Journal of Hydrology 499 (2013) 91–99

Contents lists available at SciVerse ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Statistical and hydrological evaluation of TRMM-based Multi-satellite Precipitation Analysis over the Wangchu Basin of Bhutan: Are the latest satellite precipitation products 3B42V7 ready for use in ungauged basins?

Xianwu Xue [a,b], Yang Hong [a,b,*], Ashutosh S. Limaye [c], Jonathan J. Gourley [d], George J. Huffman [e], Sadiq Ibrahim Khan [a], Chhimi Dorji [f], Sheng Chen [a]

[a] School of Civil Engineering and Environmental Sciences, University of Oklahoma, Norman, OK 73072, United States
[b] Advanced Radar Research Center, National Weather Center, Norman, OK 73072, United States
[c] ZP11/ Earth Science Office, NASA Marshall Space Flight Center 320 Sparkman Dr., Huntsville, AL 35805, United States
[d] NOAA/National Severe Storms Laboratory, Norman, OK 73072, United States
[e] NASA Goddard Space Flight Center, Greenbelt, MD 20771, United States
[f] Department of Hydro-Met Services, Ministry of Economic Affairs, Thimphu, Bhutan

ARTICLE INFO

Article history:
Received 5 October 2012
Received in revised form 6 June 2013
Accepted 23 June 2013
Available online 1 July 2013
This manuscript was handled by Konstantine P. Georgakakos, Editor-in-Chief, with the assistance of Hervé Andrieu, Associate Editor

Keywords:
CREST model
A-priori parameter estimation
Hydrologic modeling evaluation
Precipitation estimation

SUMMARY

The objective of this study is to quantitatively evaluate the successive Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA) products and further to explore the improvements and error propagation of the latest 3B42V7 algorithm relative to its predecessor 3B42V6 using the Coupled Routing and Excess Storage (CREST) hydrologic model in the mountainous Wangchu Basin of Bhutan. First, the comparison to a decade-long (2001–2010) daily rain gauge dataset reveals that: (1) 3B42V7 generally improves upon 3B42V6's underestimation both for the whole basin (bias from −41.15% to −8.38%) and for a $0.25^\circ \times 0.25^\circ$ grid cell with high-density gauges (bias from −40.25% to 0.04%), though with modest enhancement of correlation coefficients (CC) (from 0.36 to 0.40 for basin-wide and from 0.37 to 0.41 for grid); and (2) 3B42V7 also improves its occurrence frequency across the rain intensity spectrum. Using the CREST model that has been calibrated with rain gauge inputs, the 3B42V6-based simulation shows limited hydrologic prediction NSCE skill (0.23 in daily scale and 0.25 in monthly scale) while 3B42V7 performs fairly well (0.66 in daily scale and 0.77 in monthly scale), a comparable skill score with the gauge rainfall simulations. After recalibrating the model with the respective TMPA data, significant improvements are observed for 3B42V6 across all categories, but not as much enhancement for the already-well-performing 3B42V7 except for a reduction in bias (from −26.98% to −4.81%). In summary, the latest 3B42V7 algorithm reveals a significant upgrade from 3B42V6 both in precipitation accuracy (i.e., correcting the underestimation) thus improving its potential hydrological utility. Forcing the model with 3B42V7 rainfall yields comparable skill scores with in situ gauges even without recalibration of the hydrological model by the satellite precipitation, a compensating approach often used but not favored by the hydrology community, particularly in ungauged basins.

1. Introduction

Precipitation is among the most important forcing data for hydrological models. It has been arguably nearly impossible for hydrologists to simulate the water cycles over regions with no or sparse precipitation gauge networks, especially over complex terrain or remote areas. Recently, the satellite precipitation products such as TMPA (Huffman et al., 2007), CMORPH (Joyce et al., 2004), PERSIANN (Sorooshian et al., 2000) and PERSIANN-CCS (Hong et al., 2004) are starting to provide alternatives for estimating rainfall data and also pose new challenges for hydrologists in understanding and applying the remotely-sensed information.

The Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis (TMPA), developed by the National Aeronautics and Space Administration (NASA) Goddard Space Flight Center (GSFC), provides a calibration-based sequential scheme for combining precipitation estimates from multiple satellites, as well as monthly gauge analyses where feasible, at fine spatial and

* Corresponding author. Address: Atmospheric Radar Research Center, National Weather Center suite 3630, 120 David L. Boren Blvd., Norman, OK 73072-7303, United States. Tel.: +1 405 325 3644.
E-mail address: yanghong@ou.edu (Y. Hong).

0022-1694/$ - see front matter
http://dx.doi.org/10.1016/j.jhydrol.2013.06.042

temporal scales ($0.25° \times 0.25°$ and 3 hourly) over 50°N–50°S (Huffman et al., 2007). TMPA is computed for two products: near-real-time version (TMPA 3B42RT, hereafter referred to as 3B42RT) and post-real-time research version (TMPA 3B42 V6, hereafter referred to as 3B42V6). 3B42V6 has been widely used in hydrological applications (Bitew and Gebremichael, 2011; Bitew et al., 2011; Khan et al., 2011a, 2011b; Li et al., 2012; Stisen and Sandholt, 2010; Su et al., 2008), however, its computation ended June 30th 2011 and 3B42V6 was replaced by the new version (TMPA 3B42 V7, hereafter referred to as 3B42V7), which has been reprocessed and available from 1998 to present. Previously, 3B42V6 has been validated by several studies (Bitew and Gebremichael, 2011; Bitew et al., 2011; Chokngamwong and Chiu, 2008; Islam and Uyeda, 2007; Jamandre and Narisma, 2013; Jiang et al., 2012; Li et al., 2012; Mishra et al., 2010; Stisen and Sandholt, 2010; Su et al., 2008; Yong et al., 2012, 2010), while the newly available 3B42V7 is evaluated in tropical cyclone systems (Chen et al., 2013) and the United States (Chen et al., in press; Kirstetter et al., 2013), it has not been extensively statistically and hydrologically validated in mountainous South Asian regions.

Therefore, the objectives of this study are designed (1) to evaluate the widely used globally-available, high-resolution TMPA satellite precipitation products over the mountainous medium-sized Wangchu basin (3550 km^2) in Bhutan, and more importantly (2) to assess improvements of the latest upgrade version (3B42V7) relative to its predecessor in terms of statistical performance and hydrologic utility. Additionally, this study aims to shed light on the suitability of recalibrating a hydrological model with the remotely-sensed rainfall information. The remainder of this paper is organized as follows: Section 2 introduces the study area, the datasets used, and the methodology, including a brief description of the CREST distributed hydrological model and its upgrade to the new version (CREST Version 2.0). The results are discussed in Section 3, and then Section 4 draws the conclusions of this study.

2. Study area, data and methodology

2.1. Study area

The Wangchu Basin, with a total drainage area of approximately 3550 km^2 is located within 89°6′–89°46′E and 27°6′–27°51′N in the west of Bhutan (Fig. 1). Wangchu Basin is the most populous part of the country with about 3/5 of the population living in 1/5 of the basin area. The basin is equipped with one streamflow gauge at the outlet Chhukha Dam Hydrological station and five rain gauge stations. The soil types are dominated by Sandy Clay Loam (75.1%) and Loam (24.9%) based on the Harmonized World Soil Database (HWSD v1.1) (FAO/IIASA/ISRIC/ISSCAS/JRC, 2009). The various vegetation types of this basin are composed of evergreen needleleaf forest (48.1%), woodland (17.8%), open shrubland (9.7%), wooded grassland (8.2%), grassland (7.6%) and other land-use types (less than 10%) (Hansen et al., 2000).

