**Hydrology and
Earth System
Sciences**

Assessment of satellite rainfall products for streamflow simulation in medium watersheds of the Ethiopian highlands

M. M. Bitew and M. Gebremichael

Department of Civil and Environmental Engineering, University of Connecticut, Storrs, Connecticut, USA

Abstract. The objective is to assess the suitability of commonly used high-resolution satellite rainfall products (CMORPH, TMPA 3B42RT, TMPA 3B42 and PERSIANN) as input to the semi-distributed hydrological model SWAT for daily streamflow simulation in two watersheds (Koga at 299 km^2 and Gilgel Abay at 1656 km^2) of the Ethiopian highlands. First, the model is calibrated for each watershed with respect to each rainfall product input for the period 2003–2004. Then daily streamflow simulations for the validation period 2006–2007 are made from SWAT using rainfall input from each source and corresponding model parameters; comparison of the simulations to the observed streamflow at the outlet of each watershed forms the basis for the conclusions of this study. Results reveal that the utility of satellite rainfall products as input to SWAT for daily streamflow simulation strongly depends on the product type. The 3B42RT and CMORPH simulations show consistent and modest skills in their simulations but underestimate the large flood peaks, while the 3B42 and PERSIANN simulations have inconsistent performance with poor or no skills. Not only are the microwave-based algorithms (3B42RT, CMORPH) better than the infrared-based algorithm (PERSIANN), but the infrared-based algorithm PERSIANN also has poor or no skills for streamflow simulations. The satellite-only product (3B42RT) performs much better than the satellite-gauge product (3B42), indicating that the algorithm used to incorporate rain gauge information with the goal of improving the accuracy of the satellite rainfall products is actually making the products worse, pointing to problems in the algorithm. The effect of watershed area on the suitability of satellite rainfall products for streamflow simulation also depends on the rainfall product. Increasing the watershed area from 299 km^2 to 1656 km^2 improves the simulations obtained from the 3B42RT and CMORPH (i.e. products that are more reliable and consistent) rainfall inputs while it deteriorates the simulations obtained from the 3B42 and PERSIANN (i.e. products that are unstable and inconsistent) rainfall inputs.

1 Introduction

Prediction of streamflow simulation in ungauged basins of the East African highlands is a challenging task due to the absence of reliable ground-based rainfall information. The region has no any ground-based radar for rainfall measurement, the rain gauge network is very sparse, and countries in the downstream of transboundary river basins have no access to the existing upstream rain gauge information. Can high-resolution satellite-based rainfall estimates provide reliable rainfall information for streamflow simulation application in this region?

During the last two decades, satellite-based instruments have been designed to collect observations mainly at thermal infrared (IR) and microwave (MW) wavelengths that can be used to estimate rainfall rates. Observations in the IR band are available in passive modes from (near) polar-orbiting (revisit times of 1–2 days) and geostationary orbits (revisit times of 15–30 min), while observations in the passive and active MW band are only available from the (near) polar-orbiting satellites. A number of algorithms have been developed to estimate rainfall rates by combining information from the more accurate (but infrequent) MW with the more frequent (but less accurate) IR to take advantage of the complementary strengths. The TMPA method (Huffman et al., 2007) uses MW data to calibrate the IR-derived estimates and creates estimates that contain MW-derived rainfall estimates when and where MW data are available and the calibrated IR estimates where MW data are not available. The TMPA products are available in two versions:

real-time version (3B42RT) and post-real-time research version (3B42). The main difference between the two versions is the use of monthly rain gauge data for bias adjustment in the post-real-time research product. The 3B42 products are released 10–15 days after the end of each month, while the 3B42RT are released about 9 h after overpass. The CMORPH method (Joyce et al., 2004) obtains the rainfall estimates from MW data but uses a tracking approach in which IR data are used only to derive a cloud motion field that is subsequently used to propagate raining pixels. The PERSIANN method (Sorooshian et al., 2000) uses a neural network approach to derive relationships between IR and MW data which are applied to the IR data to generate rainfall estimates. The resolutions of these (often dubbed as 'high-resolution') products are 0.25° and 3 hourly, although finer resolutions are also available for CMORPH and PERSIANN. Besides these widely known products, there are also other high-resolution products, such as, Hydro-estimator (Scofield and Kuligowski, 2003), NRL-blended (Turk and Miller, 2005), PMIR (Kidd and Muller, 2009), and GSMaP (Ushio and Kachi, 2009).

It is well known that the satellite rainfall values are just estimates that are subject to a variety of error sources (gaps in revisit times, poor direct relationship between remotely sensed signals and rainfall rate, atmospheric effects that modify the radiation field) and require a thorough validation. The validation efforts can be grouped into two categories. The first is the direct comparison of the satellite rainfall estimates to the rain gauge networks and ground-based radar estimates (Dinku et al., 2007, 2008; Hirpa et al., 2010; Bitew and Gebremichael, 2010). The second is the evaluation of satellite rainfall estimates based on their predictive ability of streamflow rate in a hydrological modeling framework. This has two advantages. One, since the evaluation is performed at the watershed scale, it is not subject to the scale discrepancy problem that arises when using rain gauge data for validation. Two, the satellite rainfall estimates are evaluated with respect to a specific application, as a driving input variable in a hydrologic model.

