Validation of TRMM Rainfall for Pangani River Basin in Tanzania[1]

R. M. F. Haque[a, *], S. Maskey[b], S. Uhlenbrook[b, c] and M. Mul[b]
[a]Bangladesh Water Development Board, Dhaka, Bangladesh
[b]UNESCO-IHE Institute for Water Education, P.O. Box 3015, 2601 DA, Delft, The Netherlands
[c]Delft University of Technology, Department of Water Resources, P.O. Box 5048, 2600 DA, Delft, The Netherlands

Abstract: Tropical Rainfall Measuring Mission (TRMM) rainfall products are one of the increasingly popular remote sensing based rainfall estimates, which are available at 0.25° x 0.25° spatial and up to 3 hourly temporal resolutions since January 1998 and covers 50° S to 50° N around the globe. This study assesses the reliability of the TRMM rainfall for Pangani River Basin in Tanzania by comparing it with the gauge rainfall data. Monthly rainfall data from 1998 to 2005 for 50 gauge stations and daily rainfall data from 1998 to 2003 from 12 stations are used for the comparison. The comparison of TRMM grid rainfall with point gauge rainfall is achieved by using nearest neighborhood and bilinear weighted interpolation methods. Application of bilinear weighted interpolation method has brought some improvement in the comparison of results over the nearest neighborhood method. In both methods, the TRMM rainfall estimates compares better with the gauge rainfall at a monthly time scale than at daily time scale. The comparison is also carried out at a sub-basin level by dividing the basin into 7 sub-basins. The sub-basin rainfall is estimated from the gauge rainfall by the interpolation method based on inverse distance and elevation weighting. The spatial rainfall comparison at sub-basin level has shown a reasonably good result. Comparing the consecutive yearly total rainfall maps, it is seen that TRMM has reasonably captured both spatial and temporal rainfall pattern but failed to detect higher intensity rainfall close to the mountainous parts of the basin. Five statistical performance measures namely the coefficient of determination, mean error (bias), root mean square error, relative root mean square error and Nash-Sutcliffe coefficient are used to quantify the comparison results. In conclusion, TRMM has performed fairly well on a monthly scale but its hydrological application in mountainous areas and for high intensity rainfall period still requires validation with the ground based gauge rainfall before taking it as a reliable alternative.

Keywords: Rainfall estimation, TRMM, nearest neighborhood, bilinear weighted interpolation, sub-basin

1 Introduction

Rainfall data availability has been highlighted as a major constraint on the effective application of water resource models, and it has been argued that quality of rainfall inputs to the model is often more important than choice of model itself as discussed by Wilk et al. (2006). The quality, availability and coverage of rain gauge data are particular obstacles to effective water resource planning in Africa as well as most developing countries as per Thorne et al. (2001), Grimes and Diop (2003) and Hughes (2006). Although radar estimates of rainfall are increasingly available in technologically developed countries like the USA and Europe, in large parts of Africa these are almost non-existent. However, several satellite products provide complete coverage of the African continent at applicable time and space scales, and this data source seems to be an obvious way forward for regional-to national-scale rainfall estimation.

Recently, rainfall products from Tropical Rainfall Measuring Mission (TRMM) are becoming a popular alternative rainfall data source, particularly in large scale applications. Satellite observation based rainfall estimates are not a direct measurement of rainfall but are only an estimate and are subject to uncertainty. It is important that these data are well validated before using for practical applications.

Bowman (2005) compared TRMM rainfall with rainfall data from 26 rain gauges from ocean buoys in tropical pacific and showed that comparison results can be improved by properly averaging in space and/or time. For individual satellite overpasses averaged over a 1° x 1° box, his comparison showed Relative Root Mean Square Error (RRMSE) of as high as 200% to 300%. Whereas for 32-day means over 1° x 1° boxes, the RRMSE reduced to 40% to 70%. Collischonn et al. (2008) compared the aggregated 3 hourly TRMM rainfall estimates to daily values with

[1] *Paper JHER002 submitted 09/08/2013; accepted for publication after peer review and subsequent revision on 18/10/2013*
[*] *Corresponding author can be contacted at rmfhaque@gmail.com*

the ground-level precipitation gauge data on a daily basis. Both daily TRMM and rain gauge-interpolated rain fields were then used as input to a large-scale hydrological model for the Tapajo´s river basin, a major tributary of the Amazon; the calculated hydrographs were then compared to observations at several stream gauges along the River Tapajos and its main tributaries. Results of the rain field comparisons showed that satellite estimates can be a practical tool for identifying damaged or aberrant rain gauges at basin scale.

Fengge et al. (2008) took the La Plata basin in South America as a good example of a case where the use of satellite-derived precipitation could be beneficial. He evaluated basin wide precipitation estimates from 9 yr (1998–2006) of TTRMM Multisatellite Precipitation Analysis (TMPA; 3B42 V.6) through comparison with available gauged data and the Variable Infiltration Capacity (VIC) semi-distributed hydrology model applied to the La Plata basin. He found in general, the TMPA estimates agreed well with the gridded gauge data at monthly time scales, most likely because of the monthly adjustment to gauges performed in TMPA. The agreement between TMPA and gauge precipitation estimates was reduced at daily time scales, particularly for high rain rates. The TMPA-driven hydrologic model simulations were able to capture the daily flooding events and to represent low flows, although peak flows tended to be biased upward. Hazarika et al (2005) also found TRMM underestimated the rain for peak monsoon period for the wet rain regimes whereas it overestimated the rain for dry rain regimes of Nepal and suggested the possibility of using satellite data for operational flood forecasting in those river basins for which real time gauge data were not available. On the other hand, due to the overestimation and underestimation of daily rainfall measured from TRMM V5 3B42 in comparison with the ground rain gauge data Islam and Uyeda (2007) recommended special attention for the application of V5 3B42 data to short term hydrological problems such as flash flood. McIntyer et al. (2008) suggested that the relative performance of the satellite-based rainfall estimation algorithms depend on what aspects of the rainfall regime are being considered. Differences between the products are large and the use of more than one product for any application is recommended.