The northern periphery of the Wangchu Basin in the Himalayas has elevations over 6000 m and maintains an annual snowpack. Lower portions of the basin are drastically different and are subject to a summer monsoon from May to October (Bookhagen and Burbank, 2010). On average, the annual month with the greatest precipitation is July or August with 161–546 mm/month based on the five rain gauge station data shown in (Table 1), and the largest resulting streamflow occurs in June or August with 251 m^3 s^{-1}. It is possible that snowmelt contributes to a portion of this peak streamflow, but the majority is driven by the summer monsoon rains. In this study, neither the precipitation products nor the model explicitly deal with frozen precipitation. These are subjects requiring additional investigation, especially in light of the forthcoming Global Precipitation Measurement Mission (GPM), which aims to quantitatively estimate frozen precipitation amounts.

2.2. In situ and satellite precipitation datasets

2.2.1. Gauged precipitation and discharge data

Daily observed precipitation data are obtained from the Hydro-Met Services Department of Bhutan from 2001 to 2010 for the 5 rain gauge stations located within the Wangchu basin. In winter, frozen precipitation is reported in the form of water equivalent and computed by melting the ice/snow with hot water in the standard vessel and deducting the hot water volume from the total volume. The Thiessen polygon method is used to interpolate the rain gauge data to the spatial distributed grid data fitting the model grid resolution (30 arcsec) (Fig. 1). We also obtained the daily discharge data at the basin outlet for the same time period.

2.2.2. TMPA 3B42 research products

TMPA precipitation products are available in two versions: near-real-time version (3B42RT) and post-real-time research version (3B42) adjusted by monthly rain gauge data. The 3B42 products have two successive versions: version 6 and the latest version 7 (3B42V6 and 3B42V7). In this study, we evaluated and compared the two high-resolution (3 h and $0.25° \times 0.25°$) satellite precipitation products: 3B42V6 and 3B42V7.

The TMPA algorithm (Huffman et al., 2007) calibrates and combines microwave (MW) precipitation estimates, and then creates the infrared precipitation (IR) estimates using the calibrated MW. After this, it combines the MW and IR estimates to create the TMPA precipitation estimates. MW data used in Version 6 are from the TRMM Microwave Imager (TMI), Special Sensor Microwave Imager (SSM/I) F13, F14 and F15 on Defense Meteorological Satellite Program (DMSP) satellites, and the Advanced Microwave Scanning Radiometer-Earth Observing System (AMSR-E) on Aqua, and the Advanced Microwave Sounding Unit-B (AMSU-B) N15, N16 and N17 on the NOAA satellite; IR data collected by geosynchronous earth orbit (GEO) satellites, GEO-IR. The 3B42V6 also use other data sources: TRMM Combined Instrument (TCI) employed from TMI and PR, monthly rain gauge data from GPCP ($1° \times 1°$) and the Climate Assessment and Monitoring System (CAMS) $0.5° \times 0.5°$ developed by CPC. Based on the lessons learned in 3B42V6, 3B42V7 includes consistently reprocessed versions for the data sources used in 3B42V6 and introduces additional datasets, including the Special Sensor Microwave Imager/Sounder (SSMIS) F16-17 and Microwave Humidity Sounder (MHS) (N18 and N19) and Meteorological Operational satellite programme (MetOp) and the 0.07° Grisat-B1 infrared data. All of these data can be freely downloaded from the website: http://trmm.gsfc.nasa.gov/ and http://mirador.gsfc.nasa.gov.

2.2.3. Evapotranspiration

The potential evapotranspiration (PET) data used in this study are from the global daily potential evapotranspiration database provided by the Famine Early Warning Systems Network (hereafter referred as FEWSPET) global data portal (see http://earlywarning.usgs.gov/fews/global/web/readme.php?symbol=pt). FEWSPET is calculated from the climate parameter extracted from global data assimilation system (GDAS) analysis fields, has 1-degree resolution, and covers the entire globe from 2001 to the present.

2.3. CREST model

The Coupled Routing and Excess Storage (CREST) Model (Khan et al., 2011a; Khan et al., 2011b; Wang et al., 2011) is a grid-based distributed hydrological model developed by the University of

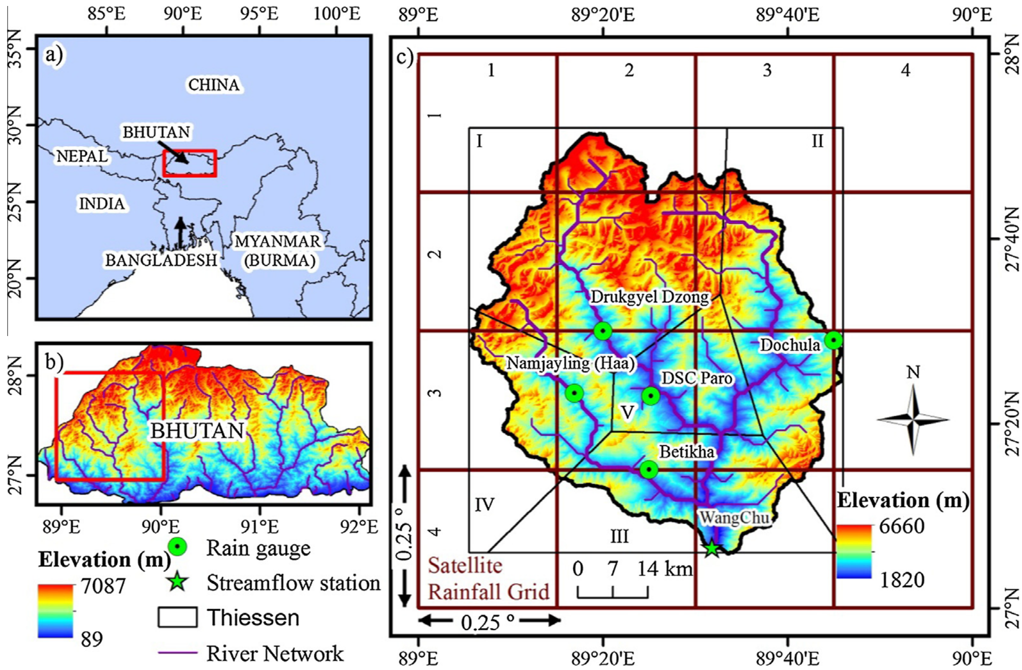

Fig. 1. Wangchu Basin Map. (a) Location of Bhutan and the surrounding countries. (b) Location of Wangchu Basin in Bhutan and its elevation. (c) Map of Wangchu Basin, rain gauges, streamflow station, topography, Thiessen polygons applied to the rain gauge data and the 0.25° × 0.25° grids of the satellite rainfall estimates.

Table 1
Monthly observed precipitation and runoff averaged from 2001 to 2010 (bold value means the maximum value of this station).

Month	Rain Gauge (mm/month)					Streamflow station
	Betikha	Dochula	Drukgyel Dzong	Namjayling Haa	DSC_Paro	Chhukha ($m^3 s^{-1}$)
January	14	19	0	12	8	26
February	49	18	9	24	18	23
March	176	17	26	33	15	25
April	346	53	29	54	34	38
May	368	105	60	69	57	55
June	390	279	123	124	81	111
July	**546**	359	185	183	**199**	222
August	383	**368**	**191**	**161**	103	**251**
September	326	217	108	120	77	180
October	182	116	71	77	63	109
November	10	11	3	4	3	52
December	4	9	1	2	1	34

Oklahoma (http://hydro.ou.edu) and NASA SERVIR Project Team (www.servir.net). It computes the runoff generation components (e.g., surface runoff and infiltration) using the variable infiltration capacity curve (VIC), a concept originally contained in the Xinanjiang Model (Zhao, 1992; Zhao et al., 1980) and later represented in the VIC Model (Liang et al., 1994, 1996). Multi-linear reservoirs are used to simulate cell-to-cell routing of surface and subsurface runoff separately. The CREST model couples the runoff generation component and cell-to-cell routing scheme described above, to reproduce the interaction between surface and subsurface water flow processes. Besides the hydrologic and basic data (DEM, flow direction, flow accumulation, slope, etc.), the CREST model employs gridded precipitation and potential evapotranspiration (PET) data as its forcing data. CREST Version 1.6 model has been applied at both global (Wu et al., 2012) and regional scales (Khan et al., 2011a, 2011b) (more applications can be found at website: http://eos.ou.edu and http://www.servir.net).

