Nat Hazards (2009) 50:109–123
DOI 10.1007/s11069-008-9324-5

ORIGINAL PAPER

Evaluation of the real-time TRMM-based multi-satellite precipitation analysis for an operational flood prediction system in Nzoia Basin, Lake Victoria, Africa

Li Li · Yang Hong · Jiahu Wang · Robert F. Adler · Frederick S. Policelli · Shahid Habib · Daniel Irwn · Tesfaye Korme · Lawrence Okello

Received: 23 September 2008 / Accepted: 11 November 2008 / Published online: 28 November 2008

Abstract Many researchers seek to take advantage of the recently available and virtually uninterrupted supply of satellite-based rainfall information as an alternative and supplement to the ground-based observations in order to implement a cost-effective flood prediction in many under-gauged regions around the world. Recently, NASA Applied Science Program has partnered with USAID and African-RCMRD to implement an operational water-hazard warning system, SERVIR-Africa. The ultimate goal of the project is to build up disaster management capacity in East Africa by providing local governmental officials and international aid organizations a practical decision-support tool in order to better assess emerging flood impacts and to quantify spatial extent of flood risk, as well as to respond to such flood emergencies more expediently. The objective of this article is to evaluate the applicability of integrating NASA's standard satellite precipitation product with a flood prediction model for disaster management in Nzoia, sub-basin of Lake Victoria, Africa. This research first evaluated the TMPA real-time rainfall data against gauged rainfall data from the year 2002 through 2006. Then, the gridded Xinanjiang Model was calibrated to Nzoia basin for period of 1985–2006. Benchmark streamflow simulations were produced with the calibrated hydrological model using the rain gauge and observed

L. Li · Y. Hong (✉) · J. Wang
School of Civil Engineering and Environmental Sciences, Center for Natural Hazard and Disaster Center, National Weather Center, University of Oklahoma, 202 W Boyd ET, CEC #334, Norman, OK 73019, USA
e-mail: yanghong@ou.edu; yanghong@agnes.gsfc.nasa.gov

R. F. Adler · F. S. Policelli · S. Habib
NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA

R. F. Adler
Earth System Science Interdisciplinary Center, University of Maryland, College Park, MD 20740, USA

D. Irwn
NASA Marshall Space Flight Center, Huntsville, AL 35812, USA

T. Korme · L. Okello
African Regional Centre for Mapping of Resources for Development (RCMRD), Nairobi, Kenya

streamflow data. Afterward, continuous discharge predictions forced by TMPA 3B42RT real-time data from 2002 through 2006 were simulated, and acceptable results were obtained in comparison with the benchmark performance according to the designated statistic indices such as bias ratio (20%) and NSCE (0.67). Moreover, it is identified that the flood prediction results were improved with systematically bias-corrected TMPA rainfall data with less bias (3.6%) and higher NSCE (0.71). Although the results justify to suggest to us that TMPA real-time data can be acceptably used to drive hydrological models for flood prediction purpose in Nzoia basin, continuous progress in space-borne rainfall estimation technology toward higher accuracy and higher spatial resolution is highly appreciated. Finally, it is also highly recommended that to increase flood forecasting lead time, more reliable and more accurate short- or medium-range quantitative precipitation forecasts is a must.

Keywords Flood prediction · Remote sensing precipitation · TRMM · Capacity building

1 Introduction

Precipitation-triggered floods are among the most devastating natural hazards and one of the most prevalent hydro-meteorological disasters in the world. With the increase in precipitation intensity, flood magnitude and frequency may increase in flood-prone regions of the world (IPCC 2007; http://www.ipcc.ch/). In addition to other factors, ever increasing settlements in vulnerable areas with increasing population and resource utilization will certainly increase the risks of floods in the near future. Such threat is going to affect the vulnerable nations the most. However, advances in flood monitoring/forecasting have been constrained by the difficulty in precisely estimating precipitation, the key forcing factor, over a range of spatial scales (point-, basin-, region-, or even global) and continuously over the time (daily to sub-daily). The recently available and virtually uninterrupted supply of satellite-based rainfall estimates is increasingly becoming a cost-effective source of data for flood prediction in many under-gauged regions around the world. The National Aeronautics and Space Administration (NASA) pioneered the development of a near-real-time global flood forecasting system for early warning, the NASA Goddard Space Flight Center (GSFC) internet-based Global Flood Monitoring (Hong et al. 2007a, b) system initially used by the Regional Visualization and Monitoring System (SERVIR) for Mesoamerica (www.servir.net). After the successful launch and operation of the SERVIR-Mesoamerica, the NASA Applied Science program has again partnered with the United States Agency for International Development (USAID) and The Africa Regional Centre for Mapping of Resources for Development (RCMRD) to implement an operational flood warning system as part of the SERVIR-Africa project. The project seeks to take advantage of spatially continuous remote sensing information as an alternative and supplement to ground-based observation in order to provide better information about the spatial extent of flood hazards.

The objective of this research is to evaluate the applicability of integrating NASA's standard precipitation product, TRMM-based Multi-satellite Precipitation Analysis 3B42 Real-Time (TMPA 3B42RT) rainfall data, with a flood prediction system for disaster management in Nzoia, sub-basin of Lake Victoria, Africa. The key datasets enabling the development of a distributed hydrological model in Africa include TMPA 3B42RT Real-time rainfall estimates, the digital elevation data from the NASA Shuttle Radar

Topography Mission (SRTM), hydrological parameter files from **Hydro**logical data and maps based on **SH**uttle **E**levation **D**erivatives at multiple **S**cales (HydroSHEDS), the Moderate Resolution Imaging Spectroradiometer (MODIS) land cover, and soil parameters provided by Food & Agricultural Organization (FAO).

The remaining part of this article is organized as follows: Sect. 2 describes the study area, data, and the flood prediction model, Sect. 3 evaluates the TMPA 3B42RT data from 2001 to 2006 at different spatial and temporal scales, Sect. 4 compares the performance of the flood prediction model at calibration and verification period with ground-based observations of streamflow, and Sect. 5 presents conclusions.