Uddin et al. (2008) found good correlation between the computed TRMM and measured data by the application of bilinear weighted interpolation method. He recommended the usefulness of bilinear weighted interpolation method to derive precipitation information at any spatial location when direct measurements are not available. TRMM-derived precipitation showed better detection of rain at low altitude stations as compared with high elevation stations, with good scores for the (Precipitation Radar) PR product for rain rates >0.5 mm/hr as discussed in Barros et al. (2000).

This study assesses the reliability of the TRMM rainfall for the Pangani River Basin in Tanzania by comparing it with the gauge rainfall data. The comparison is carried out at two stages. Firstly, the gauge rainfall is compared with the TRMM rainfall estimated at the gauge location. For this point comparison, the TRMM rainfall of 0.25×0.25 grid is interpolated for the location of the gauge by nearest neighborhood and bilinear interpolation methods. Secondly, the comparison is carried out at sub-basin level by dividing the basin into 7 sub-basins. The sub-basin rainfall is estimated from the gauge rainfall by inverse distance and elevation weighting method.

2 Pangani River basin gauge rainfall data

The study area selected for this paper is the Pangani river basin, one of the nine river basins in Tanzania as shown in Figure 1. The total drainage area of the basin is about 43,650 km^2 with 95% of the basin area in Tanzania and 5% in Kenya. It is located in the north eastern part of Tanzania between Latitudes 2^{o}55'S to 5^{o}40'S and Longitude 36^{o}20'E to 39^{o}2'E. The source of the Pangani River originates at Mt Kilimanjaro (the highest mountain in Africa) and Mt Meru. The Pare and Usambara Mountains, bordering the west side of the catchment, also contributes substantial amount to the Pangani river flow. The basin is subject to high variability, in terms of altitude and local climate. Rainfall is subject to high inter- and intra-seasonal variability. There are five main sub-basins: Ruvu, Kikuletwa, Mkomazi, Luengere and Pangani. The Kikuletwa and Ruvu tributaries form the headwaters of the Pangani River and originate in the high rainfall areas surrounding Mt Meru and Mt Kilimanjaro, respectively. Elevation in the basin varies from sea level at Pangani (the basin outlet) to 5895 m on top of Mount Kilimanjaro. There is high spatial variability of rainfall in the basin, mainly characterised by the topography.

The high altitude slopes above the forest line on Mt Meru and Mt Kilimanjaro have an Afro-Alpine climate and receive in excess of 2500 mm of rainfall per year. The middle slopes of these mountains, and the Pare and Usambara Mountains are characterized by a humid to sub-humid tropical climate. The lower Mkomazi catchment and the lower Pangani catchment have a sub-humid to semi-arid climate. The central and western parts of the Basin have a semi-arid to arid climate. Rainfall varies between 300 and 600 mm/year.

Majority of the rain gauges of Pangani river basin are situated on an axis along the mountain ranges of Mount Kilimanjaro, North-South Pare Mountains and the Usambara Mountains. There is actually no rainfall station in western part of the basin. The data available was either of short record or contained a lot of gaps.