The CREST model used in this study is the upgraded version CREST V2.0. The main features of the latest version are: (1) enhancement of the computation capability using parallel distribution techniques to make the model more efficient than the previous version (Wang et al., 2011); (2) model implementation with options of either spatially uniform, semi-distributed, or distributed parameter values; (3) automatic extraction of *a-priori* model parameter estimates from high-resolution land cover and soil texture data. The physically-based parameters, K_{sat} and WM, can be derived from land cover types and soil texture data based on a look-up table (Chow et al., 1988); (4) a modular design framework to accommodate research, development and system enhancements (see Fig. 2(a)); and (5) inclusion of the optimization scheme SCE-UA (Duan et al., 1992; Duan et al., 1993) to enable automatic calibration of the CREST model parameters (see Fig. 2(a)). Table 2 shows 11 parameters and their descriptions, ranges and default values. Fig. 2(b) shows the vertical profile of hydrological processes in a grid cell. It shows the precipitation is intercepted by a canopy to generate throughfall, and then the throughfall is separated into surface runoff and infiltration components by the variable infiltration curve. Finally, two linear reservoirs are employed to simulate sub-grid cell routing.

2.4. Evaluation statistics

In order to quantitatively analyze the performance of 3B42V6 and 3B42V7 precipitation products against rain gauge observations

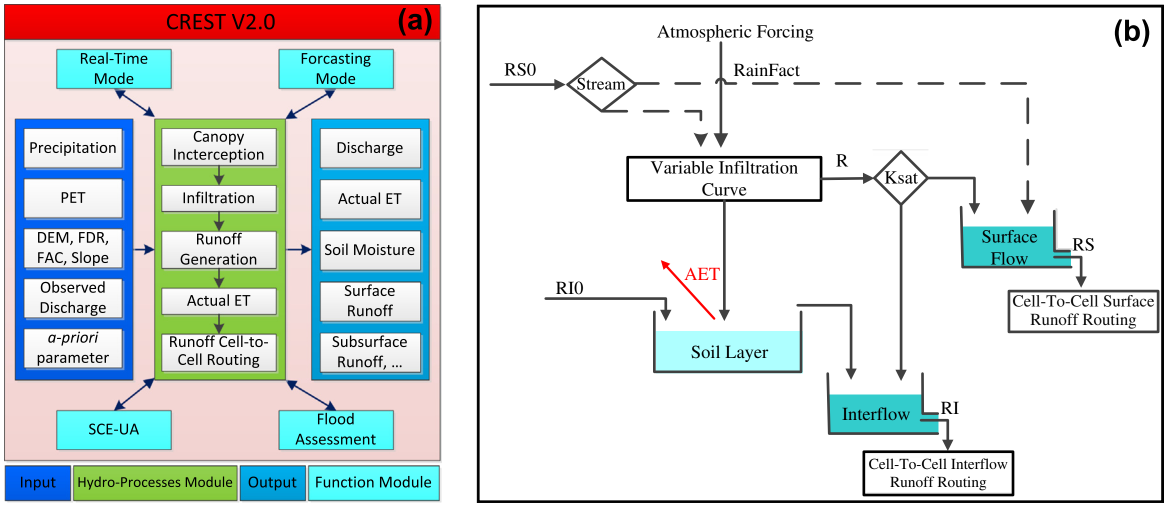

Fig. 2. (a) The framework of the CREST model version 2.0 and (b) vertical profile of hydrological processes in a grid cell.

and the effect on streamflow simulation, three widely used validation statistical indices were selected in this study. The relative Bias (%) was used to measure the agreement between the averaged value of simulated data (in this study, we call both TMPA products and simulated streamflow as "simulated data", "SIM" was used in the formulae) and observed data (such as rain gauge and observed streamflow in this study, "OBS" was used in the formulae). The root mean square error (RMSE) was selected to evaluate the average error magnitude between simulated and observed data. We also use correlation coefficient (CC) to assess the agreement between simulated and observed data.

$$Bias = \left[\frac{\sum_{i=1}^{n} SIM_i - \sum_{i=1}^{n} OBS_i}{\sum_{i=1}^{n} OBS_i}\right] \times 100 \tag{1}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(OBS_i - SIM_i)^2}{n}} \tag{2}$$

$$CC = \frac{\sum_{i=1}^{n}(OBS_i - \overline{OBS})(SIM_i - \overline{SIM})}{\sqrt{\sum_{i=1}^{n}(OBS_i - \overline{OBS})^2 \sum_{i=1}^{n}(SIM_i - \overline{SIM})^2}} \tag{3}$$

where n is the total number of pairs of simulated and observed data; i is the ith values of the simulated and observed data; $\underline{SIM}$ and $\overline{OBS}$ are the mean values of simulated and observed data respectively. Nash–Sutcliffe Coefficient of Efficiency (NSCE) is also used to assess the performance of model simulation and observation.

$$NSCE = 1 - \frac{\sum_{i=1}^{n}(OBS_i - SIM_i)^2}{\sum_{i=1}^{n}(OBS_i - \overline{OBS})^2} \tag{4}$$

Table 2
Parameters to be calibrated in CREST V2.0, their description, ranges and default values.

Parameter	Description	Numeric Range	Default Value
Ksat	Soil saturated hydraulic conductivity (mm/d)	0–2827.2	500
RainFact	Multiplier on the precipitation field	0.5–1.2	1.0
WM	Mean soil water capacity	80–200	120
B	Exponent of the variable infiltration curve	0.05–1.5	0.25
IM	Impervious area ratio	0–0.2	0.05
KE	Ratio of the PET to actual evapotranspiration	0.1–1.5	1.0
coeM	Overland runoff velocity coefficient	1–150	90
coeR	Multiplier used to convert overland flow speed to channel flow speed	1–3	2
coeS	Multiplier used to convert overland flow speed to interflow flow speed	0.001–1	0.3
KS	Overland reservoir discharge parameter	0–1	0.6
KI	Interflow reservoir discharge parameter	0–1	0.25

3. Results and discussion

3.1. Comparison of precipitation inputs

To better understand the impact of precipitation inputs on hydrologic models, the accuracy of the satellite precipitation against the in situ rain gauge observations should be assessed first. This section compares the TMPA and gauge observations over the time span of January 1, 2001–December 31, 2010 considering the basin-average precipitation and within a grid cell containing the dense rain gauge observations (Fig. 1). Fig. 3 shows that both 3B42V6 and 3B42V7 systematically underestimate, though at different levels, with biases of −41.15% and −8.38% and CCs of 0.36 and 0.40 at daily scale, respectively. Similar statistics are found at 0.25° grid-cell scale. Fig. 4 indicates that 3B42V6 largely underestimates with a bias of −40.25% and low CC of 0.37, while 3B42V7 has practically no bias (0.04%) and a relatively higher CC value 0.41.

