performance as aggregated over the validation period.

The simulated stream flow in 2012 using 3B42 data with and without bias correction is shown in Fig. 6. Bias correction of the SREs did not significantly improve the accuracy of the simulated hydrograph pattern. However, there is slight improvement in terms of flow magnitude since the simulated stream flow became slightly closer to the observed discharge suggesting the importance of bias correction. Simulated stream flow can have a difference of up to 2200 m$^3$/s when simulations are based on with and without bias correction of SREs.

4.3. Graphical evaluation of forecast skill

There exists major difference between the two rainfall data sources which were used in this study. The real time SREs are made available only until the model is run for forecasting (future day rainfall amount is not available and assumed zero) whereas ECMWF provides forecasts of rainfall for future days. In Fig. 7a, the HEC–HMS model was initiated on 25th of August 2008 to issue flood forecasts from 26th of August (1-day lead time) to 31st of August 2008 (6-days lead time). As shown in the figure, the observed stream flow was below its corresponding threshold (5%) while both real time SREs (3B42RT) and ECMWF based forecasted floods were above their threshold for all lead times. This suggests the model incorrectly forecasted a flood event which did not occur in reality. The ECMWF based flood magnitude is almost equal to that of SREs for 1–3 days lead time. However, the former has slightly larger magnitudes for 4–6 days lead time. The flood forecasts based on both rainfall data sources resulted in significant overestimation of the observed flood magnitude for short lead time (1–3). However, the difference between the observed and forecasted flood magnitude became small as the lead time increased with overestimation by real time SREs and underestimation by ECMWF.

Fig. 7b shows flood forecasts after initiating the model on the 26th of August 2008. The observed discharge is below its flood threshold for 1–5 days lead time which suggests a flood did not occur in reality. However, both forecasts incorrectly forecasted a flood event for these lead times. The model accurately detected the flood event on the 1st of September 2008 (6 days lead time). Forecasted flood magnitudes using the two rainfall sources were very close to that of the observed flood for short lead time (2–4 days). For longer lead times, the ECMWF based forecast led to slight overestimation while the SREs based forecast led to underestimation.

Fig. 7c shows that the model forecasted flood occurrence for all lead times. However, flood was actually observed only on the 1st and 2nd of September (5 and 6 days lead times). The forecasted and observed flows have somewhat equal magnitudes almost for all lead times with their difference increasing with lead time.

<table>
<thead>
<tr>
<th>Source of rainfall input</th>
<th>Lead time</th>
<th>HIT</th>
<th>MISS</th>
<th>FA</th>
<th>NE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bias corrected TRMM-3B42 RT</td>
<td>1 day</td>
<td>13</td>
<td>0</td>
<td>19</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>2 day</td>
<td>14</td>
<td>1</td>
<td>19</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>3 day</td>
<td>15</td>
<td>1</td>
<td>18</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>4 day</td>
<td>16</td>
<td>0</td>
<td>18</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>5 day</td>
<td>17</td>
<td>0</td>
<td>18</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>6 day</td>
<td>18</td>
<td>0</td>
<td>17</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>1 day</td>
<td>13</td>
<td>0</td>
<td>23</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>2 day</td>
<td>14</td>
<td>0</td>
<td>23</td>
<td>6</td>
</tr>
<tr>
<td>ECMWF</td>
<td>3 day</td>
<td>15</td>
<td>0</td>
<td>23</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>4 day</td>
<td>16</td>
<td>0</td>
<td>24</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>5 day</td>
<td>17</td>
<td>0</td>
<td>24</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>6 day</td>
<td>18</td>
<td>0</td>
<td>24</td>
<td>1</td>
</tr>
</tbody>
</table>

Note: These statistics are obtained for a stream flow threshold corresponding to 5% exceedance probability (Medium flood threshold).