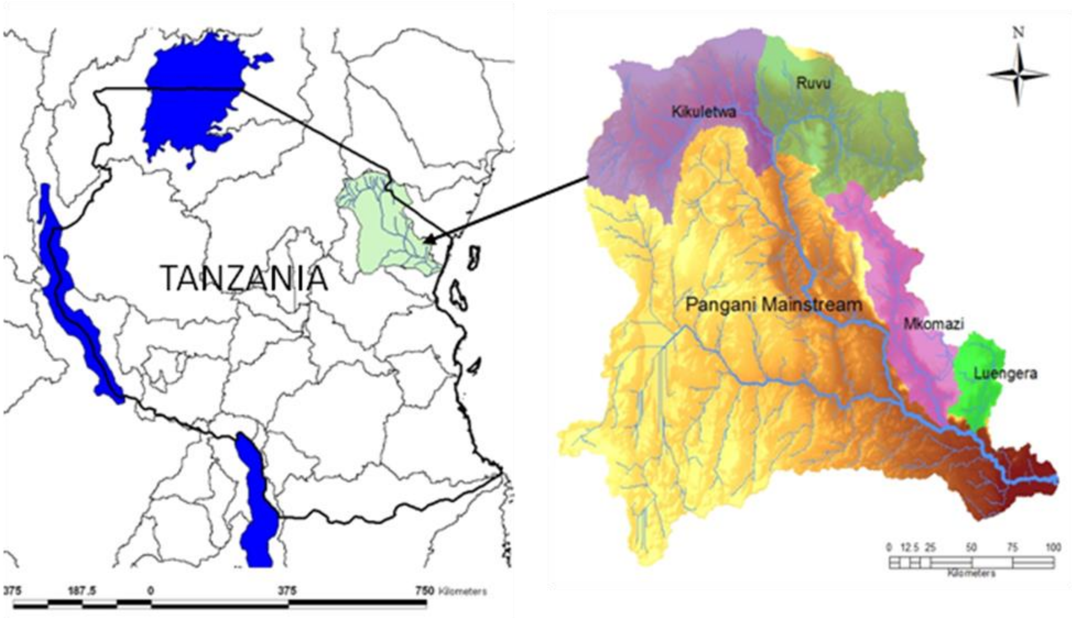

Figure 1 Location of Pangani river basin in Tanzania

3 TRMM satellite mission and rainfall products

The accurate measurement of the spatial and temporal variation of tropical rainfall around the globe remains one of the critical unsolved problems of meteorology. TRMM, during its mission and broad sampling footprint between 35°N and 35°S, is providing some of the first detailed and comprehensive dataset on the four dimensional distribution of rainfall and latent heating over the vastly under sampled oceanic and tropical continental regimes. The TRMM satellite, a joint project between the United States (under the leadership of NASA's Goddard Space Flight Centre and Japan (under the leadership of the National Space Development Agency), and the first spacecraft designed to monitor rain over the tropics, was successfully launched on November 27, 1997. It was placed in low earth orbit as the first precipitation radar (PR) in space, along with a 9-channel Special Sensor Microwave/Imager (SSM/I)-like passive microwave imager (TMI), an Advanced Very High Resolution Radiometer (AVHRR)-like visible-infrared radiometer (VIRS), a lightning sensor and a cloud sensor as discussed in Zhong et al. (2002).

3.1 TRMM product review

There are two main rainfall products based called 3B42 and 3B43. The 3B42 estimates are produced in four stages: (i) the microwave estimates precipitation are calibrated and combined; (ii) infrared precipitation estimates are created using the calibrated microwave precipitation; (iii) the microwave and (infrared) IR estimates are combined; and (iv) rescaling to monthly data is applied. Each precipitation field is best interpreted as the precipitation rate effective at the nominal observation time as per JAXA (2006).

The 3B42 is to provide precipitation estimates in the TRMM regions that has the (nearly-zero) bias of the "TRMM Combined Instrument" precipitation estimate and the dense sampling of geosynchronous IR imagery. The 3B42 is composed of two separate algorithms, which are (i) to produce monthly IR calibration parameters, and (ii) to calibrate the merged-IR precipitation data to produce the daily adjusted merged-IR precipitation and (Root Mean Square) RMS precipitation error estimates as per JAXA (2006).

For the 3B43 monthly data, in the estimates it includes the dependent data. It processes the TRMM Microwave Imager (TMI) 1B01, 2A12 and Combiner (COMB) 3B31 and Merged IR data (3A44). Additionally, Geosynchronous Precipitation Index (GPI) is used to convert Visible and Infrared Scanner (VIRS) radiance to precipitation rate as discussed in JAXA (2006). Algorithms 3B43 is executed once per calendar month to produce the single, best estimate precipitation rate and RMS precipitation error estimate field 3B43 by combining the 3 hourly merged HQ/IR estimates 3B42 with the monthly accumulation Climate Assessment and Monitoring System (CAMS) or Global Precipitation Climatology Centre (GPCC) rain gauge analysis as per JAXA (2006).

The purpose of algorithm 3B42 and 3B43 is to produce TRMM rainfall retrievals merged high quality (HQ)/ infrared (IR) precipitation and root mean square (RMS) precipitation error estimate. The data include retrievals from six different algorithms: VIRS, TMI, (Precipitation Radar) PR and the combination with other satellite (GPI, GPCP and SSM/I).

4 Methodology

4.1 Data used for comparison

4.1.1 Gauge rainfall data

The precipitation data for the Pangani River basin is obtained via the "Pangani Basin Water Office (PBWO) in Moshi". Due to the long time series, data from ground based rain gauges is used for this research. The data that has been used could be divided into two groups: (i) 50 rain gauges with monthly data from 1970 to 2005; and (ii) 32 rain gauges with daily data from 1916 to 2003.

All rain gauges are standard manual rain gauges and part of Tanzania rain gauge network maintained by the Tanzanian Meteorological Agency in conformance with the World Meteorological Organization (WMO) standard. The rain gauges in the Pangani basin are primarily located near permanent settlements.

4.1.2 TRMM rainfall data

For this study, eight years (1998 to 2005) observations from TRMM have been used. The data set used for this study is 3-hourly TRMM and other rainfall estimate (3B42 V6), daily TRMM and other rainfall estimate (3B42 V6 derived) and monthly TRMM and other data sources rainfall estimate (3B43 V6). 3B43 V6 consists of accumulated rainfall in millimeters, based on multi-satellite precipitation analysis. The data has an area averaged over 0.25^0 x 0.25^0 longitude and latitude grid boxes (approximately 25km x 25km). The Lat-Lon plot type provides a time-averaged data plot for a specified area. This data could be downloaded as an ASCII output from the principal web page of TOVAS.

The TRMM 3B42 records precipitation every 3 hours and the TRMM imager is saved from 12 UTC Coordinated Universal Time (UTC) of the previous day to 12 UTC of the analysis day. Normally, the daily data rain gauges record 24-hr precipitation accumulation ending at the time of observation and this is reported as the rainfall for the day of observation as suggested by Dingman (2002). To create similar dataset for comparison, the 3-hourly TRMM precipitation data are accumulated to 24-hr depths depending on the rainfall observation time of the rain gauge located in the basin. The rainfall observation time in Pangani River basin, Tanzania is 9:00 AM local time. The 3-hourly TRMM precipitation data are accumulated from 9:00 AM on January 1, 1998 to 9:00 AM January 2, 1998 and this 24-hr depth is considered as the rainfall for January 2, 1998. Those days with missing and suspicious TRMM data for any hour are excluded from the analysis. Daily precipitation data from 12 daily rain gauges are assembled in a similar format, excluding the dates with missing data. Then the daily rain gauge data files and TRMM data files are compared to create a new set of data files that included only those dates without any missing data for both rain gauge and satellite. In this study, data from rain gauges stations over Pangani River basin, Tanzania is used to compare with TRMM rainfall data using several statistical measures.