Figs. 3(d) and 4(d) present the inter-comparison of monthly precipitation estimates to gain further information about the precision and variations at longer time scales. The monthly data for both basin-based and grid cell-based analyses were accumulated from daily data over the same time span from January 2001 to December 2010. At monthly time scale, both the basin-based and grid cell-based data show that 3B42V7 has better agreement with the monthly rain gauge data. Both Figs. 3 and 4 indicate that the latest V7 algorithm significantly corrects the underestimation bias in its predecessor version V6.

Fig. 5(a) and (b) show the frequency distribution of daily precipitation for different precipitation intensities (PI) for the basin-averaged and the grid cell-based precipitation time series, respectively. Fig. 5(a) shows that for the basin-averaged data, both 3B42V6 and 3B42V7 overestimate at the low PI range (less than 5 mm/day), but they underestimate at the medium and high PI ranges. However, 3B42V7 is in better agreement with the rain

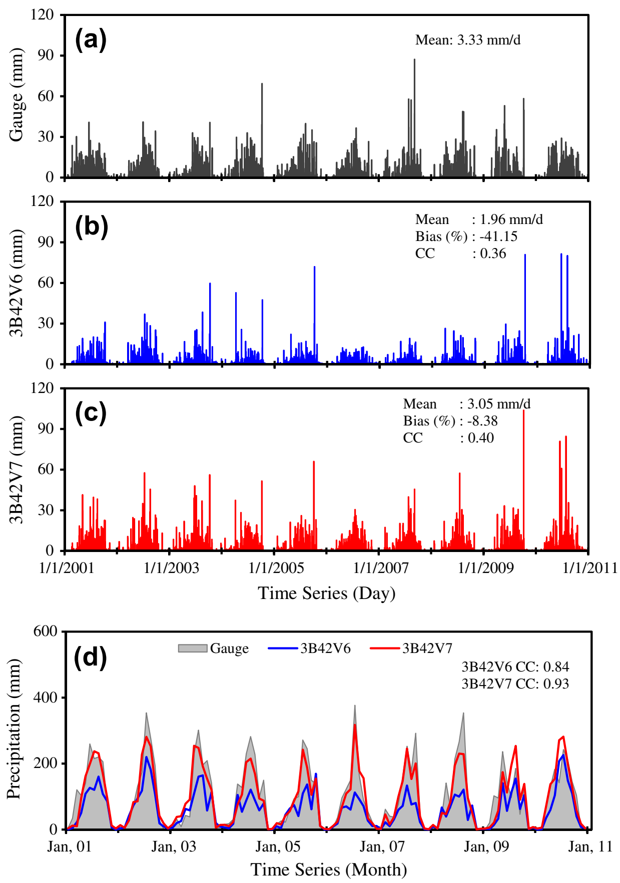

Fig. 3. Basin-averaged precipitation of (a) Gauge, (b) 3B42V6 and (c) 3B42V7 for the period January 2001–December 2010. Monthly data for both 3B42V6 and 3B42v7 are shown in (d).

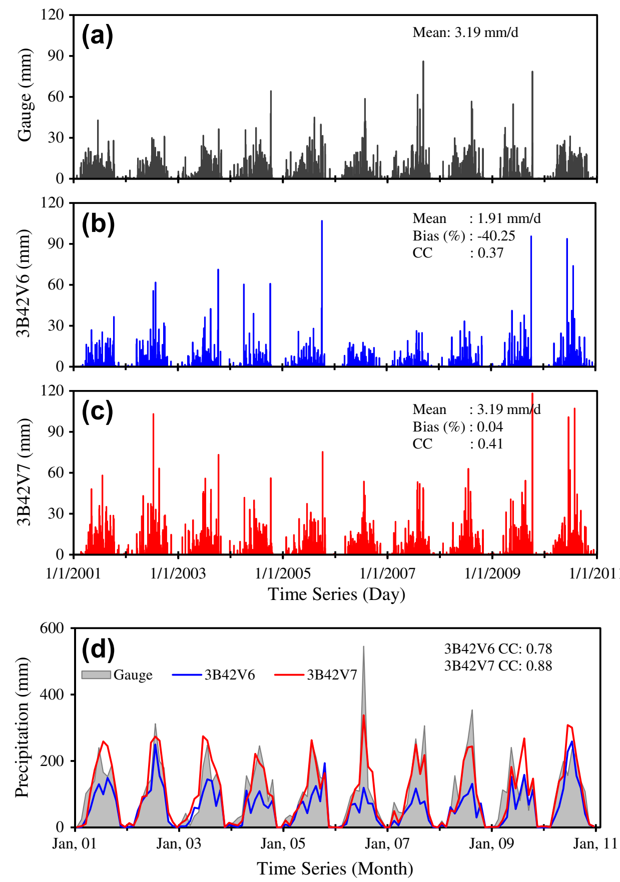

Fig. 4. As in Fig. 3, but for a single grid cell (Grid32).

gauge observations than 3B42V6 for the basin-averaged comparison across all PIs. Similarly, better agreement has been found in Fig. 5(b) for the new Version-7 products at the grid cell scale, except for values greater than 30 mm/day where there is overestimation.

3.2. Streamflow simulation scenarios

Although different precipitation products vary in accuracy and spatiotemporal resolutions, they might have similar hydrological prediction (i.e., streamflow simulation) skill after re-calibrating the model using the respective precipitation products (Jiang et al., 2012; Stisen and Sandholt, 2010). In the previous section, we compared the 3B42V6 and 3B42V7 precipitation products against the rain gauge observations; the next step is to evaluate how these two TMPA products affect streamflow simulations. Their hydrological evaluation is performed under two scenarios:

I. In situ gauge benchmarking: Calibrate the CREST model with 5 years of rain gauge data (January 2001 through December 2005). Then, replace the rain gauge forcing with precipitation from 3B42V6 and 3B42V7 for an independent validation period from January 2006 through December 2010 using the rain gauge-calibrated model parameters.
II. Product-specific calibration: Recalibrate the CREST model using the 3B42V6 and 3B42V7 precipitation data, respectively, over the same calibration period and then use the product-specific parameter sets to simulate streamflow over the same validation period as Scenario I.

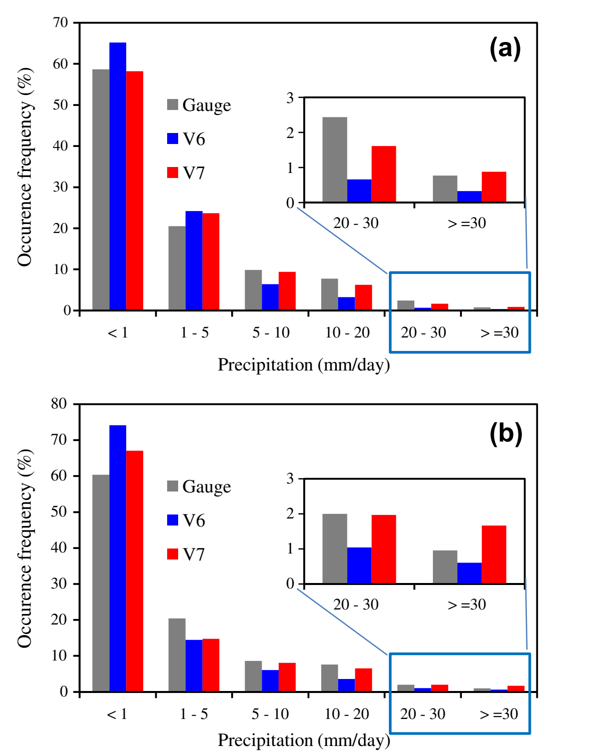

Fig. 5. Occurrence frequencies of rain gauge, 3B42V6 and 3B42V7 for (a) basin-averaged data and (b) single grid cell (Grid32).