The purpose of this study is to assess the capability and limitation of satellite rainfall products as input into a hydrological model for streamflow simulation in a mountainous and medium-size watershed in Ethiopia, for four different satellite rainfall products and two different watershed sizes. This study is limited to the following specific cases: the Soil and Water Assessment Tool (SWAT) hydrological model, two watersheds (299 km^2, and 1656 km^2), and four satellite precipitation products (TMPA 3B42, TMPA 3B42RT, CMORPH, and PERSIANN). SWAT is a semi-distributed hydrological model widely used for research and application according to Gassman et al. (2007) over 250 peer-reviewed journal articles existed by 2007 on SWAT-related work. SWAT was also successfully used to model Ethiopian Highland watersheds in previous studies (e.g., Easton et al., 2010).
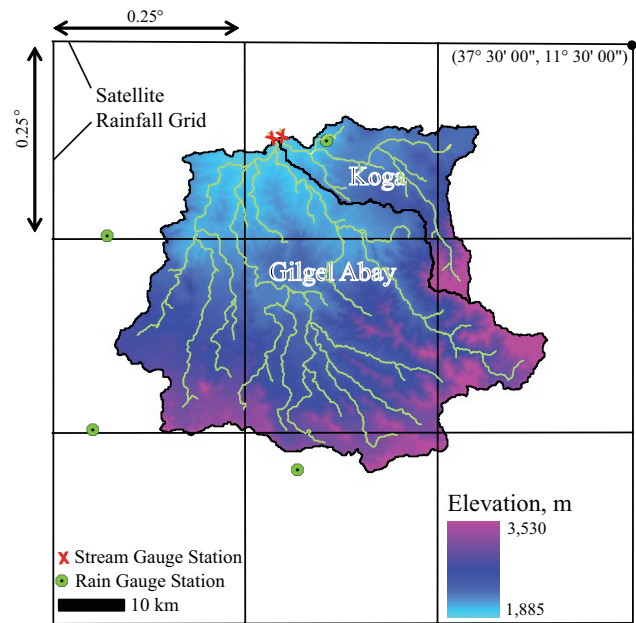

Fig. 1. The study region in Ethiopian highlands consisting of two adjoining watersheds: Koga (299 km^2) and Gilgel Abay (1656 km^2). Also shown are satellite rainfall grids (0.25° × 0.25°) and locations of four rain gauge stations in the study region, and two stream gauge stations at the outlets of the watersheds.

2 Data and method

2.1 Study region

The study region consists of two gauged adjoining watersheds (Koga and Gilgel Abay) in the Ethiopian part of the East African highlands (Fig. 1). Koga watershed has a drainage area of 299 km^2 and is located within 37°2′ E–37°20′ E and 11°8′ N–11°25′ N, and Gilgel Abay has a drainage area of 1656 km^2 and is located within 36°48′ E–37°24′ E and 10°56′ N–11°23′ N. The climate is semi-humid with a mean annual rainfall of 1300 mm, more than 70% of which falls in the summer monsoon season. The watersheds have similar landscape characteristics: complex topography with elevations ranging from 1890 m to 3130 m (for Koga), and 1880 m to 3530 m (for Gilgel Abay); land use characterized by cropland, pasture and forest shrubs (55%, 20%, and 25%, respectively for Koga, and 74%, 15%, and 11%, respectively, for Gilgel Abay), and soils characterized by clay, clay loam and silt loam (42%, 39%, and 19%, respectively for Koga, and 33%, 34%, and 33%, respectively, for Gilgel Abay). There are four rain gauges in the study region, and a stream gauge at the outlet of each watershed. These rain gauges were not used in the derivation of the TMPA 3B42 products.
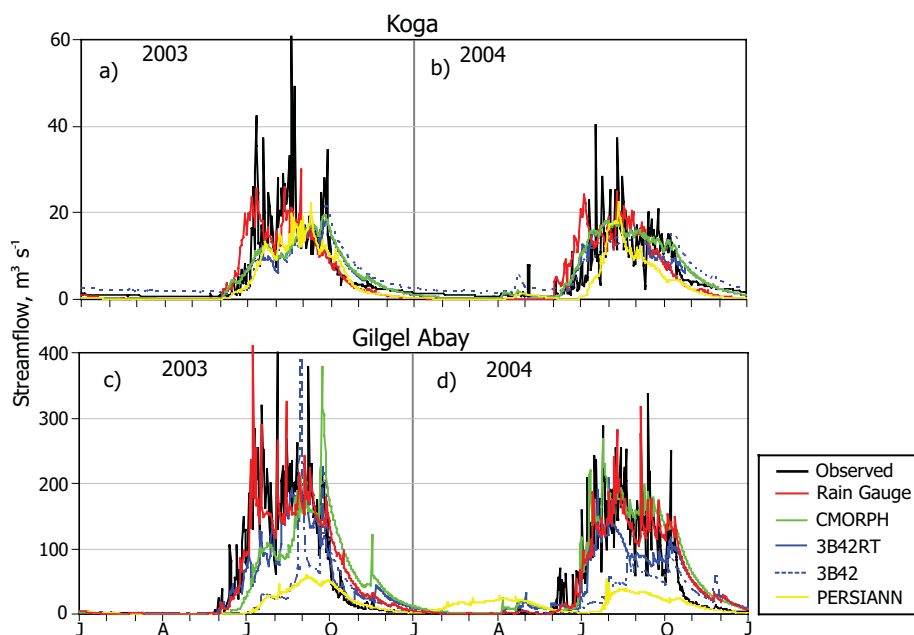
Fig. 2. Comparison of SWAT simulated (based on CMORPH, 3B42RT, 3B42, PERSIANN, and rain gauge network rainfall inputs, separately) and observed daily streamflow hydrographs during the calibration period of 2003 through 2004, for **(a, b)** Koga and **(c, d)** Gilgel Abay watersheds. In the legend, "Observed" indicates the observed streamflow, while the others (Rain Gauge, CMORPH, 3B42RT, 3B42, PERSIANN) indicate the source of rainfall data used in SWAT simulation.