4.2 Indicators used for comparison

The comparison results are evaluated using five indicators, namely the coefficient of determination (R^2), Nash-Sutcliff coefficient of efficiency (C_oE), root mean square error (RMSE), relative root mean square error (RRMSE) and mean error bias. These indicators are defined below.

R^2 is one of the most common and most powerful statistics. A correlation is a single number that describes the degree of relationship between two variables. It is good measure of linear association or phase error. Visually, how close the points of a scatter plot are to a straight line. It is possible for a set of estimated values with large errors to still have good coefficient of determination with the observations. It is sensitive to outliers and goes from 0 to 1 as discussed by Murphy (1995). For this study, the coefficient of determination addresses the question how well did the TRMM retrievals correspond to the observed rainfall ground values.

$$\text{Mean error (bias)} = \frac{1}{n}\sum_{i=1}^{n}(Gi - Fi) \qquad (1)$$

RMSE = $\sqrt{\frac{1}{n}\sum_{r=1}^{n}(Gi-Fi)^2}$ (2)

RRMSE = $\frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n}(Gi-Fi)^2}}{\bar{\bar{G}}}$ (3)

C_oE = 1.0 - $\frac{\sum_{i=1}^{n}(G_i-Fi)^2}{\sum_{i=1}^{n}(G_i-\bar{\bar{G}})^2}$ (4)

G_i = Rain gauge measurements

F_i = Satellite estimates

$\bar{\bar{G}}$ = mean of the rain gauge measurement

n = is the number of observations.

4.3 Point to grid comparison

The rain gauges and the satellite instruments make fundamentally different measurements. The gauges provide in situ high time-resolution measurements at a point. The TRMM instruments make remote, volume-averaged measurements of hydrometeors in the atmosphere, from which the area averaged surface rain rate over the instantaneous fields-of-view is inferred. By scanning across the orbit track, TRMM can provide a snapshot of the rain rate over an extended region. Thus, both instruments provide only limited samples of the precipitation falling within a region. The rain gauges have good time sampling but poor spatial sampling; while the satellite has good spatial sampling but poor time sampling as observed by Bowman (2005).

It is frequently the case that when TRMM observes rain within a region it is not raining at the gauge. Similarly, it often rains at the gauge between TRMM overpasses. Using only gauge data from the time of the TRMM overpass neglects a great deal of information collected by the gauge between satellite overpasses as discussed by Bowman (2005). Only the individual pixel that contains the gauge neglects information collected by the satellite at nearby locations. Because rainfall is correlated with itself in both space and time, more information can be obtained by properly averaging the data in space and/or time.

4.4 Nearest neighborhood method

Here the pixel values for the grid within which rain gauges are cited are used to make comparison. Monthly TRMM precipitation data is downloaded by using the gauge station geographic coordinates (Lat/Lon) for a specific period. When the station's coordinate (Lat/Lon) is specified in TOVAS, for a single point, it gives the time series for the nearest grid point and no averaging or interpolating is performed.

4.5 Bilinear weighted interpolation method

As TRMM 3B43 V6 data with a 0.25^0 grid is a bit coarse for local interpretation, instead of using the grid value for a particular location, or the averaging of adjacent pixel values on 0.25^0 grid, the bilinear weighted interpolation is presented here which is more realistic. The data from TRMM 3B43 dataset is used for bilinear weighted interpolation to resample the data for a specific point of interest as per Brito et al. (2003), which in the present study is the spatial location of the rain gauge. The logic that has been adopted is to average four adjacent pixels for a location of interest as dicussed by Kuhel and Sacchi (2003). The location (X, Y) has been kept in a 2 x 2 grid in such a manner that it occupies a central position in the grid. An illustration is included in Figure 2. In the bilinear interpolation, a simplistic rationale is followed, i.e. for calculating the pixel value of a particular position (X, Y), four adjacent pixel values are used. The closer the pixel is to the position (X, Y), the more influence (weight) it will carry. The method is not merely a falling function of distance from the pixel. Rather it considers a weighted approach based on its spatial locations in a two-dimensional space as discussed by Gribbon and Bailey (2004). The derivation of bilinear interpolation weights can be expressed as follows as per Arnold et al. (2002):

$$X = S_x X_1 + (1-S_x)X_2 \quad (5)$$

$$S_x = \frac{X - X_1}{X_2 - X_1} \quad (6)$$

$$Y = S_y Y_1 + (1-S_y)Y_2 \quad (7)$$

$$S_y = \frac{Y - Y_1}{Y_2 - Y_1} \tag{8}$$

where $0 \le S \le 1, X_1 \le X \le X_2$ where the pairs $(X_i Y_i)$ are the (Long, Lat) cordinates of the 4 grid cells as shown in Figure 2.