2 Study area, data, and hydrological model

2.1 Nzoia River basin

In general, repeated flooding is a serious problem in East Africa, particularly in the Lake Victoria Basin, which affects the lives of 30 million people of that region (Osano et al. 2003). The region around Lake Victoria is prone to flooding because of heavy rains and overflowing of the tributary rivers and streams. People in the heavily populated regions of Kenya, Uganda, and Tanzania live their life under a constant threat of flooding every year. In late May of 2002 alone, widespread flooding throughout Kenya displaced up to 60,000 people. Table 1 shows the severity of the problem in East Africa due to flooding in 2006 which reportedly claimed the lives of 1,000 people and displaced another 150,000 people.

The initial focus area of this project was to implement flood early warning information for Nzoia, a sub-basin of the Lake Victoria region because of its territorial, geographic, and epidemiological importance for the region. The Nzoia River sub-basin, lying in East Africa, covers approximately 12,696 km^2, and lies in the upstream of the Lake Victoria basin and Nile River basin (Fig. 1). It is bounded by the latitudes: 34°–36°E, and longitudes: 0°03′–1°15′ N, and the basin elevation ranges from 1,134 to 2,700 m. It encompasses three geographical regions: the highlands around Mount Elgon and the Cherangany Hills, the upper plateau, which includes Eldoret, and the lowlands. The region receives an average of 1,350 mm of rain annually and is an important cereal and sugarcane-farming region of Kenya, producing at least 30% of the national output of both maize and sugar. The total length of the river is 252 km with an average gradient of 4 m km^{-1}.

2.2 Data used in this study

The key datasets enabling the development of a distributed hydrological model in Africa include TMPA 3B42RT (Huffman et al. 2007; http://trmm.gsfc.nasa.gov), the digital elevation data from SRTM (Rabus et al. 2003; http://www2.jpl.nasa.gov/srtm/),

Table 1 Estimated impact of flooding in Kenya and East Africa (*Source*: RCMRD)

	Kenya 2006	East Africa 2006
Property loss ($ US)	250 million	400 million
Displaced persons	300,000	850,000
People needing emergency relief service	250,000	700,000
Deaths	260	1,000

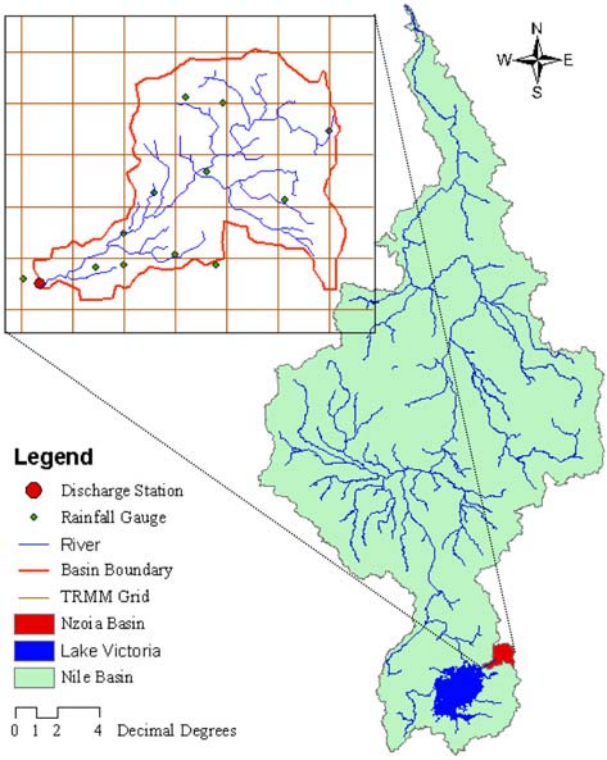

Fig. 1 Study area and gauge stations

SRTM-derived hydrological parameter files of HydroSHEDS (Lehner et al. 2008), soil parameters provided by the FAO 2003 (http://www.fao.org/AG/agl/agll/dsmw.htm), the MODIS land classification map is used as a surrogate for land use/cover, with 17 classes of land cover according to the International Geosphere–Biosphere Programme classification (Friedl et al. 2002). Details of several key datasets used in the study are described below.

2.2.1 NASA TMPA

A recent development in precipitation estimation from space involves combination of infrared measurements from geostationary satellites with passive microwave measurements from polar-orbiting satellites. The TMPA 3B42RT product (Huffman et al. 2007) is a near-real-time precipitation rate product at time and space scales (3-h, 0.25° × 0.25° latitude–longitude) over the latitude band 50° N–S. This product makes use of TRMM's highest quality observations, along with a high quality passive microwave-based rain estimates from 3 to 7 polar-orbiting satellites (e.g., Advanced Microwave Scanning Radiometer (AMSR-E), Special Sensor Microwave Imager from the Defense Satellite Meteorological Program (SSMI/DMSP), and Advanced Microwave Sounding Unit (AMSU-A)), and all the geosynchronous IR sensors (e.g., Meteorological Satellite (Meteosat), Geostationary Operational Environmental Satellites (GOES), GMS (Geostationary Meteorological Satellite)). The combined quasi-global rain map at 3-h resolution is produced by using TRMM to calibrate, or adjust, the estimates from all the other satellites, and then combining

all the estimates into the TMPA final product. The technique uses as much microwave data as possible, and uses the geo-IR estimates to fill in the gaps during the 3-h analysis. The calibrations are computed using the latest monthly accumulations of matched geo-IR and microwave data to ensure stability. The TMPA is a TRMM standard product. A real-time version of the TMPA merged product was introduced in February 2002 and is available on the U.S. TRMM web site (http://trmm.gsfc.nasa.gov).