2.2 Hydrological model and calibration

2.2.1 SWAT hydrological model

SWAT, developed by the United States Department of Agriculture (USDA) – Agricultural Research Service (ARS) (Arnold et al., 1998), is a continuous, semi-distributed hydrologic model that runs on a daily time step. Hydrologic response units (HRUs), defined by combinations of land cover and soil combinations, are the computational elements of SWAT. The daily water budget in each HRU is computed based on daily precipitation, runoff, evapotranspiration, percolation, and return flow from the subsurface and groundwater flow. Runoff volume in each HRU is computed using the Soil Conservation Service (SCS) curve number method (SCS, 1986). A complete description of the SWAT model can be found in Arnold et al. (1998). We obtained the following SWAT inputs: elevation data from the 30-m USGS NED digital elevation model dataset (http://hydrosheds.cr.usgs.gov), soil texture from the FAO Soil and terrain data map of East Africa (SEA) dataset, land use from the Ethiopian Woody Biomass Inventory Strategic Planning Project, meteorological data from the nearby meteorological stations of the National Meteorological Agency of Ethiopia, and rainfall data from satellite rainfall estimates and rain gauge measurements.

2.2.2 Parameter specification and calibration

Automatic calibration of all the SWAT model parameters could be time consuming and less practical (Eckhardt and Arnold, 2001). In order to reduce the number of calibration parameters, we performed sensitivity analysis using the LH-OAT method available within SWAT, which combines the Latin Hypercube (LH) sampling method with the One-factor-At-a-Time (OAT) method (Van Griensven et al., 2006). We found nine most sensitive parameters, and focused our automatic and manual calibration exercise on these parameters. Our objective function was maximizing the Nash-Sutcliffe efficiency between simulated and measured daily streamflow.

We calibrated the model parameters for each watershed and rainfall input source, separately, over a two-year period (2003–2004) by comparing the simulated and observed daily streamflow hydrographs. The resulting model parameter estimates are shown in Table 1. There are large differences in the parameter estimates obtained from the different rainfall inputs. Comparison between simulated and observed streamflow hydrographs is shown in Fig. 2. In general, the simulation results are satisfactory for Koga. For Gilgel Abay, the calibration results for the 3B42RT, CMORPH and rain gauge simulations show satisfactory calibration results, whereas the results for 3B42 and PERSIANN are not satisfactory. As can be seen from Fig. 3b, the 3B42 and PERSIANN products give annual rainfall estimates that are lower than the annual streamflow estimates, while the other rainfall products

Table 1. SWAT model parameter estimates for each watershed and rainfall input source.

Parameter Type	Model Parameter	Variable	Unit	Parameter Values for Koga (Gilgel Abay)				
				Rain gauge	CMORPH	3B42RT	PERSIANN	3B42
Routing	Hydraulic conductivity of main channel alluvium	CH_K2	mm h^{-1}	1.1 (60)	23.6 (60)	29.3 (60)	143.7 (60)	0.01 (60)
Groundwater	Base flow alpha factor	Alpha_BF	day^{-1}	0.97 (0.75)	0.97 (0.75)	0.97 (0.75)	0.97 (0.75)	0.97 (0.75)
HRU	Curve number	CN2*	–	62 (50)	69 (57)	69 (57)	73 (72)	72 (67)
Basin	Surface runoff lag coefficient	Surlag	–	0.001 (8)	0.001 (8)	0.001 (8)	8.94 (8)	0.001 (0.1)
Routing	Manning's "n" value for main channel	CH_N2	–	0.04 (0.04)	0.02 (0.04)	0.01 (0.04)	0.06 (0.04)	0.116 (0.04)
HRU	Soil hydraulic conductivity	Sol_K*	mm hr^{-1}	0.02 (0.0175)	0.02 (0.0175)	0.02 (0.0175)	0.02 (0.0175)	0.02 (0.0175)
HRU	soil evaporation compensation factor	ESCO		0.92 (0)	1 (0)	0.99 (0)	1 (0)	1 (0)
HRU	Maximum canopy storage	canmx	–	2.34 (2.5)	2.34 (0)	1.44 (0)	0 (0)	0.39 (0)
Groundwater	Deep aquifer percolation fraction	Rchrg_dp	–	0.01 (0.25)	0.01 (0)	0.01 (0)	0.01 (0)	0.01 (0)
Groundwater	Groundwater delay	Gw_delay	day	31 (25)	31 (25)	31 (25)	31 (25)	31 (25)
Groundwater	Threshold depth of water in the shallow aquifer for return flow to occur	Gwqmn	mm	0 (1500)	0 (0)	0 (0)	0 (0)	0 (0)