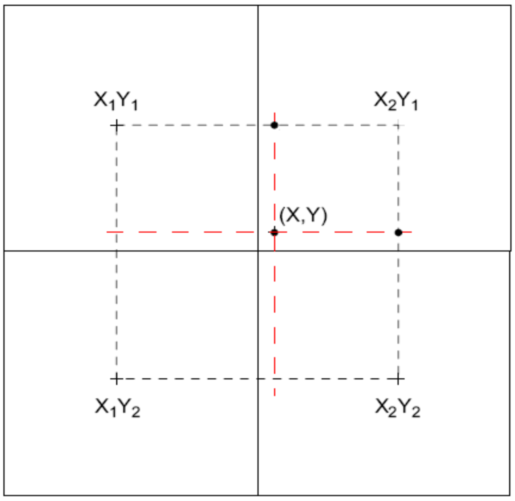

Figure 2 A graphical representation of weight estimation on 2 x 2 grid

Considering the above, the actual weight at any point (*X, Y*) in a two-dimensional space can be computed by equation (9) as suggested by Arnold et al (2002).

$$I(X,Y) = (1-S_x)(1-S_y)I(X_1,Y_1) + S_x(1-S_y)I(X_2Y_1) + (1-S_x)S_yI(X_1Y_2) + S_xS_yI(X_2Y_2) \tag{9}$$

where *I* is the actual pixel value.

In the current study, it was found that some improvement in the estimation of R^2 by the use of bilinear weighted interpolation method over the nearest neighborhood method for 50 monthly rainfall stations data comparison. R^2 is increased by 5% to 80% for the 50 monthly rainfall stations in Pangani River basin by the use of bilinear weighted interpolation method and it ranges from 0.18 to 0.72. Especially those stations having poor R^2 values got better R^2 by the use of bilinear weighted interpolation method. For example, comparison statistics between TRMM and gauge monthly data for 8 stations of out 50 stations are provided in Table 1.

By the use of bilinear weighted interpolation method, the overestimation (negative mean error) by TRMM monthly data has been significantly reduced for Mweka College of wildlife and Uru Estate station which are located in Kikuletwa and Ruvu catchment of Pangani River basin. On the other hand, Osaki forest station and Lyamungo have got underestimation (positive mean error) which are located in Kikuletwa catchment. By the same method no significant improvement in mean error has been observed for TPC Langasani, Osaki forest, Mwakinyumbi Sisal Estate and Hale plantations which are located in Kikuletwa and Pangani main stream respectively. It was found some improvement in the R^2 for all the 8 stations mentioned in Table 1.

The standard deviation of bias for the 50 monthly rainfall stations data of Pangani basin is 23.70 mm/month for bilinear weighted interpolation method whereas it is 31.31 mm/month for nearest neighbourhood method. This is an acceptable improvement over the previous method. It deserves a special mention here that the bilinear weighted interpolation method has not brought any improvement in daily comparison over nearest neighborhood method.

Table 1 Comparison statistics between TRMM and gauge monthly data for 8 stations by the use of nearest neighborhood and bilinear weighted interpolation method

Station Name	Altitude	Nearest neighborhood method			Bilinear weighted interpolation method		
		Mean error	C_oE	R^2	Mean error	C_oE	R^2
Shume forest station ID 9438012	1891	-43.14	-8.99	0.57	-46.47	-9.97	0.57
Mweka College of wildlife ID 9337098	1452	-65.14	-0.22	0.2	-10.2	0.27	0.28
Osaki forest station ID 9337121	1432	-9.68	0.05	0.1	47.62	0.09	0.18
Uru Estate ID 9337140	1415	-62.94	-0.37	0.26	-8.87	0.39	0.41
Lyamungo ID 9337021	1334	-43.48	0.04	0.16	25.28	0.23	0.27
TPC Langasani ID 9337028	715	-27.26	0.06	0.48	-23.27	0.22	0.52
Mwakinyumbi Sisal Estate ID 9538019	248	10.03	0.64	0.65	10.38	0.64	0.66
Hale plantations ID 9538010	202	-17.89	0.46	0.57	-17.37	0.49	0.59

The average gauge monthly rainfall of the station Shume forest is 21.32 mm/month and 64.46 mm/month from TRMM. Figure 3 plots the variation between the two estimates which is prominent for the whole time period (1998-2005). This station has an altitude of 1891m and located on the higher part of Usambara mountain. High rainfall occurs around Usambara due to orographic effect. The precipitation in the upper part could be different from lower part. Besides, Environmental factors such as wind speed or the error in the gauge itself could result a negative bias of -43.14mm/month. The variation and poor CoE can be attributed to the weaker detection of TRMM at higher altitude station. The station Osaki forest (ID 9337121) has got the lowest R^2 (0.18). The variation between the two estimates is the reason for this poor correlation as shown in Figure 4.

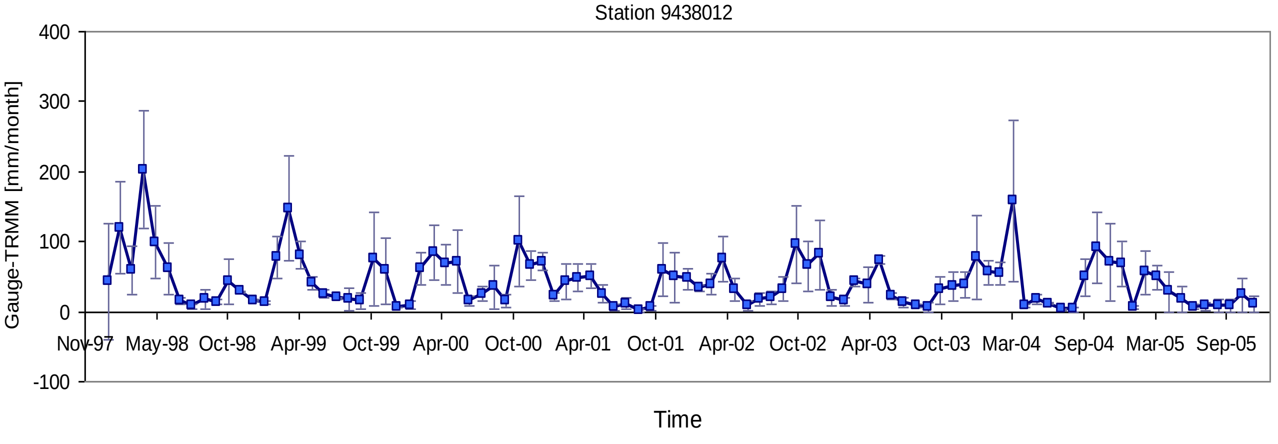

Figure 3 Variation between monthly gauge and TRMM rainfall for the station Shume forest station