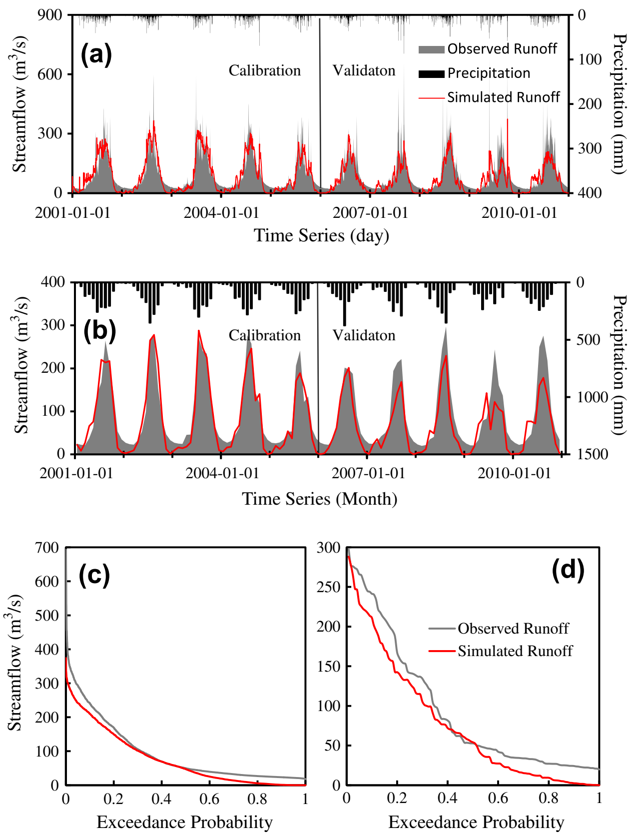

Fig. 6. Comparison of observed and simulated streamflow using gauge data as input: (a) daily calibration (2001.1.1–2005.12.31) and validation (2006.1.1–2010.12.31), (b) monthly data, c) exceedance probabilities using daily data from 2001.1.1 to 2010.12.31) and (d) exceedance probabilities using monthly data.

Scenario I, gauge benchmarking, is widely used by the hydrological community especially over gauged basins, while Scenario II is arguably deemed as an alternative for application to ungauged basins where only rainfall from remote-sensing platforms are available for use.

3.2.1. Scenario I: CREST benchmarked by in situ gauge data

(1) Rain gauge calibration and validation

The CREST model parameters are calibrated using rain gauge inputs for the period from January 2001 to December 2005 using the automatic calibration method (SCE-UA) by maximizing the NSCE value between the simulated and observed daily streamflow. The calibrated model is subsequently validated for the period from January 2006 to December 2010. Fig. 6 compares the simulated streamflow forced by rain gauge data with the observed streamflow in terms of time series plots and exceedance probability plots at daily and monthly scales. Fig. 6(a) and (b) show that general agreement exists between the observed and simulated streamflow. However, the simulated streamflow consistently underestimates the peaks, especially in the validation period and in relatively low flow seasons as well. The exceedance probabilities in Fig. 6(c) and (d) also show underestimation at low and high streamflow observations, while the simulations match relatively well in the intermediate ranges. As summarized in Tables 3 and 4, the statistical indices show that there is very good agreement between the simulated and observed hydrographs in the calibration period for both daily and monthly time scale, and reasonable simulations occurred in the validation period as well. Based on the criteria of the statistical indices in Moriasi et al. (2007), the model calibration and validation results indicate that the CREST model is well benchmarked by the in situ data at the daily and monthly time scale, so it can be used to evaluate the utility of the satellite precipitation products for hydrological prediction (i.e., streamflow) in this basin.

Table 3
Comparison of daily observed and simulated streamflow under two calibration scenarios.

Precipitation products	Scenario I				Scenario II			
	NSCE	Bias (%)	CC	RMSE	NSCE	Bias (%)	CC	RMSE
Calibration period								
Gauge	0.76	−9.73	0.89	45.38	–	–	–	–
3B42V6	0.23	−52.94	0.80	81.99	0.63	−1.70	0.80	56.55
3B42V7	0.66	−26.98	0.86	54.65	0.78	−4.81	0.88	43.94
Validation period								
Gauge	0.59	−29.59	0.83	57.85	–	–	–	–
3B42V6	0.17	−57.78	0.78	82.98	0.65	−8.67	0.81	54.00
3B42V7	0.63	−25.15	0.83	55.26	0.72	−3.02	0.86	47.72

Table 4
As in Table 3, but for monthly data.

Precipitation products	Scenario I				Scenario II			
	NSCE	Bias (%)	CC	RMSE	NSCE	Bias (%)	CC	RMSE
Calibration period								
Gauge	0.91	−9.75	0.96	25.18	–	–	–	–
3B42V6	0.25	−53.01	0.88	72.08	0.75	−1.66	0.87	41.41
3B42V7	0.77	−27.06	0.94	39.76	0.91	−4.83	0.95	25.41
Validation period								
Gauge	0.70	−29.59	0.88	43.63	–	–	–	–
3B42V6	0.19	−57.81	0.89	71.35	0.79	−8.65	0.89	36.29
3B42V7	0.80	−25.25	0.94	35.58	0.89	−3.08	0.95	26.53

(2) Impacts of satellite precipitation forcing

The gauge-benchmarked model is then forced by the TMPA 3B42V6 and 3B42V7 products from 2001 to 2010 using the model parameters calibrated by rain gauge data during the period from 2001 to 2005. Figs. 7 and 8 compare the daily and monthly time series of the simulated and observed hydrographs for both the calibration and validation periods. While 3B42V6 largely missed the high peak flows at both daily and monthly time series, 3B42V7 adequately captured a majority of the peak flows, especially at the smoothed monthly scale. The daily and monthly statistical comparisons in Tables 3 and 4 show that the daily and monthly simulations forced by rain gauge data had better skill (NSCE = 0.76/0.91, BIAS = −9.73%/−9.75%, CC = 0.89/0.96) than those based on 3B42V6 and 3B42V7 in the calibration period, which is expected. Interestingly, the 3B42V7-forced model simulations had very similar to and slightly better performance compared to the rain gauge-forced simulations in the validation period. A likely explanation is one of the rain gauge stations (i.e. the Dochula) had missing data from September 2006 to December 2010, which apparently degrades the hydrologic skill of this product. Overall, simulations forced by 3B42V7 were a significant improvement over 3B42V6. This clearly shows the improvements of the new version-7 algorithm upon its predecessor V6 products both statistically and now hydrologically.