2.2.2 NASA SRTM

The Earth's topography is an essential constraint and boundary condition in various aspects of Earth System Science researches from hydrological models predicting floods and runoff, to atmospheric boundary layer friction theories. Conventional topographic mapping technologies have produced maps of uneven quality—some with astounding accuracy, some far less adequate because of the lack of universal datum (Farr et al. 2007). Digital Elevation models (DEM), in the form of topographic maps, provide a base and context for airborne navigation systems and for a range of field activities in the civilian and military sectors. The Shuttle Radar Topography Mission (SRTM) is a joint collaboration between NASA and the National Geospatial-Intelligence Agency. SRTM collected interferometric radar data to generate a near-global topography data product. The Shuttle Radar Topography Mission produced the most complete, the highest resolution (30-m) digital elevation model of the Globe. It used dual radar antennas to acquire interferometric radar data, processed to digital topographic data at 1 arc-second resolution. Please refer to Farr et al. 2007 for more details of the development, flight operations, data processing, and products provided for users of this revolutionary dataset.

2.2.3 HydroSHEDS

HydroSHEDS is a new and innovative product that provides hydrographic information in a consistent and comprehensive format for regional- and global-scale applications. It offers a suite of geo-referenced datasets (vector and raster) at various scales, including river networks, watershed boundaries, drainage directions, and ancillary data layers such as flow accumulations, distances, and river topology information. HydroSHEDS is based on high-resolution elevation data obtained during a Space Shuttle flight for NASA's SRTM at 3 arc-second resolution. The original SRTM data have been hydrologically conditioned using a sequence of automated procedures (Lehner et al. 2008). The goal of developing HydroSHEDS data bank is to generate key data layers to support regional and global watershed analysis, hydrological modeling, and freshwater conservation planning at a quality, resolution, and extent that had previously been unachievable. Available resolutions range from 3 arc-second (approx. 90 m at the equator) to 5 min (approx. 10 km at the equator) with seamless near-global extent. Quality assessments indicate that the accuracy of HydroSHEDS significantly exceeds that of existing global watershed and river maps (Lehner et al. 2008). HydroSHEDS has been developed by the Conservation Science Program of World Wildlife Fund (WWF), in partnership with the U.S. Geological Survey (USGS), the International Centre for Tropical Agriculture (CIAT), The Nature Conservancy (TNC), and the Center for Environmental Systems Research (CESR) of the University of Kassel, Germany. HydroSHEDS data are free for non-commercial use. More information on HydroSHEDS and related data can be accessed by logging on to http://www.worldwildlife.org/hydrosheds and http://hydrosheds.cr.usgs.gov.

2.2.4 *Ground observations and other datasets*

Multi-year (2002 through 2006) daily observed precipitation of 12 gauges located within the Nzoia basin and daily discharge at the basin outlet (shown in Fig. 1) are provided by RCMRD. Daily historical meteorological data (e.g., daily maximum and minimum temperatures, and mean wind speed) were downloaded from NOAA national data center (NCDC climate data online. http://cdo.ncdc.noaa.gov/CDO/georegion) and were used to estimate potential evapotranspiration. The MODIS land classification map is used as a surrogate for land use/cover, with 17 classes of land cover according to the International Geosphere–Biosphere Programme classification (Friedl et al. 2002). Soil information from the Food and Agriculture Organization (FAO 2003, see http://www.fao.org/AG/agl/agll/dsmw.htm), SRTM DEM, and flow direction of HydroSHEDS with 30 s spatial resolution were used to extract catchment properties and hydrological routing parameters required by distributed hydrological model.

2.3 A conceptual physical distributed hydrological model

Module-structured gridded Xinanjiang Model (Zhao and Liu 1995) is a conceptual, physically based, distributed hydrological model. It was set up in the basin according to the resolution of GTOPO30 DEM and coupled with a flow routing scheme with direction file from HydroSHEDS. The model has been successfully and widely applied in humid and semi-humid regions in China since its development in the 1970s (Zhao 1992). It has been used for analyzing the impacts of climate change (Jiang et al. 2007) and other purposes mainly for flood forecasting (Jayawardena and Zhou 2000). Its runoff generation method has been used widely in distributed hydrological simulations recently (Li et al. 2006, 2008). Especially, its function to describe the soil moisture variation is employed in the Variable Infiltration Capacity (VIC) model (Wood et al. 1992).

The structure and flow chart of the distributed model is shown in Fig. 2a. First, runoff of each grid was calculated with runoff-generation module of Xinanjiang model. Then they were concentrated to the outlet along the flow path with a runoff routing scheme according to the concentration time and multi-linear reservoirs (Li et al. 2005).

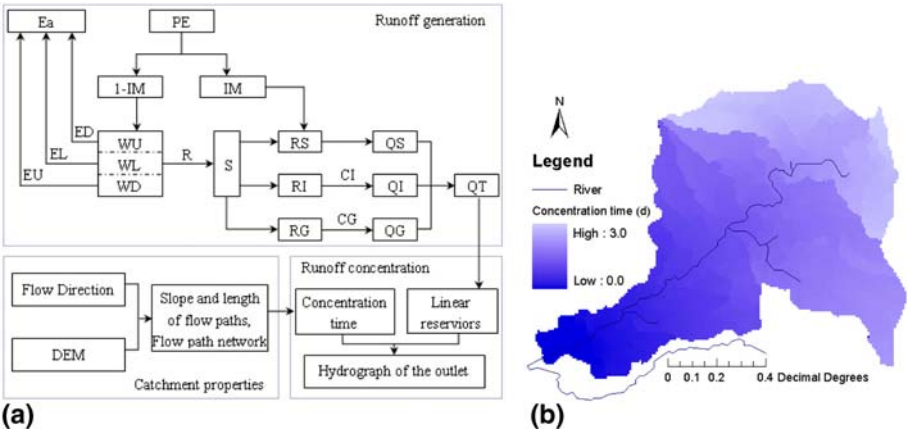

Fig. 2 **a** The structure and flow chart of the distributed hydrological model. **b** Concentration time of each grid in Nzoia basin

2.3.1 Runoff generation

As shown in Fig. 2a, runoff generation module is run in each grid according to the balance of precipitation and potential evapotranspiration (P-PET) at three soil layers (i.e., WU, WL, and WD). The residual P-PET begins to infiltrate and generate surface runoff RS, subsurface runoff RI, and groundwater RG. Then RI and RG will flow out and become QI and QG via recession constants CI and CG. Finally, total runoff (QT) generated in the surface of each grid is obtained by summing up QS, QG and QI, which ultimately becoming the input for runoff routing module.