4.4. Evaluation of forecast skill using categorical statistic

The forecast skill was assessed between 1st of August to 12th of September of the years 2008 and 2012 (86 days in total) using categorical statistics. These statistics are presented using contingency table (Table 4) and include: hits (HIT), false alarms (FA), misses (MISS) and non-event (NE). Using the bias corrected real time SREs (3B42 RT) for 1-day lead time, 11 days out of 43 in 2008 were forecasted as no flood days as both the observed and forecasted flows were below the specified thresholds. There appears to be slightly more false alarms than hits while the model almost did not miss any event. Unexpectedly, the number of hits has slightly increased as the lead time increases while the number of false alarm decreased. The forecast skill did not significantly deteriorate with lead time when evaluated in terms of these statistics. Replacing the SREs with ECMWF as model input did not change the number of hits. However, the number of false alarms increased noticeably.

The forecast skill score for 86 days (from August 1 to September 12 of 2008 and 2012) was evaluated in terms of categorical statistics. These are Probability Of Detection (POD), Frequency Of Hits (FOH), and Frequency Of Misses (FOM). The POD values for SREs based forecasts detect about 50–60% of the observed moderate floods (Table 5). Unexpectedly, the POD values increased as the lead time increased though the increment was very small. Replacing the SREs with ECMWF rainfall led to a slight decrease in POD and an increase in FOM which implies increased miss of actually occurred floods. Values of FOH were very close to 1.0 for both data sets.

5. Conclusion

Reliable forecasting tools provide an important input to early warning which increases preparedness of floodplain residents for
Fig. 6. Comparison of observed and simulated (with and without bias correction of SREs) hydrograph of Benue basin at Makurdi station for the 2012 flood season.

Fig. 7. Observed and forecasted discharge using the TRMM-3B42 RT rainfall data (after bias correction) and ECMWF rainfall data for medium flood threshold for 1–6 days lead time for 2008 at Makurdi station. The forecast is initialized on 25-August (a), 25-August (b) and 27-August (c).
climate extremes. Forecast information also helps to avoid or minimize flood damage and enable farmers to produce food using recession crop cultivation under uncertain climate.

In the present study, a parsimonious rainfall-runoff model was set-up and successfully implemented for a data scarce basin which is shared by Cameroon and Nigeria. The hydro-meteorological data were obtained from different sources which included gauge, satellite, and weather forecast data. However, the gauge data was very limited due to lack of data sharing practice in the region. To overcome the problem of data availability, we mainly relied on satellite rainfall estimate (SRE) and rainfall forecasts from a weather model. The post real time SREs (TRMM-3B42) was used for model calibration and validation while the real time SREs (TRMM-3B42 RT) and the weather forecast (ECMWF) rainfall data were used to evaluate forecast skill.

Our results suggest the importance of bias correction if SREs need to receive application in flood forecasting. However, improved data sharing practice (e.g. reservoir release and rainfall data) is needed in the basin to further improve the bias correction. NEC-HMS model still reasonably captured the pattern and volume of the observed stream flow hydrograph for normal, dry and wet (flood) years. An exceptional poor model performance was found for the 2012 flood which is the worst in terms of magnitude and impact in the study area. The river stage – discharge rating curve used in this study could not be validated and may have affected assessment results.

Graphical inspection and objective functions based evaluation shows that use of SREs and weather forecast data as model input enable us to issue flood forecasts in the Benue basin for 1–6 days lead time. The best forecast skill was achieved in detecting flood occurrence as compared to estimating flood magnitude. Both data sets resulted in detection of 50–60% of the actual flood events.

Our evaluation results did not favor selection of one of the two rainfall data sets (3B42 RT or ECMWF). ECMWF based flood forecasts have slightly more false alarms than SRE based forecasts. However, the former has a slightly better detection capability than SREs. Forecast skill using both data sources did not significantly deteriorate with increasing lead time. For some instances, better forecast skill was attained for long lead time. Future studies should evaluate the importance of correcting the systematic error in the ECMWF rainfall product and added value of probabilistic forecasts. Considering the various sources of uncertainties, the forecast results obtained in this study are encouraging.

The TRMM based rainfall products will be replaced by the Global Precipitation Mission (GPM) rainfall products after mid-2016. It is expected that there will not be significant difference in terms of the rainfall accuracy. However, it is worth evaluating the new GPM rainfall data in the future when it becomes readily available for the study area before it is used for forecasting.

Acknowledgements

This work was supported by the project: Turning Tragedy into Opportunity: Water Management Solutions for Flood Recession and Dry Season Farming in Nigeria. The project was led by the International Water Management Institute (IWMI).

References