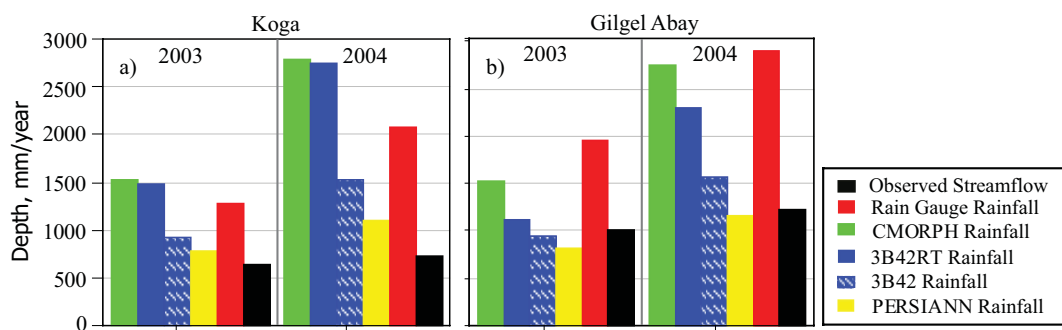* Values represent average values of spatial distribution.

Fig. 3. Comparison of annual rainfall depth derived from each rainfall source (CMORPH, 3B42, 3B42RT, PERSIANN, and rain gauges) to annual observed streamflow at the outlet of each watershed, for **(a)** Koga and **(b)** Gilgel Abay watershed.

give rainfall estimates substantially higher than the streamflow depth. This indicates that the lack of satisfactory calibration results for 3B42 and PERSIANN over Gilgel Abay is reflective of the substantial underestimation bias in the 3B42 and PERSIANN rainfall estimates.

Let us now compare the model parameter values resulting from the different rainfall inputs. For illustration, we focus on two parameters, CN2 (the Soil Conservation Service Curve Number) and Surlag (surface runoff lag coefficient), that control the overland flow, and two satellite rainfall

inputs, CMORPH and PERSIANN. Increasing the CN2 values result in increasing runoff. The calibrated average CN2 value for Koga is 69 with CMORPH input and 73 with PERSIANN input. This sequence of increasing CN2 is in agreement with the sequence of increasing the degree of rainfall underestimation by the products (see Fig. 3). The product that has the larger negative bias in rainfall (i.e. PERSIANN) results in higher CN2 values. Surlag is a lag factor for watersheds that control surface runoff storage by lagging a portion of the runoff that would have otherwise been released to the main channel. Higher Surlag value corresponds to more runoff release to the channel. Again, as expected, we find a higher Surlag value for PERSIANN compared to CMORPH. These differences in model parameters are compensating in one way or another for the difference in the input rainfall estimates for generating runoff. In other words, readjusting the model calibration parameter values to increase the streamflow simulation accuracy compensates for the underestimation error in satellite rainfall estimates by decreasing the performance of the simulation of other water balance components. Therefore, caution must be exercised when using satellite simulations of other water balance components when the model is calibrated only on the basis of streamflow.

2.3 Approach and performance statistics

We used rainfall data from each source (3B42RT, 3B42, CMORPH, PERSIANN, and rain gauges) for the validation period 2006 to 2007 as input into SWAT with model parameter estimates corresponding to each rainfall source (e.g., CMORPH rainfall for 2006–2007 would be used as input into SWAT model calibrated using the 2003–2004 CMORPH rainfall input) and watershed to simulate daily streamflow. We assess the performance accuracy of each simulation by comparison with observed streamflow. The comparison is made based on visual inspection of hydrographs and exceedance probabilities, and through the following performance statistics: coefficient of determination (R^2), relative bias (Rbias), and Nash-Sutcliffe efficiency (NSE):

$$R^2 = \left[\frac{\sum\limits_{i=1}^{n} \left(\text{SIM}_i - \overline{\text{SIM}} \right) \left(\text{OBS}_i - \overline{\text{OBS}} \right)}{\sqrt{\sum\limits_{i=1}^{n} \left(\text{SIM}_i - \overline{\text{SIM}} \right)^2} \sqrt{\sum\limits_{i=1}^{n} \left(\text{OBS}_i - \overline{\text{OBS}} \right)^2}} \right]^2,$$

$$\text{Rbias} = \frac{\sum\limits_{i=1}^{n} \left(\text{SIM}_i - \text{OBS}_i \right)}{\sum\limits_{i=1}^{n} \text{OBS}_i},$$

$$\text{NSE} = 1 - \left[\frac{\sum\limits_{i=1}^{n} \left(\text{SIM}_i - \text{OBS}_i \right)^2}{\sum\limits_{i=1}^{n} \left(\text{OBS}_i - \overline{\text{OBS}} \right)^2} \right],$$

where SIM is the simulated daily streamflow, OBS is the observed daily streamflow, n is the total number of pairs of simulated and observed data, and the bar indicates average value over n. NSE indicates how well the plot of the observed value versus the simulated value fits the 1:1 line, and ranges from $-\infty$ to 1, with higher values indicating better agreement (Legates and McCabe, 1999). R^2 measures the variance of observed values explained by the simulated values. Rbias measures the relative error in total streamflow volume.

3 Results and discussion

We simulated daily streamflow for the validation period 2006–2007 from SWAT using rainfall input from each source and corresponding model parameters. In this section, we discuss the simulation results.