Potential evapotranspiration was calculated with simplified FAO Penman–Monteith equation (Allen et al. 1998) with meteorological station observations of wind speed, mean, maximal and minimal daily temperature, and station location (longitude, latitude and altitude).

2.3.2 Runoff concentration

The runoff routing method was based on DEM and flow path of each grid which was extracted from direction file of HydroSHEDS. The routing method includes two parts: multi-linear reservoirs and concentration time. The time from each pixel to the outlet of the basin was calculated using length and slope of the flow path. At the same time a series of linear reservoirs were used to simulate the hydrograph from each grid to the next downstream grid. In the routing method, concentration time of each grid can be calculated with following equations:

$$T_i = \sum_{j=1}^{n_i} t_j \tag{1}$$

$$t_j = \frac{l_j}{v_j} = \frac{l_j}{K_V S_j^{\frac{1}{2}}} \tag{2}$$

where, T_i is the concentration time from center of the ith grid to the outlet; n_i is the number of grids passed through by flow of the ith grid from the center to the outlet, t_j is the time of flow from center of the jth grid to the next downriver grid center, l_j is the surface distance from center of the jth grid to the next downstream grid center; v_j is the average flow velocity from center of the jth grid to the next downstream grid center, S_j is the slope from center of the jth grid to the next downriver grid center, and K_v is defined as the velocity constant, implying the effect of hydraulic factors, such as roughness and hydraulic radius, on flow velocity. And the effect of roughness is more apparent than others (Li et al. 2005). The larger roughness results in smaller velocity constant. Figure 2b shows an example for concentration time distribution of each grid in Nzoia basin.

3 Precipitation comparison

3.1 Validation indices

Widely used validation statistical indices were selected for the evaluation of TMPA 3B42RT precipitation estimates against the gauged rainfall observations. Relative bias (Bias) was used to assess the systematic bias of satellite precipitation. The mean absolute

error (MAE) measured the average magnitude of the error. The RMSE also measures the average error magnitude but gives greater weight to the larger errors. The correlation coefficient (CC) is used to assess the agreement between TRMM precipitation and gauged observations.

$$\text{Bias} = \frac{\sum_{i=1}^{N} \widehat{P}(i) - \sum_{i=1}^{N} P(i)}{\sum_{i=1}^{N} P(i)} \times 100\% \tag{3}$$

$$\text{MAE} = \frac{\sum_{i=1}^{N} \left| \widehat{P}(i) - P(i) \right|}{N} \tag{4}$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N} \left| \widehat{P}(i) - P(i) \right|^2}{N}} \tag{5}$$

$$\text{CC} = \frac{\sum_{i=1}^{N} (P(i) - \bar{P}) \cdot \sum_{i=1}^{N} (\widehat{P}(i) - \bar{\widehat{P}})}{\sqrt{\sum_{i=1}^{N} (P(i) - \bar{P})^2 \cdot \sum_{i=1}^{N} (\widehat{P}(i) - \bar{\widehat{P}})^2}} \tag{6}$$

where $\widehat{P}$ is the precipitation estimation (e.g., TMPA 3B42RT) and P is the gauge observations; $\bar{\widehat{P}}$ and $\bar{P}$ are the mean of gauge and TRMM precipitation, respectively.

Several categorical verification statistics, which measure the correspondence between the estimated and observed occurrence of events, are also used in this study. They are probability of detection (POD), false alarm ratio (FAR), and critical success index (CSI).

$$\text{POD} = \frac{\text{hits}}{\text{hits} + \text{misses}} \tag{7}$$

$$\text{FAR} = \frac{\text{false alarms}}{\text{hits} + \text{false alarms}} \tag{8}$$

$$\text{CSI} = \frac{\text{hits}}{\text{hits} + \text{misses} + \text{false alarms}} \tag{9}$$

where "misses" means the number of sampling times that gauge classifies as rain but the TMPA 3B42RT algorithm shows no rain, "hits" indicate both the algorithm and gage detect rainfall events, and false alarms means the algorithm show raining signals but rain did not occur according to the gages. POD measures the fraction of observed events that were correctly diagnosed, and is sometimes called the "hit rate". FAR gives the fraction of diagnosed events that were actually nonevents. CSI gives the overall fraction of correctly diagnosed events by the TMPA 3B42RT algorithm. Perfect values for these scores are POD = 1, FAR = 0, and CSI = 1.

3.2 Evaluation of precipitation validation results and discussion

Point-based and areal-based daily rainfall data were computed from the 3B42 RT and gauge observations over the time span of 2002–2006, respectively. Point-based time series

were simply compared with the spatiotemporally co-located gauges and TMPA's grids while the areal-based basin time series were averaged over an entire basin.

Validation results shown in Table 2 and Fig. 3 indicate that the point-based (areal-based) TMPA 3B42RT time series systematically overestimate 22% (15%) with 0.51 (0.57) "goodness of fit" at daily scale. By plotting the rainfall occurrence as a function of rain intensity, Fig. 3 indicates the overestimation, i.e., positive bias, mostly occur at rainfall intensity less than 1 mm/d or higher than 10 mm/d while underestimation has occurred in between the given range. The results also imply that the areal-based average method effectively reduces the Bias, MAE, RMSE, among other indicators. As a matter of fact, the FAR rather becomes worse, increasing from 0.136 to 0.188, because of the fact that averaging the 3B42RT data from points to area amplifies the local grid-based FAR to the basin-wide FAR.