3.2.2. Scenario II: CREST calibrated by individual TMPA products

To further assess the effects of TMPA 3B42 (V6 and V7) products on streamflow, the CREST model is recalibrated and validated with

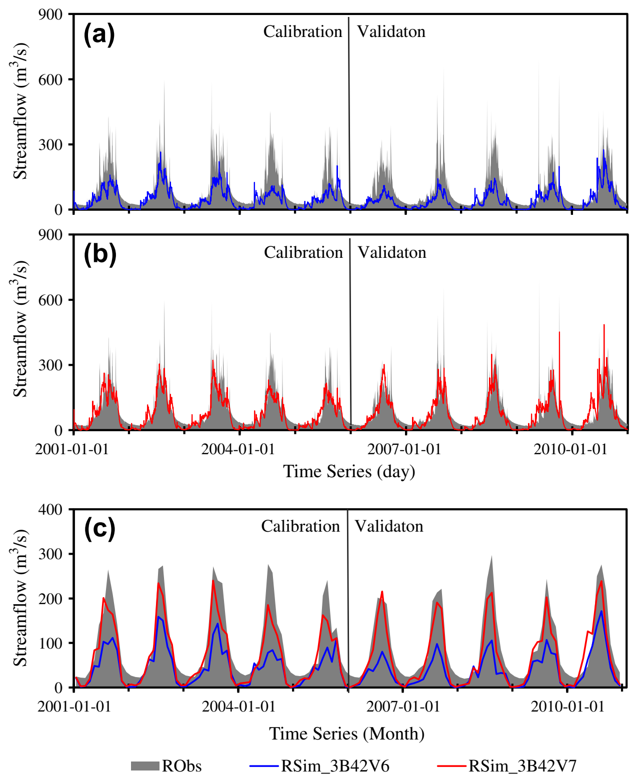

Fig. 7. Comparison of CREST simulated streamflow from 3B42V6 (blue line) and 3B42V7 (red line) with gauge-calibrated parameters and observed stream flow in both calibration (2001.1.1–20.5.12.31) and validation (2006.1.1–2010.12.31) period. (a) Daily data from 3B42V6; (b) Daily data from 3B42V7; (c) Monthly data from both 3B42V6 and 3B42V7. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

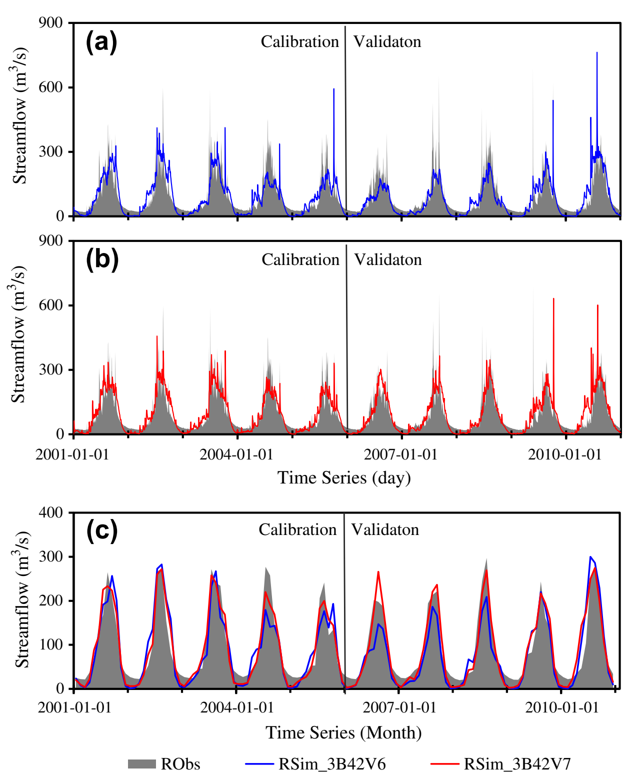

Fig. 8. As in Fig. 9, but the parameters were recalibrated using 3B42V6 and 3B42V7, respectively.

3B42V6 and 3B42V7 for the same period as Scenario I. This scenario is often used as an alternative strategy for remote sensing precipitation over ungauged basins. As shown in Fig. 8, all simulations are significantly improved after the recalibration, and they capture most of the daily and monthly peak flows. Comparatively, the CREST model simulations based on 3B42V7 inputs have better skill than those based on 3B42V6. As summarized in Tables 3 and 4, simulations have good statistical agreement with observed streamflow at daily and monthly scale.

The statistical indices of daily NSCE, Bias and CC in Table 3 were selected for visual comparison of the two modeling scenarios. Fig. 9 indicates that the product-specific recalibration in Scenario II has obviously improved the NSCE and CC values and reduced the Bias values for both the calibration and validation periods. It is noted that the recalibration forcing with 3B42V7 in Scenario II has much higher NSCE and smaller Bias than 3B42V6, and very comparable CC values, all of which improved upon the rain gauge-benchmarked model.

3.3. Discussion of parameter compensation effect from Scenario II

Table 5 shows the optimum parameter sets forced by 3B42V6 and 3B42V7, relative to the gauge forcing, for the calibration period from 2001 to 2005 using the SCE-UA algorithm. Note that the parameter values of Ksat and WM are spatially distributed but have been basin-averaged and summarized in Table 5. It shows

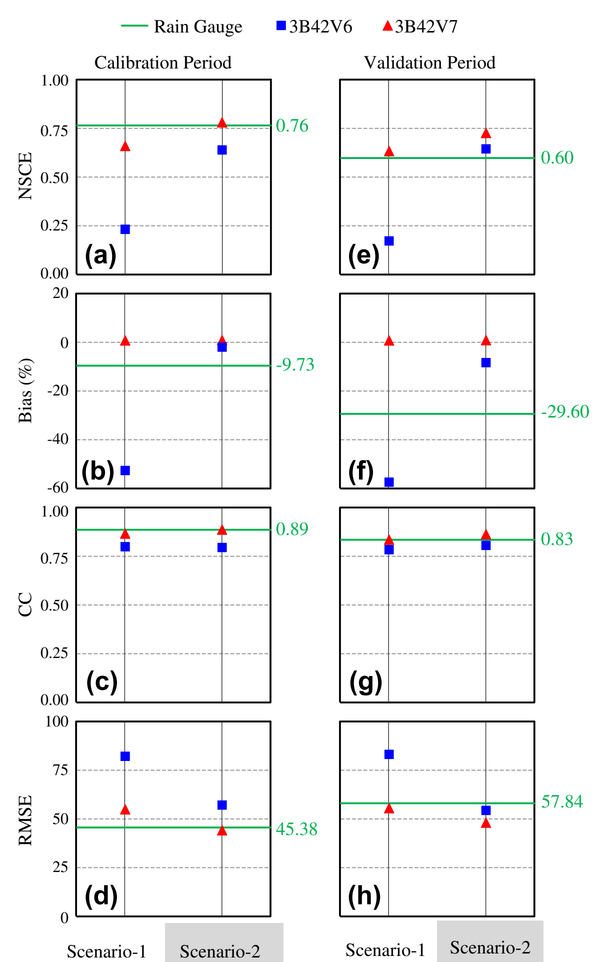

Fig. 9. Comparison of the streamflow performance statistics of the TMPA 3B42V6 and 3B42V7 precipitation for the two simulation scenarios in both calibration period (a–d) and validation period (e–h).

Table 5
CREST model parameter values calibrated with different precipitation inputs for the calibration period of January 2001–December 2005.

Parameters	Gauge	3B42V6	3B42V7
RainFact	0.87	1.34	0.98
Ksat	56.90	33.09	52.73
WM	166.50	142.71	166.52
B	1.48	1.48	1.48
IM	0.20	0.20	0.20
KE	0.10	0.05	0.13
coeM	88.05	63.67	67.95
coeR	2.68	1.33	1.44
coeS	0.43	0.47	0.67
KS	0.99	0.71	0.78
KI	0.20	0.13	0.14

that 3B42V7-calibrated parameters have less deviation from the gauge-calibrated parameter values than 3B42V6. For example, RainFact is the adjustment factor of the precipitation either due to canopy interception or bias. Table 5 shows that 3B42V6 increases the RainFact value from 0.87 to 1.34, to compensate its underestimation as shown in Figs. 3 and 4, while 3B42V7's estimated value (0.98) is closer to 1 and the Gauge value (0.87). Another example is KE, the ratio of potential evapotranspiration to the satellite PET data. Table 5 reveals that 3B42V6 demands a reduced KE value from 0.10 to 0.05 in order to partition more precipitation into runoff while 3B42V7 only slightly increases it from 0.10 to 0.13, possibly to partially offset the above RainFact increase, amongst other parametric interactions. The third example is Ksat, the soil saturated hydraulic conductivity. Table 5 shows that the Ksat of 3B42V6 reduced from 56.90 to 33.09 while V7 only changed slightly from 56.90 to 52.73. Regarding the mean water capacity, WM, 3B42V6 decreased from 166.50 to 142.71 to hold less water in the soils while 3B42V7 did not change much from the gauge-calibrated value, which is presumably closer to the truth. It also shows the overland flow coefficient, coeM, the average channel flow speed, coeR, the overland flow recession coefficient, KS, and the interflow recession coefficient, KI, all had reduced values to retain more water in the river basin after recalibrating the parameters to both of the satellite products. Not surprisingly, Table 5 also shows some opposite changes of values such as KE for 3B42V7 and coeS, the surface-interflow conversion factor, for both 3B42V6 and 3B42V7, resulting in a slight decrease in streamflow.