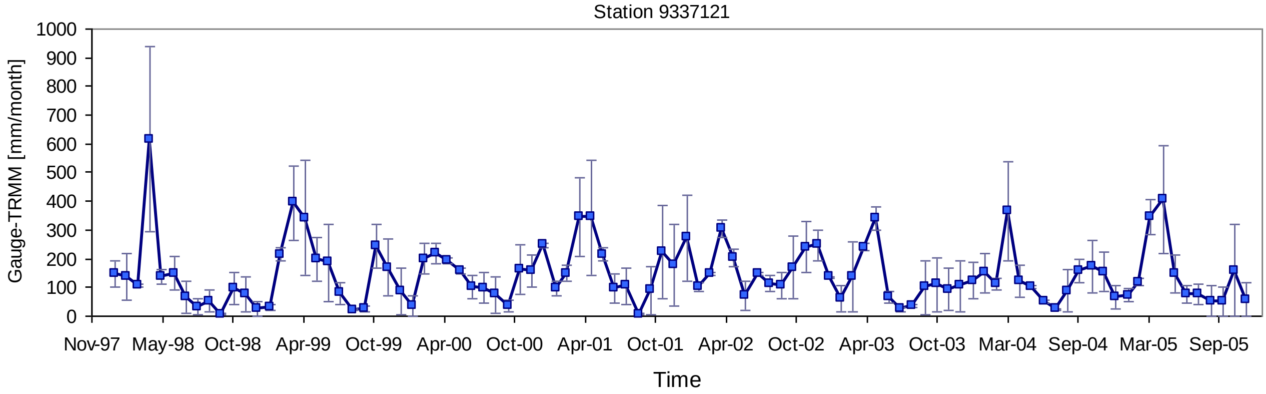

Figure 4 Variation between monthly gauge and TRMM rainfall for the station Osaki forest station

This station is located at the slope of Mount Kilimanjaro where high rainfall occurs due to orographic effect and TRMM has limitation in detecting orographic enhancement of rainfall as such clouds are warm-topped. Due to the unique location of this gauge more careful gauge measurement is required for future comparison with TRMM data.

4.6 Sub-basin level comparison

TRMM provides rainfall on 0.25^0 x 0.25^0 grids which is bit course to make comparison with point gauge rainfall. Hence the DEM and sub-basins of Pangani are resampled into 0.05^0 x 0.05^0 grids. Here the comparison has been done between the 0.05^0 x 0.05^0 interpolated point gauge rainfall vs. 0.25^0 x 0.25^0 TRMM monthly estimates.

Algorithm used for deriving spatial rainfall from gauge rainfall

This interpolation method calculates points without data from the sum of weighted observation in the neighborhood. The weights are proportional to the distance and elevation by equation (11) as suggested by Maskey and Venneker (2006), the influence of a rain gauge reduces with increasing distance and elevation to it. A disadvantage of the method is the arbitrary choice for the exponent and there is no estimation for error.

$$p_j = \sum_{i=1}^{N} w_{i,j} p_i \quad (10)$$

$$w_{i,j} = f_{i,j}^n / \sum_{i=1}^{N} f_{i,j}^n \quad (11)$$

p_j = station weighted rainfall for the facet j [L/T]

N = No of rainfall stations relevant to the grid cell. [-]

p_i = Observed rainfall at station i [L/T]

$w_{i,j}$ = Station weight for the given station i and facet j [-]

The facet parameter f and the value of the exponent n for different facets are given in (Table 2).

Table 2 Parameters of stations weights per facets (in this table d is the distance between the two stations and Δz is the differences in elevations between the two stations) as suggested by Maskey and Venneker (2006)

Facet	Parameters f	Exponent n
Distance	d	-2
Elevation	$\lvert \Delta z \rvert$	-1

The quality of the interpolated data is determined by Jack knife cross validation method. The radius of influence, the distance (F_d) and elevation (F_z) weighting importance factors (Appendix-A) are determined by this method. Here radius of influence (R), F_d and F_z are adjusted by R^2 between the measured and estimated interpolated monthly rainfall for the 35 stations. Quality of the interpolated data is determined by R^2 between the measured and estimated rainfall of a station. The interpolated monthly rainfall estimates are obtained by interpolation of 35 monthly point rainfall stations for the time period 1998 to 2004 by Hykit (Appendix-A). TRMM 3B43 monthly product is validated against the interpolated gauge data at sub-basin level. Pangani is divided into seven sub-basins. The comparison statistics (mean error, RMSE, RRMSE, C_oE, R^2) are calculated in Table 3 to compare and quantify the accuracy of TRMM retrievals.