Table 2 Statistics of point and areal-based basin-mean 3B42RT's precipitation

Statistics	Mean		Bias (%)	CC	MAE	RMSE	POD	FAR	CSI
	Gage	3B42RT							
Point-average	3.94	4.82	22.49	0.51	3.70	6.24	0.953	0.136	0.82
Areal-average	3.12	3.59	15.06	0.57	2.80	4.82	0.996	0.188	0.89

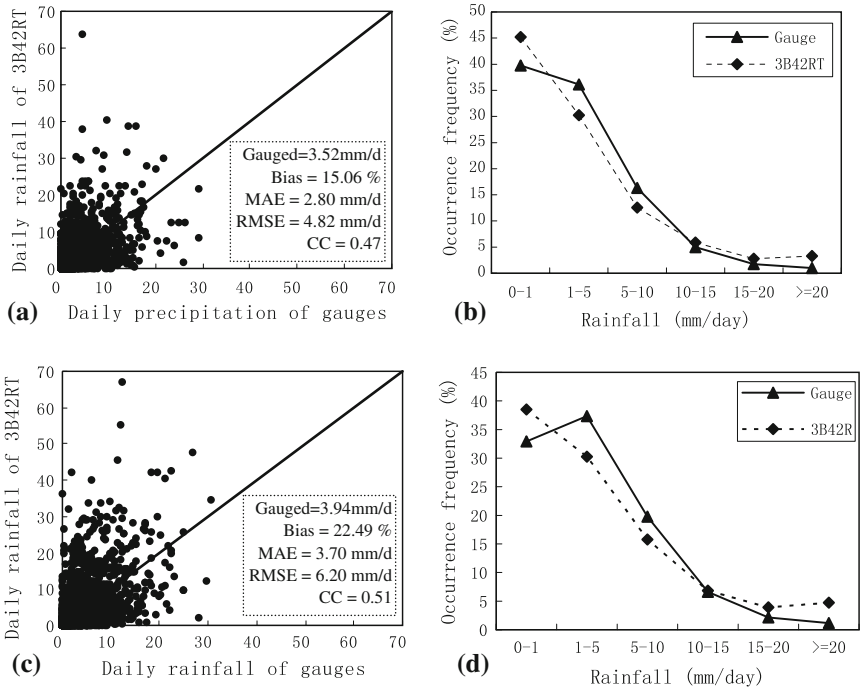

Fig. 3 Temporal evaluation of 3B42RT's data. **a** Scatter plots of areal-based gauge's and 3B42RT's daily precipitation data. **b** Rainfall occurrence of areal-based gauge observation and 3B42RT as a function of rainfall intensity, respectively. **c** Scatter plots of point-based gauge's and 3B42RT's daily precipitation data. **d** Rainfall occurrence of point-based gauge observation and 3B42RT as a function of rainfall intensity, respectively

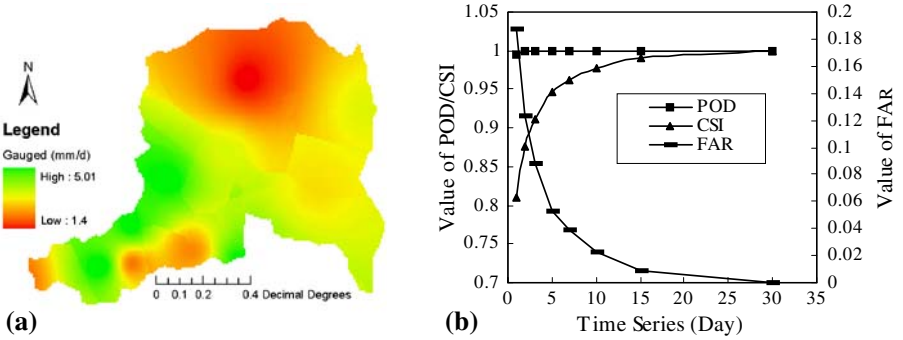

Fig. 4 **a** Spatial distribution of daily gauge rainfall data derived from 2002 to 2006. **b** Change of verification statistics as a function of time scales

Spatially, Fig. 4a shows lower basin receives heavier rainfall than upper basin according to the 6-year gauge-averaged rainfall (2002–2006). In general, the 3B42RT data follow the similar spatial distribution patterns with slight overestimation over downstream and underestimation over upper stream (figure not shown here). 3B42RT's data with different aggregation time step was also compared with gauged observation in order to evaluate the change of its estimation performance as a function of time scale (shown in Fig. 4b). When aggregating the data from 1-day to 5-day scale, dramatic improvements of the categorical verification indices can be seen: CSI, from 0.81 to 0.96 and reduction of FAR from 0.188 to 0.05 (Fig. 4b).

Evaluation of the 6-year TMPA 3B42RT daily rainfall data against gauge observations indicates that the 3B42RT has potential for hydrological modeling in this area, with good correlation coefficients and very good skill scores given its high POD, CSI and small FAR. However, though 15–20% bias ratio may be acceptable for hydrological modeling in lower medium basins, it also prompted us to implement an additional simulation scheme at next section: to calibrate the hydrological model with gauge rainfall data and then simulate the discharges with systematically bias-removed 3B42RT, i.e., 3B42V6 (Huffman et al. 2007).

4 Hydrological model calibration and prediction

4.1 Calibration and evaluation criteria of hydrological model performance

Quantitative comparison was done between predicted and observed flows. We used two most-popular statistical indices to assess the hydrological model performance. First, relative bias ratio was calculated according to Eq. 3 between observed and modeled streamflows. Second, for statistical "goodness of fit" of simulated flows, we used a widely used performance indicator—the Nash-Sutcliffe Coefficient of Efficiency (NSCE). The NSCE is an indicator of the model's ability to predict about the 1:1 line. The NSCE is defined as:

$$\mathrm{NSCE} = \frac{\sum (Q_{i,o} - \overline{Q_0})^2 - \sum (Q_{i,o} - Q_{i,c})^2}{\sum (Q_{i,o} - \overline{Q_0})^2} \tag{10}$$

where $Q_{i,o}$ is the observed discharge of ith day; $Q_{i,c}$ is the simulated discharge of ith day; $\overline{Q_0}$ is average value of all the daily observed discharge; and $\overline{Q_c}$ is the average value of all daily simulated discharge.