3.1 Koga watershed

Comparisons of simulated and observed streamflow for Koga watershed are given in Fig. 4. Let us first discuss the results for 2006. According to Fig. 4a, all simulations capture the overall shape of observed streamflow hydrographs, but underestimate the large flood peaks, with the rain gauge simulations showing better performance than the satellite simulations. The 3B42RT and CMORPH simulations are identical. Figure 4c shows that all simulations underestimate the frequency of the extreme events with probabilities of exceedance lower than 5%; the underestimations are severe for satellite rainfall simulations compared to the rain gauge simulations. According to Fig. 4e, the R^2 values for the time series of daily streamflow between simulated and observed values vary in the range 0.4 to 0.6; the satellite rainfall simulations underestimate the total streamflow volume by 10% to 20%, while the rain gauge simulations give almost accurate results; the NSE values, ranging from 0.4 to 0.5, indicate that all the simulations exhibit moderate skills in reproducing daily streamflow.

Do the performance accuracy results hold in 2007? According to Fig. 4b, the 3B42RT, CMORPH and PERSIANN simulations capture the monsoonal pattern but underestimate all floods. The 3B42 simulations fail to see any of the flood events, while the rain gauge simulations show superior performance, better than any of the satellite simulations. According to Fig. 4d, the 3B42RT, CMORPH and PERSIANN simulations underestimate the frequency of all extremes events with probabilities of exceedance lower than 25%, while the 3B42 simulations do not even see any of the extremes. The rain gauge simulations reproduce the frequency of extreme events very well. According to Fig. 4f, the R^2 values for the time series of daily streamflow between simulated and observed values are moderate (about 0.75) for all simulations except for the 3B42 simulation (0.05). All simulations underestimate the total streamflow volume; the degree
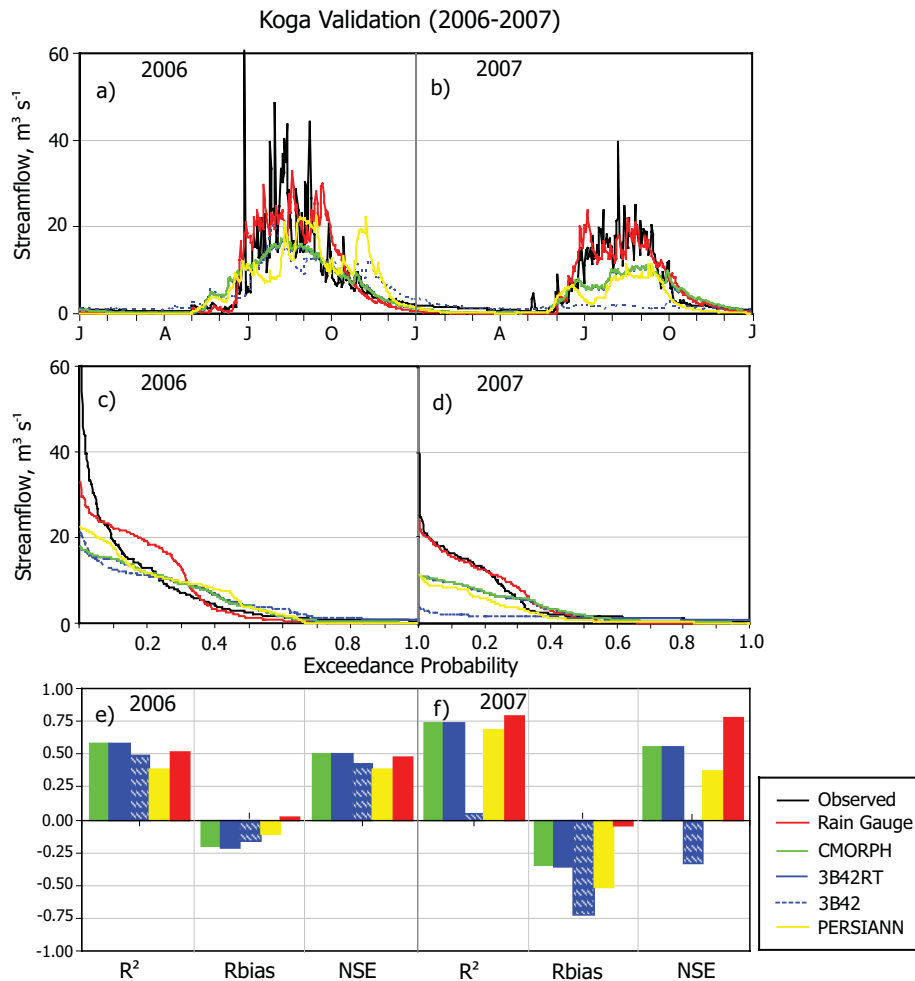
Fig. 4. Comparisons of SWAT simulated (based on CMORPH, 3B42RT, 3B42, PERSIANN, and rain gauge network rainfall inputs, separately) and observed daily streamflow for Koga watershed, in terms of **(a, b)** time series, **(c, d)** exceedance probabilities, and **(e, f)** performance statistics, for year **(a, c, e)** 2006, and **(b, d, f)** 2007.

of underestimation is very high for 3B42 (Rbias = −0.73), moderate for the other satellite rainfall products (Rbias ranging from −0.34 to −0.52) and very low (Rbias = −0.05) for the rain gauge simulations. Except for the 3B42, all simulations have positive NSE values, ranging from 0.4 to 0.8, indicating moderate skills of the simulations in reproducing the observed hydrographs. The 3B42 simulation has a negative NSE value indicating no skill in the simulations compared to simply using the mean as a predictor. Comparison of the performance statistics for 2006 and 2007 reveals that the 3B42RT, CMORPH and PERSIANN simulations show relatively stable performance over time (although with some variations in statistics) while the 3B42 simulations show large performance fluctuations from modest skill in 2006 to no skill in 2007.