Table 3 Comparison statistics between TRMM and interpolated monthly rainfall data for 7 sub-basins

Sub-basin	Mean error or bias [mm/month]	RMSE [mm/month]	RRMSE	C_oE (Nash-Sutcliffe)	Co-eff det^n.
sub_1 (Kikuletwa)	3.54	37.94	0.61	0.66	0.69
sub_2 (Ruvu)	2.71	35.63	0.56	0.71	0.71
sub_3 (Part of Pangani main stream)	9.44	33.58	0.58	0.68	0.71
sub_4 (Part of Pangani main stream)	0.19	25.79	0.49	0.76	0.77
sub_5 (Mkomazi)	-0.78	36.94	0.55	0.65	0.68
sub_6 (Luengera)	20.83	57.65	0.59	0.60	0.65
sub_7 (Part of Pangani main stream)	15.53	42.37	0.58	0.61	0.67

An analysis of the results during the seven years period revealed a negative bias in sub-basin 5, that was the only sub-basin where TRMM estimated higher rainfall compared to gauge rainfall. It was seen that after sub-basin 6, sub-basin 7 got the lowest C_0E (0.60) and coefficient of determination (0.67) which was easily understandable as there were very few ground rainfall stations in this sub-basin concentrated at the out let of the basin. The coefficient of determination for 7 sub-basins ranged from 0.65 to 0.77.

5 Conclusion

TRMM satellite rainfall data can play an important part in data sparse regions as well as places where point rain gauges are not available due to lack of access. This can bring a benefit for many countries where point rain gauges are not sufficient to capture the spatial and temporal variability of rainfall. TRMM has the inability to capture high rainfall periods and the rainfall influenced by the orographic effect in the mountainous regions. The use of bilinear weighted interpolation has not given a significant change in comparison statistics but it can be claimed that it is the most realistic way to make comparison between gauge and TRMM rainfall estimates. As the TRMM grid (0.25^0 x 0.25^0) is bit coarse, in reality it is seen that number of rainfall stations are situated in the same grid and TRMM is giving the same rainfall estimate for all the stations in the same grid although they are far apart and their individual rainfall estimate is quite different from each other. Use of bilinear weighted interpolation method gives separate set of estimates for different rainfall stations. Comparison between the interpolated monthly gauge rainfall and TRMM rainfall data has given better results at sub-basin level. This is particularly interesting in the perspective of using this data for hydrological applications.

Acknowledgement

The authors are highly indebted to the Zeeland Fellowship Program of The Netherlands for providing scholarship to the first author of the paper for the M.Sc. in Water Science and Engineering at the UNESCO-IHE, Institute for Water Eductaion, Delft, the Netherlands. They are also thankful to B. M. C Fischer of TU Delft and Dr. Marloes Mul for their support and assistance in data collection. First author is very much grateful to Dr. Ataur Rahman for his valuable suggestions and comments on the work. We would like to also thank the anonymous reviewers of this paper for his positive and constructive comments. Authors would like to acknowledge the TRMM Online Visualization and Analysis System (TOVAS) for TRMM data.

References

Arnold DN, Boffi D, Falk RS (2002). Approximation by quadrilateral finite element, *Mathematics of Computation,* 71, 909–922.

Barros A, Joshi M, Putkonen J, Burbank DW (2000). A study of the 1999 monsoon rainfall in a mountainous region in central Nepal using TRMM products and rain gauge observations. *Geophysical Research Letters*, 27, 22, 3683-3686.

Bowman, Kenneth P (2005). Comparison of TRMM precipitation retrievals with rain gauge data from ocean buoys. *Journal of Climate*, 18, 178–190.

Brito AE, Chan SH, Cabrera SD (2003). SAR image super resolution via 2-D adaptive extrapolation. *Multidimensional Systems and Signal Processing*, 14, 83–104.

Collischonn BW, Collischonn W, Tucci CEM (2008). Daily hydrological modeling in the Amazon basin using TRMM rainfall estimates. *Journal of Hydrology*, 360, 1-4, 207-216.

Dingman SL (2002). Physical Hydrology. Second edition, Waveland Pr Inc.

Fengge S, Yang H, Lettenmaier DP, (2008). Evaluation of TRMM Multisatellite Precipitation Analysis (TMPA) and its Utility in Hydrologic Prediction in the La Plata Basin. *Journal of Hydrometeorology*, 9, 4, 622-640.

Gribbon KT, Bailey DG (2004). A novel approach to real-time bi-linear interpolation. In: Proceedings of the Second IEEE International Workshop on Electronic Design, Test and Applications, pp. 126.

Grimes DIF, Diop M, (2003). Satellite-based rainfall estimation for river flow forecasting in Africa. Part 1. Rainfall estimates and hydrological forecasts. *Hydrological Sciences Journal*, 48, 4, 567-584.

Hazarika MK, Kafle TP, Sharma R, Karki S, Shrestha RM, Samarkoon L (2005). Statistical approach to discharge prediction for flood forecasting using TRMM data. In: Proceedings of the 5th Annual Mekong Flood Forum, Ho Chi Minh City, Vietnam, May 17–18, 2007.

Hughes DA (2006). Comparison of satellite rainfall data with observations from gauging station networks. *Journal of Hydrology*, 327, 2-4, 399-410.

Islam MN and Uyeda H (2007). Use of TRMM in determining the climatic characteristics of rainfall over Bangladesh. *Remote Sensing of Environment,* 108, 3, 264-276.

JAXA, Japan Aerospace Exploration Agency (2006). TRMM data users hand book, JAXA, Japan.

Kuhel H, Sacchi MD, 2003. Least squares wave equation migration for AVP/AVA inversion. *Geophysics*, 68, 262–273.

Murphy AH (1995). The coefficient of correlation and determination as measures of performance in forecast verification. *American Meteorological Society*, Weather Forecasting, 10, 681–688.