Two interrelated procedures are involved in the model calibration. First, reduction of Bias is looked as the main objective to calibrate runoff-generation parameters, such as the ratio of actual ET to PET and water capacity of the three soil layers, in order to best match the total runoff volume generated by the model to the observations. Meanwhile, improvement of the NSCE is deemed as the objective to calibrate runoff-routing parameters such as concentration time and coefficient of linear reservoir in order to adjust the routed hydrograph to best match with the discharge observations. The ideal scenario is when Bias = 0 and NSCE = 1.

4.2 Calibration and simulation results

Initially, the Xinanjiang model parameters were estimated using digital datasets of soil, MODIS land cover types, SRTM elevation, and HydroSHEDS. Afterward, we went through two phases of model set-up: (1) 1985–2001 hydrological model warming up period and (2) 2002–2006 hydrological model calibration and verification.

4.2.1 Model warming up

As the observed gauge rainfall and discharges data are available since 1985, we first calibrated the model with the data from 1985 to 2001 for warming up rather than using TMPA 3B42RT data which are available only after 2001. Figure 5a shows calibrated runoff slightly underestimated (−5.2%) with relatively good NSCE (0.73). The 17-year averaged daily discharge is 108.6 (m^3/s) from station observations and 102.9 (m^3/s) from the calibrated model calculation, respectively. However, Table 3 and Fig. 5b also indicate that the warming-up calibration during the 1985–2001 period did not perform consistently for the subsequent discharge prediction period (2002–2006) that was forced by 3B42RT rainfall, resulting in a relatively large bias = 26.6% and lower NSCE = 0.53.

4.2.2 Model calibration and simulation

Because TMPA 3B42RT rainfall data were only available after 2001, we implemented the hydrological model for the period 2002–2006 during which both satellite and gage data are available. After recalibration against gage rainfall for this period, we observed that the bias ratio and NSCE score became 0.94% and 0.84, respectively (Fig. 6a). Afterward, the calibrated hydrological model was forced by 3B42RT rainfall data and the results shown in Fig. 6b over-predicted daily stream flows 20.4% but with promising NSCE score (0.67). Based on the rainfall validation results from Table 2 in Sect. 3, it is shown that 3B42RT

Table 3 Summary of the calibration and overall results

Phases	Calibration and results				Prediction and results			
	Time span	Forcing rainfall	Bias (%)	NSCE	Time span	Forcing data	Bias (%)	NSCE
Warming up	1985–2001	Gauge	−5.20	0.73	2002–2006	3B42RT	26.6	0.53
Implementation	2002–2006	Gauge	−0.94	0.84	2002–2006	3B42RT	20.37	0.67
					2002–2006	Bias-corrected 3B42RT	3.60	0.71

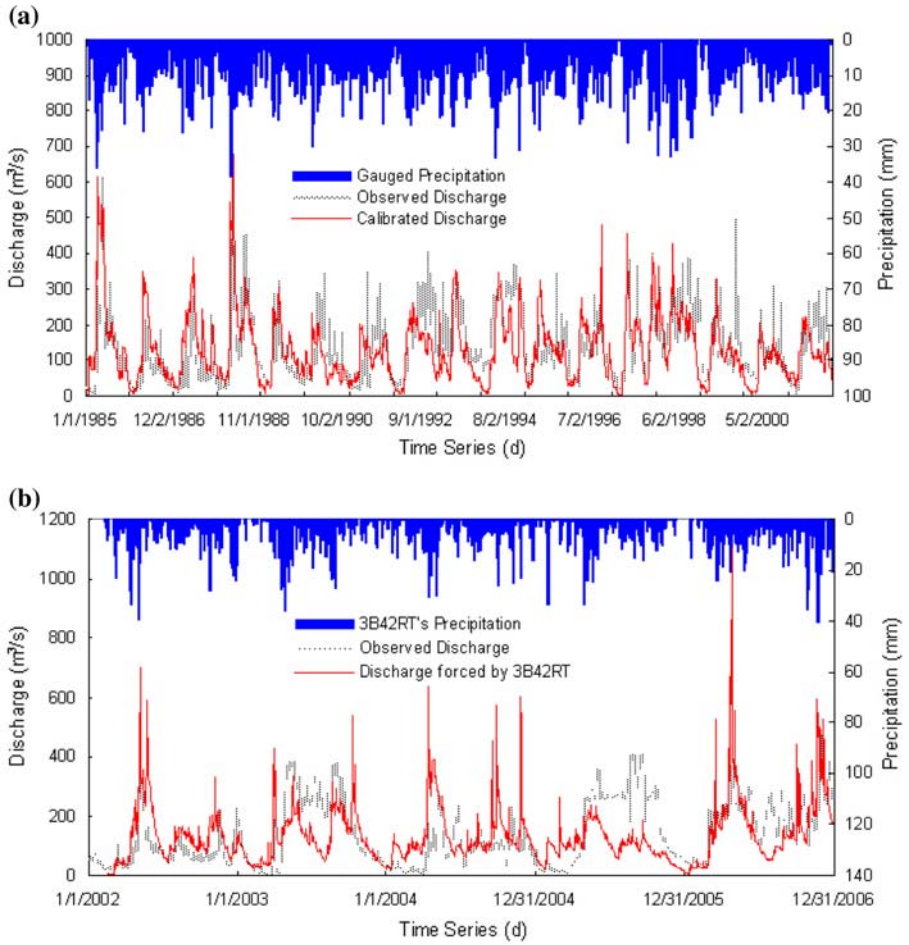

Fig. 5 Hydrograph of daily stream flow for (**a**) warming up period (1985–2001) forced by gage rainfall and (**b**) prediction period (2002–2006) forced by 3B42RT

generally over-estimated the rainfall 15–20%. Therefore, we further simulated the stream flows forced by the bias-corrected 3B42RT, i.e., 3B42V6 (Huffman et al. 2007). This reduced the overall bias of discharge prediction to 3.6% at the same time improving the NSCE score to 0.71 (Table 3). Showing the accumulation of daily discharges from 2002 through 2006, Fig. 6c also demonstrates the improvement of NSCE and reduction of bias with forcing data from bias-corrected 3B42RT over the original 3B42RT data.