In addition to the analysis of the parameters properties, water balance analysis is another important indicator for analyzing the effect of the parameter recalibration. Thus the difference of water balance components over 10-year (2001–2010) simulations is further examined using rain gauge and TMPA 3B42 rainfall, respectively. In CREST model, the water balance budgeting partitions the precipitation after canopy interception into actual evapotranspiration (ET), runoff depth (i.e. streamflow) and water storage change (ΔS), as shown in Fig. 10. As expected, precipitation is the dominant runoff generation input so in Fig. 3, all satellite rainfall forced simulations underestimated the streamflow compared to rain gauge results in scenario I. However, in scenario II, the model was recalibrated with respective satellite rainfall, the increased partition of the satellite driven streamflow simulations comes at the expense of a significant decrease of water storage due to the effect of the parameters value changes (shown in Fig. 10). In the gauge rainfall driven simulations, 27.90% of precipitation will be stored in this basin, however, 26.43% (27.18%) of precipitation is water storage in scenario I while 8.95% (16.09%) in scenario II for 3B42V6 (3B42V7).

From the above discussion, it is clear that the overall effect of the recalibrated parameter sets is to largely compensate for rainfall

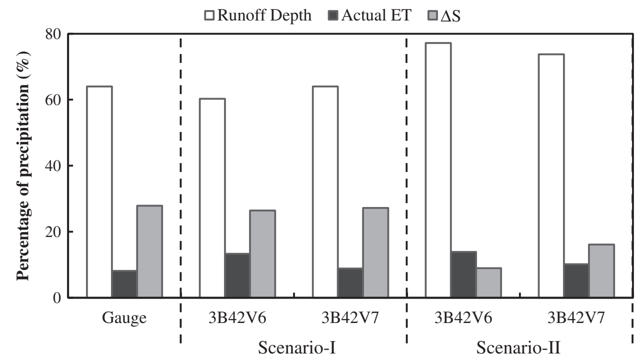

Fig. 10. Relative change of the water balance components using rain gauge and satellite rainfall based on 10-year annual averages (2001–2010) hydrological simulations in scenarios I and II.

underestimation in 3B42V6 while less so for 3B42V7. The effect of arriving at a very similar simulation with different combinations of parameter settings has been called "Equifinality" of the hydrological model (Aronica et al., 1998; Beven and Freer, 2001; Zak and Beven, 1999). This study clearly shows how different parameter settings can compensate for errors in the satellite rainfall forcing and can thus improve model predictions of streamflow. It is possible that the current model structural deficiency, i.e., not accounting for snowmelt process, is compensated by the model re-calibration. However, this parameter compensation effect comes with the price of having a locally optimized model with parameter values unrepresentative of reality. This might limit the model's predictive capability at internal sub-basins, or under different initial conditions. This is particularly concerning under scenarios involving climate change. In any case, the recalibration strategy could be especially problematic for 3B42V6 (Bitew and Gebremichael, 2011; Jiang et al., 2012), however the 3B42V7 product gives higher confidence for use in ungauged basins even without the need for recalibration.

4. Summary and conclusions

Satellite precipitation products are very important for regional and global hydrological studies, particularly for remote regions and developing countries. This study first focuses on statistically assessing the accuracy of the TMPA 3B42V6 product vs. its latest successive version 3B42V7, and then hydrologically evaluates their streamflow prediction utility using the CREST distributed hydrologic model in the mountainous Wangchu Basin of Bhutan.

The two versions of TMPA satellite products are statistically compared with a decade-long (2001–2010) rain gauge dataset at daily and monthly scales. In general, 3B42V7 consistently improves upon 3B42V6's underestimation both for the whole basin (bias improved from −41.15% to −8.38%) and for a $0.25^\circ \times 0.25^\circ$ grid cell with high-density gauges (bias improved from −40.25% to 0.04%), though with modest enhancement of correlation coefficients (CC) (from 0.36 to 0.40 for entire basin and from 0.37 to 0.41 for the grid cell). 3B42V7 also improves upon 3B42V6 in terms of occurrence frequency across the rain intensity spectrum. Apparently the results show that the new algorithm 3B42V7 has much improved accuracy upon 3B42V6, in concert with other studies in different areas (Chen et al., 2013a,b; Kirstetter et al., 2013). The improvement from V6 to V7 is mainly a combination of three factors: (1) the enhanced TMPA Level-2 retrieval algorithms (Chen et al., 2013a; Kirstetter et al., 2013), (2) incorporation of the global gauge network (i.e. GPCC) data with improved climatology and anomaly analysis (Huffman et al., 2011), and (3) additional satellite observations incorporated (Huffman and Bolvin, 2012).

For the hydrological evaluation, two scenario-based calibration and validation experiments are conducted over the same 10-year time span. Scenario I, in situ gauge benchmarking, is widely used by the hydrological community especially over gauged basins, while Scenario II, input-specific recalibration, is arguably deemed as an alternative for application to ungauged basins where only remote-sensing rainfall data may be available for use. In Scenario I, the 3B42V6-based simulation shows lower hydrologic prediction skill in terms of NSCE (0.23 at daily scale and 0.25 at monthly scale) while 3B42V7 performs fairly well (0.66 at daily scale and 0.77 at monthly scale), a comparable skill score with the simulations using the gauge benchmark. For the precipitation-specific calibration in Scenario II, significant improvements are observed for 3B42V6 across all statistics. These enhancements are not as obvious for the already-well-performing 3B42V7-calibrated model, except for some reduction in bias (from −26.98% to −4.81%). This behavior is consistent with previous studies (Bitew and Gebremichael, 2011; Bitew et al., 2011; Jiang et al., 2012). This study offers unique insights into 3B42V6 and 3B42V7 products in a mountainous South Asian basin.

In concert with several other studies by Chen et al. (2013a) and Kirstetter et al. (2013) in the US and Chen et al., 2013b in the tropics, this study also reveals the latest 3B42V7 algorithm has a noticeable improvements from 3B42V6 both in terms of accuracy (i.e., correcting the underestimation) and in its promising hydrological, even with or without recalibration of the hydrological model with respective rainfall inputs. The parameter compensation effect is often recognized but still used by the hydrology community. This approach has been noted to be problematic due to unrealistic parameter settings which may ultimately limit the model's predictive capability under conditions of climate change and differing initial conditions.

Acknowledgements

The current study was supported by the NASA/Marshall Space Flight Center Grants NNM11AB34P and NNMi2428088Q to the University of Oklahoma.