3.2 Gilgel Abay watershed

Comparisons of simulated and observed streamflow for the larger watershed, Gilgel Abay, are given in Fig. 5. Let us start with the 2006 results. Figure 5a shows that the 3B42RT, CMORPH and rain gauge simulations capture remarkably the observed streamflow hydrographs, while the 3B42 and PERSIANN simulations fail to capture satisfactorily the observed hydrographs resulting in substantial underestimation. Figure 5c shows that all the satellite simulations underestimate the frequency of extreme events; the underestimation is moderate in the case of 3B42RT and CMORPH but severe in the case of 3B42 and PERSIANN simulations; the rain gauge simulation performs very well. Figure 5e shows that the R^2 values for the time series of daily streamflow between simulated and observed values are higher (0.75) for the 3B42RT, CMORPH and rain gauge simulations compared to the 3B42 (0.50) and PERSIANN (0.37) values. All
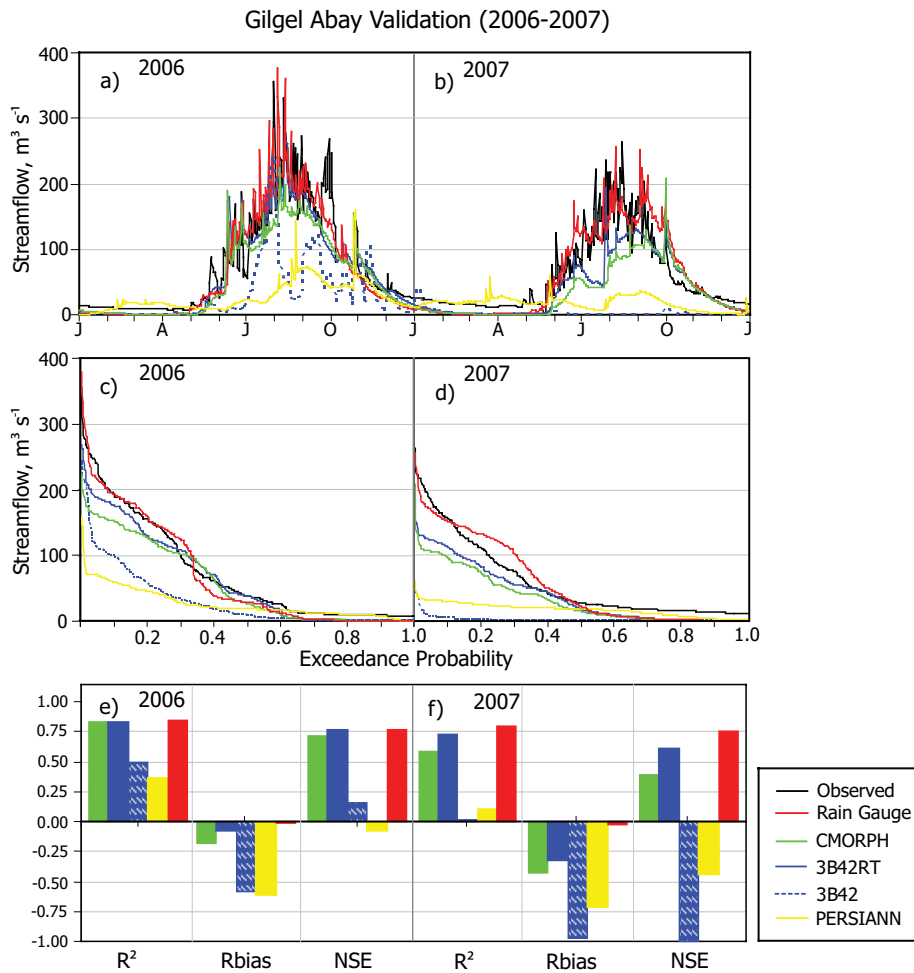
Fig. 5. Same as in Fig. 4 but for Gilgel Abay.

simulations underestimate the total streamflow volume; the underestimation is negligible for the rain gauges, moderate (Rbias = −0.08, −0.18) for the 3B42RT and CMORPH, and severe (Rbias = −0.58, −0.62) for the 3B42 and PERSIANN simulations. The NSE values are high (about 0.75) for the 3B42RT, CMORPH and rain gauge simulations, low (0.16) for the 3B42, and negative for PERSIANN. The performance accuracy of all satellite simulation is lower in 2007 than it is in 2006; however, the 3B42RT, CMORPH and rain gauge simulations still have modest skills in reproducing daily streamflow, while both 3B42 and PERSIANN show no skills.