5 Summary and discussion

Satellite-based rainfall and geospatial datasets are potentially useful for cost-effective detection and early warning of natural hazards, such as floods, specifically for regions of the world where local data are sparse or non-existent. Many researchers seek to take

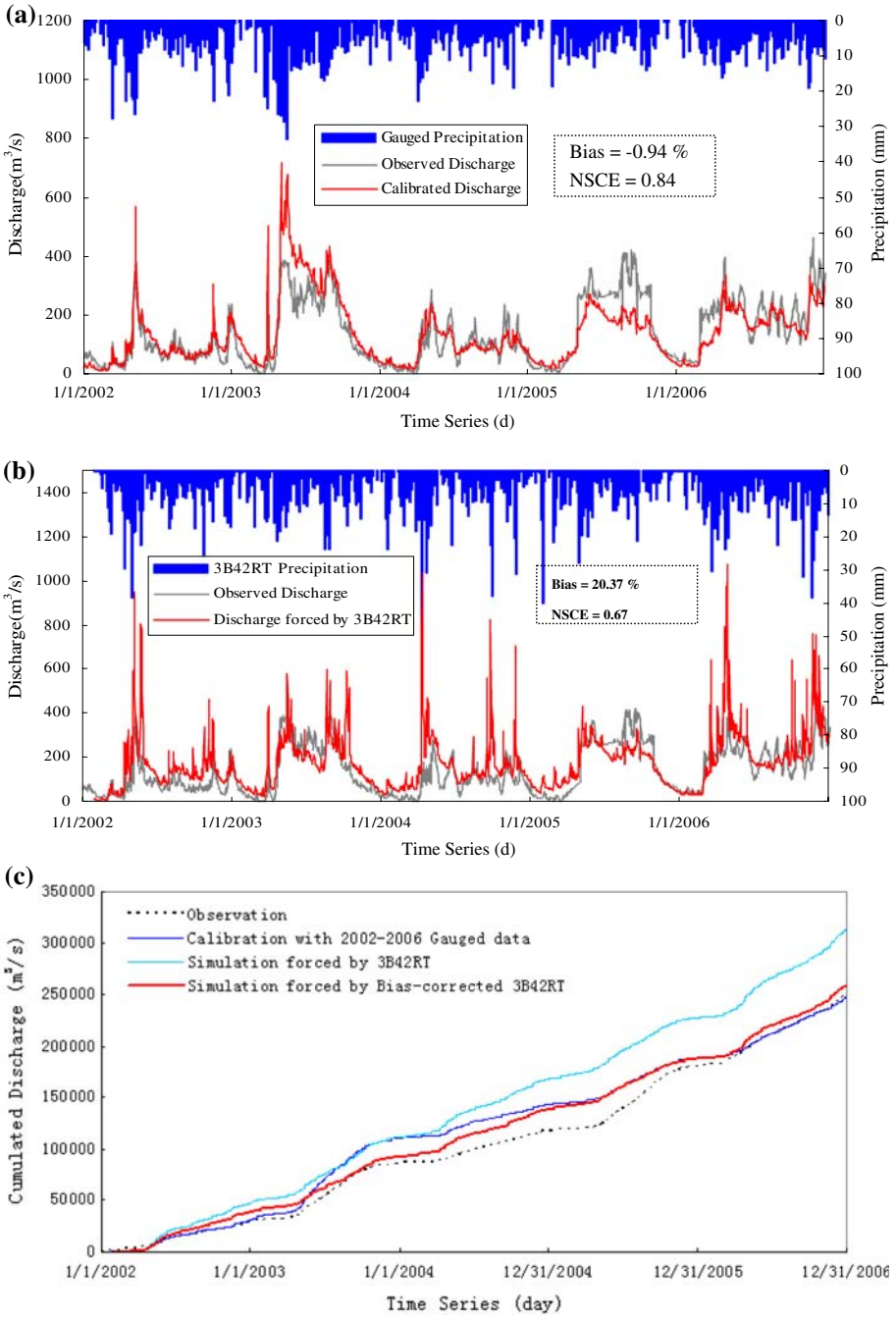

Fig. 6 Hydrological model calibration and simulation over period of 2002–2006. **a** Model calibration using gage rainfall. **b** Model simulation using TMPA 3B42RT. **c** Accumulation of discharges data over the period of 2002–2006

advantage of the recently available and virtually uninterrupted supply of satellite-based rainfall information as an alternative and supplement to the ground-based observation in order to implement a cost-effective flood prediction in many under-gauged regions around the world. The goal of this study is to build disaster management capacity in East Africa by providing local governmental officials and international aid organizations a practical decision-support tool so as to better asses emerging flood impacts and to quantify spatial extent of water hazard risk, as well as to respond to such flood emergencies more expediently.