References

Aronica, G., Hankin, B., Beven, K.J., 1998. Uncertainty and equifinality in calibrating distributed roughness coefficients in a flood propagation model with limited data. Adv. Water Resour. 22 (4), 349–365.
Beven, K.J., Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. J. Hydrol. 249 (1–4), 11–29.
Bitew, M.M., Gebremichael, M., 2011. Evaluation of satellite rainfall products through hydrologic simulation in a fully distributed hydrologic model. Water Resour. Res. 47 (6), W06526.
Bitew, M.M., Gebremichael, M., Ghebremichael, L.T., Bayissa, Y.A., 2011. Evaluation of high-resolution satellite rainfall products through streamflow simulation in a hydrological modeling of a small mountainous watershed in Ethiopia. J. Hydrometeorol. 13 (1), 338–350.
Bookhagen, B., Burbank, D.W., 2010. Toward a complete Himalayan hydrological budget: spatiotemporal distribution of snowmelt and rainfall and their impact on river discharge. J. Geophys. Res. 115 (F3), F03019.
Chen, Sheng et al., in press. Evaluation of the successive V6 and V7 TRMM multi-satellite precipitation analysis over the continental United States. Water Resour. Res.
Chen, Y., Ebert, E.E., Walsh, K.J.E., Davidson, N.E., 2013. Evaluation of TRMM 3B42 precipitation estimates of tropical cyclone rainfall using PACRAIN Data. J. Geophys. Res.: Atmos. 118 (5), 2184–2196.
Chokngamwong, R., Chiu, L.S., 2008. Thailand daily rainfall and comparison with TRMM products. J. Hydrometeorol. 9 (2), 256–266.
Chow, V.T., Maidment, D.R., Mays, L.W., 1988. Applied Hydrology: McGraw-Hill Series in Water Resources and Environmental Engineering. McGraw-Hill, Inc., New York.
Duan, Q., Sorooshian, S., Gupta, V., 1992. Effective and efficient global optimization for conceptual rainfall-runoff models. Water Resour. Res. 28 (4), 1015–1031.
Duan, Q.Y., Gupta, V.K., Sorooshian, S., 1993. Shuffled complex evolution approach for effective and efficient global minimization. J. Optim. Theory Appl. 76 (3), 501–521.
FAO/IIASA/ISRIC/ISSCAS/JRC, 2009. Harmonized World Soil Database (version 1.1), FAO, Rome, Italy and IIASA, Laxenburg, Austria.
Hansen, M.C., Defries, R.S., Townshend, J.R.G., Sohlberg, R., 2000. Global land cover classification at 1 km spatial resolution using a classification tree approach. Int. J. Remote Sens. 21 (6), 1331–1364.
Hong, Y., Hsu, K.-L., Sorooshian, S., Gao, X., 2004. Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. J. Appl. Meteorol. 43 (12), 1834–1853.
Huffman, G.J., Bolvin, D.T., Nelkin, E., Adler, R.F., 2011. Highlights of Version 7 TRMM Multi-satellite Precipitation Analysis (TMPA). In: C.K.a.G.H. Eds (Editor), 5th Internat. Precip. Working Group Workshop, Workshop Program and Proceedings. Reports on Earth Sys. Sci., 100/2011, Max-Planck-Institut für Meteorologie, Hamburg, Germany.
Huffman, G.J., Bolvin, D.T., 2012. TRMM and other Data Precipitation Data set Documentation. Laboratory for Atmospheres, NASA Goddard Space Flight Center and Science Systems and Applications.
Huffman, G.J. et al., 2007. The TRMM Multisatellite Precipitation Analysis (TMPA): quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. J. Hydrometeorol. 8 (1), 38–55.
Islam, M.N., Uyeda, H., 2007. Use of TRMM in determining the climatic characteristics of rainfall over Bangladesh. Remote Sens. Environ. 108 (3), 264–276.
Jamandre, C.A., Narisma, G.T., 2013. Spatio-temporal validation of satellite-based rainfall estimates in the Philippines. Atmospheric Research 122, 599–608.
Jiang, S. et al., 2012. Comprehensive evaluation of multi-satellite precipitation products with a dense rain gauge network and optimally merging their simulated hydrological flows using the Bayesian model averaging method. J. Hydrol. 452–453, 213–225.
Joyce, R.J., Janowiak, J.E., Arkin, P.A., Xie, P., 2004. CMORPH: a method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. J. Hydrometeorol. 5 (3), 487–503.
Khan, S.I. et al., 2011a. Hydroclimatology of Lake Victoria region using hydrologic model and satellite remote sensing data. Hydrol. Earth Syst. Sci. 15 (1), 107–117.
Khan, S.I. et al., 2011b. Satellite remote sensing and hydrologic modeling for flood inundation mapping in Lake Victoria basin: implications for hydrologic prediction in ungauged basins. Geosci. Remote Sens. IEEE Trans. 49 (1), 85–95.
Kirstetter, P.-E., Hong, Y., Gourley, J.J., Schwaller, M., Petersen, W., Zhang, J., 2013. Comparison of TRMM 2A25 Products, Version 6 and Version 7, with NOAA/NSSL Ground Radar–Based National Mosaic QPE. Journal of Hydrometeorology 14, 661–669.
Li, X.-H., Zhang, Q., Xu, C.-Y., 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.
Liang, X., Lettenmaier, D.P., Wood, E.F., Burges, S.J., 1994. A simple hydrologically based model of land surface water and energy fluxes for general circulation models. J. Geophys. Res. 99 (D7), 14415–14428.
Liang, X., Wood, E.F., Lettenmaier, D.P., 1996. Surface soil moisture parameterization of the VIC-2L model: Evaluation and modification. Global Planet. Change 13 (1–4), 195–206.
Mishra, A., Gairola, R.M., Varma, A.K., Agarwal, V.K., 2010. Remote sensing of precipitation over Indian land and oceanic regions by synergistic use of multisatellite sensors. J. Geophys. Res. 115 (D8), D08106.
Moriasi, D.N. et al., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. ASABE 50 (3), 885–900.
Sorooshian, S. et al., 2000. Evaluation of PERSIANN system satellite–based estimates of tropical rainfall. Bull. Am. Meteorol. Soc. 81 (9), 2035–2046.
Stisen, S., Sandholt, I., 2010. Evaluation of remote-sensing-based rainfall products through predictive capability in hydrological runoff modelling. Hydrol. Process. 24 (7), 879–891.
Su, F., Hong, Y., Lettenmaier, D.P., 2008. Evaluation of TRMM Multisatellite Precipitation Analysis (TMPA) and its utility in hydrologic prediction in the La Plata Basin. J. Hydrometeorol. 9 (4), 622–640.
Wang, J. et al., 2011. The coupled routing and excess storage (CREST) distributed hydrological model. Hydrol. Sci. J. 56 (1), 84–98.
Wu, H., Adler, R.F., Hong, Y., Tian, Y., Policelli, F., 2012. Evaluation of global flood detection using satellite-based rainfall and a hydrologic model. J. Hydrometeorol. 13 (4), 1268–1284.
Yong, B. et al., 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.
Yong, B. et al., 2010. Hydrologic evaluation of Multisatellite Precipitation Analysis standard precipitation products in basins beyond its inclined latitude band: a case study in Laohahe basin, China. Water Resour. Res. 46 (7), W07542.
Zak, S.K., Beven, K.J., 1999. Equifinality, sensitivity and predictive uncertainty in the estimation of critical loads. Sci. Total Environ. 236 (1–3), 191–214.
Zhao, R.J., 1992. The Xinanjiang model applied in China. J. Hydrol. 135 (1–4), 371–381.
Zhao, R.J., Zhuang, Y.G., Fang, L.R., Liu, X.R., Zhang, Q.S., 1980. The Xinanjiang Model, Hydrological Forecasting Proceedings Oxford Symposium. IAHS Publication, Oxford University, pp. 351–356.