Maskey S, Venneker R (2006). Generation of input fields for large Scale hydrological models. UNESCO-IHE, LSHM Doc 0604, 2006-2005.

Thorne V, Coakley P, Rokke P, Grimes D, Dugdale G (2001). Comparison of TAMSAT and CPC rainfall estimates with rainfall, for southern Africa. International Journal Remote Sensing 22, 10, 1951-74.

Uddin S, Al-Dousari A, Ramadan A, Al Ghadban A (2008). Site-specific precipitation estimate from TRMM data using bilinear weighted interpolation technique: An example from Kuwait. *Journal of Arid Environments*, 72, 7, 1320-1328.

Wilk J, Kniveton D, Anderson L, Layberry R, Todd M, Hughes D, Ringrose S, Vanderpost C (2006). Estimating rainfall and water balance over the Okavango River Basin for hydrological applications. *Journal of Hydrology*, 331, 1-2, 18-29.

Zhong L, Chiu H, Rui W, Teng G, Serafno (2002). Online Analysis and Visualization of TRMM and other Precipitation data sets. GSFC Earth Data and Information Services Centre Distributed Active Archive Centre NASA/ Goddard Space Flight Centre, Greenbelt Maryland, USA.

Appendix A Special Data Interpolation kit: Hykit

The interface of spatial data interpolation kit Hykit is shown in Figure 5 below.

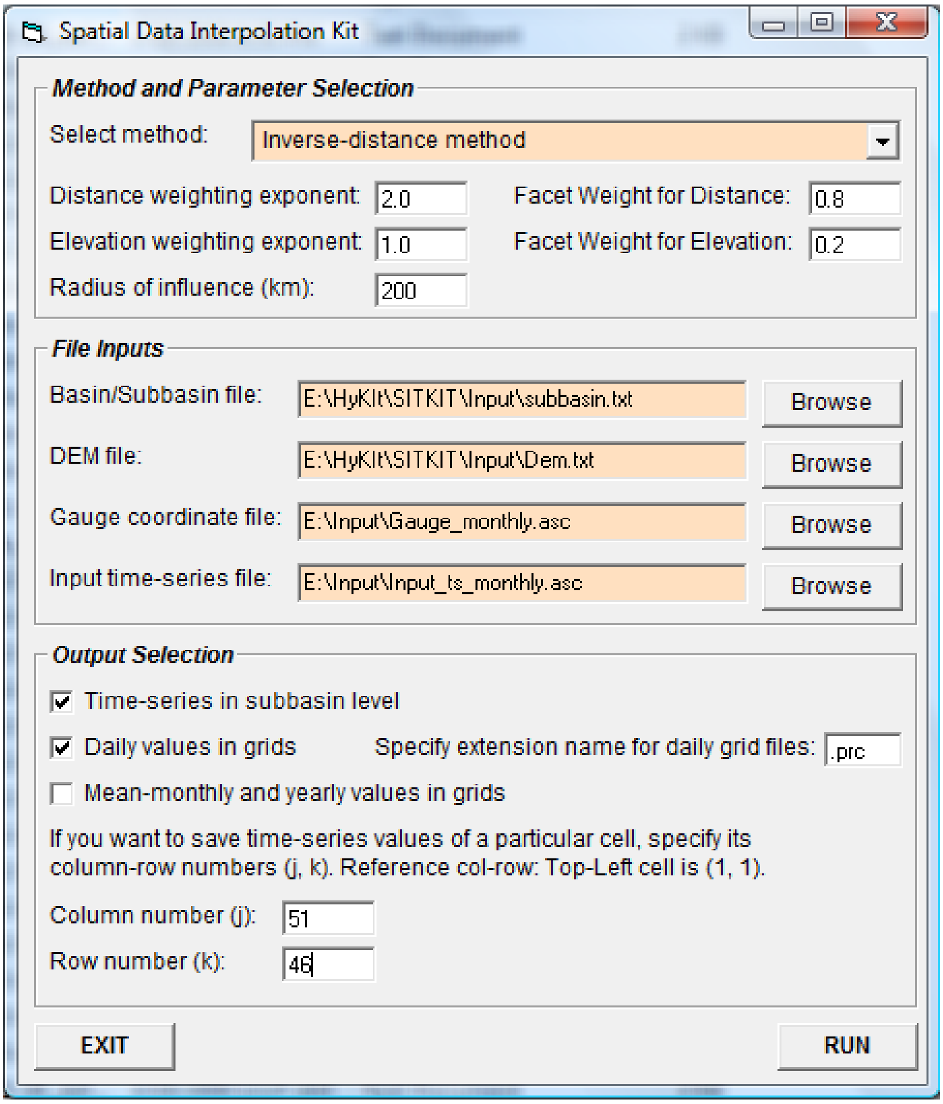

Figure 5 Interface of Hykit

A station's influence in the inverse distance interpolation is assumed to decrease as its distance from the target grid cell increases. *D* is the distance between the station and the target grid cell and τ is the distance weighting exponent typically set to 2, which is equivalent to an inverse-distance-squared weighting function. Elevation weighting allows the model to focus on the vertical range that is specific to a target grid cell, thereby accommodating climate profiles that may vary in slope across the altitudinal range of the data. Here the elevation exponent is typically set to 1.0 which is equivalent to a 1-dimentional inverse distance weighting function. The distance (F_d) and elevation (F_z) weighting importance factors apply a measure of scaling to the vertical dimension by controlling the relative importance of distance and elevation in the model. The influence of horizontal distance on the inter-station correlation seems to be greater overall than that of vertical distance. Thus, F_d is typically set to 0.8 and F_z to 0.2 as shown in Figure 5.