In this study, the applicability of TMPA 3B42RT rainfall estimates for distributed hydrological modeling over a flood-prone region, Nzoia, a sub-basin of Lake Victoria and Nile River (Fig. 1) was evaluated. This research first validated the TMPA 3B42RT rainfall products with gauged daily rainfall observations from the year 2002 through 2006 (Table 2; Figs. 3, 4). Afterward, a spatially distributed hydrological model was setup and calibrated to Nzoia basin. The key datasets enabling the development of a distributed hydrological model there include TMPA 3B42 Real-time rainfall estimates, the digital elevation data from the NASA SRTM mission, hydrological parameter files of Hydro-SHEDS, the MODIS Land cover, and soil information provided by FAO. Then benchmark streamflow simulation performance was produced by the calibrated hydrological model using the rain gauge and observed streamflow data provided by African RCMRD, local partner of this project. Finally, continuous discharges forced by TMPA Real-time data from 2002 through 2006 were simulated and reasonable results were obtained in comparison with the benchmark performance according to the designated statistical indices such as bias ratio and NSCE. Figures 5 and 6 and Table 3 summarize the overall performance of the model calibration and prediction results. In general, the results demonstrate that the TMPA 3b42RT precipitation data for flood prediction in this basin is acceptable, although the 7-year (1985–2001) gauged-rainfall warming-up does not appear to yield greater benefit for consequent discharge modeling forced by 3B42RT at later time period in terms of relatively high bias ratio (26.6%) and low NSCE score (0.53). During the implementation time period (2002–2006), both gage calibration and 3B42RT simulation show good results. Moreover, we also identified that improved flood prediction performance could be achieved with systematically bias-corrected TMPA 3B42RT rainfall data over original real-time data.

Although the results justify to suggest to us that TMPA 3B42RT can be acceptably used to drive hydrological models for flood prediction purpose in Nzoia basin, full realization of the potential of seamless satellite-based rainfall estimates requires further investigation of optimal calibration strategy for integrating remote sensing data into a real-time hydrological modeling system for vast ungauged regions of the world (Hossain and Lettenmaier 2006). On the other hand, results of this study also demand continuous progress in spaceborne rainfall estimation technology in terms of both the accuracy and spatiotemporal resolutions of rainfall estimates. In this regard, future deployment of the Global Precipitation Measuring mission (http://gpm.gsfc.nasa.gov) would largely facilitate this particular study. Finally, it can also be concluded that to increase flood forecasting lead time, more reliable and more accurate short- or medium-range quantitative precipitation forecasts is a must.

Acknowledgments The financial support from NASA Applied Science Program SERVIR-Africa Project and from University of Oklahoma is gratefully acknowledged. The authors also thank the RCMRD (Tesfaye Korme and Lawrence Okello) for providing gauged rainfall and streamflow observations over Nzoia Basin.

References

Allen RG, Pereira LS, Raes D, Smith M (eds) (1998) Crop evapotranspiration-guidelines for computing crop water requirements—FAO irrigation and drainage paper 56, Rome

Farr TG et al (2007) The shuttle radar topography mission. Rev Geophys 45:RG2004. doi:10.1029/2005RG000183

Friedl MA et al (2002) Global land cover mapping from MODIS: algorithms and early results. Remote Sens Environ 83(1–2):287–302. doi:10.1016/S0034-4257(02)00078-0

Hong Y, Alder RF, Hossain F, Curtis S, Huffman GJ (2007a) A first approach to global runoff simulation using satellite rainfall estimation. Water Resour Res 43(8):W080502. doi:10.1029/2006GL028010

Hong Y, Alder RF, Negri A, Huffman GJ (2007b) Flood and landslide applications of near real-time satellite rainfall products. Nat Hazards 43(2):285–294. doi:10.1007/s11069-006-9106-x

Hossain F, Lettenmaier DP (2006) Flood prediction in the future: recognizing hydrologic issues in anticipation of the global precipitation measurement mission. Water Resour Res 42:W11301. doi:10.1029/2006WR005202

Huffman GJ, Adler RF, Bolvin DT, Gu G, Nelkin EJ, Bowman KP, Hong Y, Stocker EF, Wolff DB (2007) The TRMM multisatellite precipitation analysis: quasi-global, multi-year, combined-sensor precipitation estimates at fine scale. J Hydrometeorol 8:38–55. doi:10.1175/JHM560.1

IPCC (2007) Climate change 2007. IPCC Fourth Assessment Report (AR4)

Jayawardena AW, Zhou MC (2000) A modified spatial soil moisture storage capacity distribution curve for the Xinanjiang model. J Hydrol 227:93–113. doi:10.1016/S0022-1694(99)00173-0

Jiang T, Chen YD, Xu C-Y, Chen X, Chen X, Singh VP (2007) Comparison of hydrological impacts of climate change simulated by six hydrological models in the Dongjiang Basin, South China. J Hydrol 336:316–333. doi:10.1016/j.jhydrol.2007.01.010

Lehner B, Verdin K, Jarvis A (2008) New global hydrography derived from spaceborne elevation data. Eos Trans AGU 89(10):93–94. doi:10.1029/2008EO100001

Li L, Hao ZC, Wang J (2005) A distributed hydrologic model based on DEM. Proc China Assoc Sci Technol 2(3):510–513

Li ZJ, Cheng Y, Xu P-Z (2006) Application of GIS-based hydrological models in humid watersheds. Proceedings of the 15th APD-IAHR & ISMH: water for life—surface and ground water resources. IIT, Madras

Li L, Hao Z-C, Wang J-H et al (2008) Impact of future climate change on runoff in the head region of the Yellow River. J Hydrol Eng 13(5):347–354. doi:10.1061/(ASCE)1084-0699(2008)13:5(347)

Osano O, Nzyuko D, Tole M, Admiraal W (2003) The fate of chloroacetanilide herbicides and their degradation products in the Nzoia Basin, Kenya. AMBIO: J Hum Environ 32(6):424–427

Rabus B, Eineder M, Roth A, Bamler R (2003) The shuttle radar topography mission—a new class of digital elevation models acquired by spaceborne radar. Photogramm Remote Sens 57:241–262. doi:10.1016/S0924-2716(02)00124-7

Wood EF, Lettenmaier DP, Zartarian VG (1992) A land-surface hydrology parameterization with subgrid variability for general circulation models. J Geophys Res 97:2717–2728

Zhao RJ (1992) The Xinanjiang model applied in China. J Hydrol 135:371–381. doi:10.1016/0022-1694(92)90096-E

Zhao RJ, Liu XR (1995) The Xinanjiang model. In: Singh VP (ed) Computer models of watershed hydrology. Water Resources Publications, USA, pp 215–232