3.3 Koga vs. Gilgel Abay

Koga and Gilgel Abay are adjoining watersheds with significant differences in the watershed area, Koga at 299 km^2 and Gilgel Abay at 1656 km^2. Figure 6 presents comparison of the performance statistics (Rbias and NSE) between the two watersheds, for each rainfall input simulation. Increasing watershed area increases slightly the performance

accuracy of the 3B42RT, CMORPH and rain gauge simulations, but decreases substantially the performance accuracy of the 3B42 and PERSIANN simulations. The increased performance accuracy of the 3B42RT, CMORPH and rain gauge simulations for larger watersheds is as expected due to the additional averaging process in larger watersheds that tends to dampen the random error in rainfall input and hydrological process approximation. The decreasing performance accuracy of the 3B42 and PERSIANN in larger watersheds is counter-intuitive and indicates that larger watersheds introduce much more errors from the unreliable rainfall estimates of 3B42 and PERSIANN than the reduction in random error gained due to more averaging. We acknowledge that the differences between the two watersheds may not be exclusively due to the watershed size, as Uhlenbrook et al. (2010) reported significant differences in the hydrological characteristics between the two watersheds.
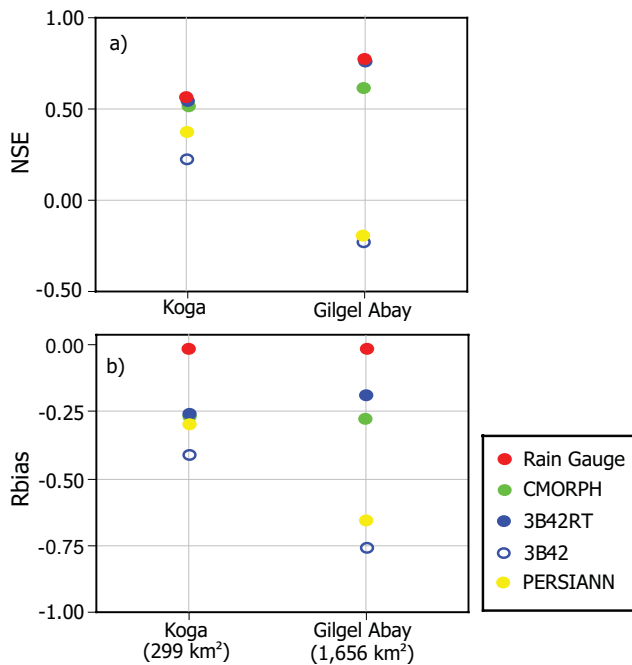
Fig. 6. Comparisons of the performance statistics (NSE and Rbias) of SWAT simulations (based on CMORPH, 3B42RT, 3B42, PERSIANN, and rain gauge network rainfall inputs, separately) as a function of watershed size, for **(a)** NSE and **(b)** Rbias.

4   Conclusions

The main purpose of this study is to assess the utility of satellite rainfall estimates as input into a hydrological model for daily streamflow simulation in the East African highlands. We limited our analyses to the following specifics: the semi-distributed hydrologic model SWAT; adjoining two watersheds, Koga at 299 km^2 and Gilgel Abay at 1656 km^2; and four types of satellite precipitation products (3B42RT, 3B42, CMORPH, and PERSIANN). Our results reveal that the utility of satellite rainfall products as input to SWAT for daily streamflow simulation strongly depends on the product type. The 3B42RT and CMORPH simulations show consistent and modest skills in their simulations but underestimate the large flood peaks. On the other hand, the 3B42 and PERSIANN simulations have inconsistent performance with poor or no skills. Let us put these results in perspective.

4.1   Microwave vs. infrared algorithm products

Depending on the main input, satellite rainfall algorithms can be grouped into two categories: those that use primarily microwave data (e.g., CMORPH, 3B42RT) and those that use primarily infrared data (e.g., PERSIANN). The conventional notion is that the microwave-based algorithms fare better than the infrared-based algorithms. Our results indicate that not only are the microwave-based algorithms better than the infrared-based algorithm, but the infrared-based

algorithm also has poor or no skills for streamflow simulations. We conclude that the infrared-based algorithm PERSIANN is not a reliable source of rainfall data in the East African highlands.

4.2   Satellite-gauge vs. satellite-only products

The conventional notion that the satellite rainfall estimates that incorporate rain gauge information perform better than the satellite-only estimates has led to the incorporation of rain gauge data into global satellite rainfall products. Our results turn this conventional notion on its head. The satellite-only product (3B42RT) performs much better than the satellite-gauge product (3B42). Apparently incorporating rain gauge data in satellite rainfall products has the undesirable consequence of deteriorating the quality of the satellite rainfall products in this region. This suggests that the algorithm used to incorporate rain gauge information in the satellite rainfall algorithms needs to be modified to account for the effects of mountainous topography and sparse rain gauge network.

4.3   Effect of watershed area

One would expect the performance accuracy of the satellite streamflow simulations to increase as the watershed area becomes larger. Our results indicate that this actually depends on the satellite rainfall product used as input. For satellite rainfall products that have relatively reliable and consistent performance (3B42RT and CMORPH), the resulting streamflow simulations will indeed have higher performance for larger watersheds. However, for satellite rainfall products that have unreliable and inconsistent performance (3B42 and PERSIANN), the resulting streamflow simulations' performance accuracy decreases as the watershed area increases from 299 km^2 to 1656 km^2, indicating that larger watersheds introduce more errors from the unreliable rainfall estimates of 3B42 and PERSIANN than the reduction in random error gained due to more averaging.

We have repeated the above analyses using different hydrologic models (HBV and MIKE SHE), and the results indicate that our conclusions are fairly insensitive to the choice of the model type (results not shown). Finally, we acknowledge the limitations of this study. As suggested by Uhlenbrook et al. (2010), there could be a large uncertainty in the observed streamflow data. The impact of watershed area was also investigated using only two watersheds. We recommend further investigation based on the analysis of a reasonably large number of watersheds with high-quality streamflow data.

Edited by: S. Uhlenbrook

References

Arnold, J. G., Srinivasan, R., Muttiah, R. S., and Allen, P. M.: Large-Area Hydrologic Modeling and Asssessment: Part I. Model Development, J. Am. Water Resour. Assoc., 34(1), 73–89, doi:10.1111/j.1752-1688.1998.tb05961.x, 1998.

Bitew, M. M. and Gebremichael, M.: Evaluation through independent Measurements: Complex Terrain and Humid Tropical Region in Ethiopia, in: Satellite Rainfall Applications for Surface Hydrology, edited by: Gebremichael, M. and Hossain, F., ISBN: 978-90-481-2914-0, Springer-Verlag, New York, 2010.

Dinku, T., Chidzambwa, S., Ceccato, P., Connor, S. J., and Ropelewski, F.: Validation of high-resolution satellite rainfall products over complex terrain, Int. J. Remote Sens., 29(14), 4097–4110, 2008.

Dinku, T., Ceccato, P., Grover-Kopec, E., Lemma, M., Connor, S. J., and Ropelewski, C. F.: Validation of satellite rainfall products over East Africa's complex topography, Int. J. Remote Sens., 28(7), 1503–1526, 2007.

Easton, Z. M., Fuka, D. R., White, E. D., Collick, A. S., Biruk Ashagre, B., McCartney, M., Awulachew, S. B., Ahmed, A. A., and Steenhuis, T. S.: A multi basin SWAT model analysis of runoff and sedimentation in the Blue Nile, Ethiopia, Hydrol. Earth Syst. Sci., 14, 1827–1841, doi:10.5194/hess-14-1827-2010, 2010.

Eckhardt, K. and Arnold, J. G.: Automatic calibration of a distributed catchment model, J. Hydrol., 251(1–2), 103–109, doi:10.1016/S0022-1694(01)00429-2, 2001.

Gassman, P. W., Reyes, M. R., Green, C. H., and Arnold, J. G.: The soil and water assessment tool: Historical development, applications, and future research directions, Trans. ASABE, 50(4), 1211–1250, 2007.

Hirpa, F. A., Gebremichael, M., and Hopson, T.: Evaluation of High Resolution Satellite Precipitation Products over Very Complex Terrain in Ethiopia, J. Appl. Meteorol. Clim., 49(5), 1044–1051, doi:10.1175/2009JAMC2298.1, 2010.

Huffman, G. J., Adler, R. F., Bolvin, D. T., Gu, G., Nelkin, E. J., Bowman, K. P., Hong, Y., Stocker, E. F., and Wolff, D. B.: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multilayer, combined-sensor, precipitation estimates at fine scale, J. Hydrometeorol., 8, 38–55, 2007.

Joyce, R. J, Janowiak, J. E., Arkin, P. A., and Xie, P.: CMORPH: A method that produces global precipitation estimation from passive microwave and infrared data at high spatial and temporal resolution, J. Hydrometeorol., 5, 487–503, 2004.

Kidd, C. and Muller, C.: The combined passive microwave-infrared (PMIR) algorithm, in: Satellite Rainfall Applications for Surface Hydrology, edited by: Gebremichael, M. and Hossain, F., Springer-Verlag, New York, 2009.

Legates, D. R. and McCabe, G. J.: Evaluating the use of "goodness of fit" measures in hydrologic and hydroclimatic model validation, Water Resour. Res., 35(1), 233–241, 1999.

Scofield, R. A. and Kuligowski, R. J.: Status and outlook of operational satellite precipitation algorithms for extreme-precipitation events, Weather Forecast., 18, 1037–1051, 2003.

SCS – Soil Conservation Service: TR-55, Urban Hydrology for Small Watersheds, Natural Resources Conservation Service, Washington, DC, 1986

Sorooshian, S., Hsu, K., Gao, X., Gupta, H. V., Imam, B., and Braithwaite, D.: Evaluation of PERSIANN system satellite-based estimates of tropical rainfall, B. Am. Meteorol. Soc., 81, 2035–2046, 2000.

Turk, F. J. and Miller, S. D.: Toward improved characterization of remotely sensed precipitation regimes with MODIS/AMSR-E blended data techniques, IEEE T. Geosci. Remote, 43, 1059–1069, 2005.

Uhlenbrook, S., Mohamed, Y., and Gragne, A. S.: Analyzing catchment behavior through catchment modeling in the Gilgel Abay, Upper Blue Nile River Basin, Ethiopia, Hydrol. Earth Syst. Sci., 14, 2153–2165, doi:10.5194/hess-14-2153-2010, 2010.

Ushio, T. and Kachi, M.: Kalman filtering applications for global satellite mapping of precipitation (GSMaP), in: Satellite Rainfall Applications for Surface Hydrology, edited by: Gebremichael, M. and Hossain, F., Springer-Verlag, New York, 2009.

van Griensven, A., Meixner, T., Grunwald, S., Bishop, T., DiLuzio, M., and Srinisvan, R.: A global sensitivity analysis tool for the parameters of multi-variable catchment models, J. Hydrol., 324, 10–23, doi:10.1016/j.jhydrol.2005.09.008, 2006.