Evaluation of Satellite Based Precipitations and Their Applicability for Rainfall Runoff Modelling in Narayani Basin of Nepal

Sunil Ghaju[1] and Knut Alfredsen[2]
[1] *ITECO-Nepal & TMS JV*
[2] *Depart of Hydraulic and Environmental Engineering, NTNU, Norway*

ABSTRACT

High spatial variability of precipitation over Nepal demands dense network of rain-gauge stations. But to set-up a dense rain gauge network is almost impossible due to mountainous topography of Nepal. Also the dense rain gauge network will be very expensive and some time impossible for timely maintenance. Satellite precipitation products are an alternative way to collect precipitation data with high temporal and spatial resolution over Nepal. In this study, the satellite precipitation products TRMM and GSMaP were analyzed. Precipitation was compared with ground based gauge precipitation in the Narayani basin, while the applicability of these rainfall products for runoff simulation were tested using the LANDPINE model for Trishuli basin which is a sub-basin within Narayani catchment. The Nash-Sutcliffe efficiency calculated for TRMM and GSMaP from point to pixel comparison is negative for most of stations. Also the estimation bias for both the products is negative indicating under estimation of precipitation by satellite products, with least under estimation for the GSMaP precipitation product. After point to pixel comparison, satellite precipitation estimates were used for runoff simulation in the Trishuli catchment with and without bias correction for each product. Among the two products, TRMM shows good simulation result without any bias correction for calibration and validation period with scaling factor of 2.24 for precipitation which is higher than that for gauge precipitation. This suggests, it could be used for runoff simulation to the catchments where there is no precipitation station. But it is too early to conclude by just looking into one catchment. So extensive study need to be done to make such conclusion.

Keywords: TRMM, GSMaP, satellite rainfall, rainfall runoff model, LANDPINE

1. BACKGROUND

For the planning, development and operation of water related projects in any region, knowledge of precipitation patterns is essential for predicting drought and floods. Precipitation data with high temporal and spatial resolution is essential input data for any hydrological model for timely forecast of river flows.

In a country like Nepal, the widely used method of rainfall estimation in a given catchment is the method of interpolation of point measurements from a network of ground based meteorological stations. Also the nature of precipitation varies due to orographic and rain shadow effects caused by high hills and mountains. This creates localized rainfall and varies greatly within small region. So for accurate estimation

of precipitation over the area, precipitation stations should be placed close together. This demands a high number of precipitation stations which is very costly regarding installation and maintenance. One cost effective alternative is to estimate rainfall using data from a combination of satellite sensors. However, the estimates made from satellite observations contain random error and bias because of the indirect nature of the relationship between the observation and the precipitation, the inadequate sampling and algorithm imperfections (Xie and Arkine, 1997). So before using the satellite products, their accuracy must be checked first.

Several studies were carried out on Satellite Rainfall Estimates (SRF) for different parts of Nepal with different satellite rainfall products. Statistical comparison with ground based rainfall stations and runoff simulations were performed at different time using several satellite rainfall products some of which are listed in references below. The results of these studies are discussed and compared with the current studies in the following sections.

In this study, two types of satellite precipitation data are selected. They are the Tropical Rainfall Measuring Mission (TRMM-3B42) and the Global Satellite Mapping of Precipitation (GSMaP) based on the temporal and spatial resolution. These products are compared to the ground based precipitation stations to access the quality, and then later analyzed for inflow forecasting in Upper Trishuli catchment within Narayani basin (Figure 1 and Figure 2).
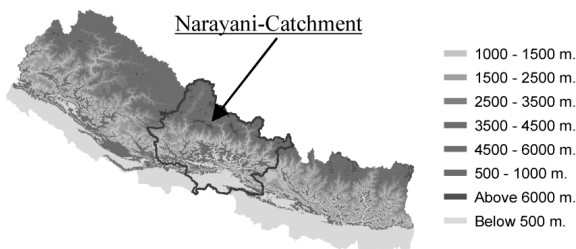

Figure 1: Elevation Map of Nepal

2. DATA SOURCE AND METHODOLOGY

2.1 Rain Gauge Data and its quality

Rain gauges collect precipitation data at particular point during a specified interval of time. In Nepal, interval of measurement is one day from 03:00 UTC to 03:00 UTC (i.e. 08:45 am to 08:45 am local time). The rain gauge data for the Narayani basin were collected from the Department of Hydrology and Meteorology, Ministry of Environment, Science and Technology, Government of Nepal. Data were collected from 2001 to 2007 for 79 different stations as indicated in Figure 2 with their station index. The data quality was also checked for individual stations.
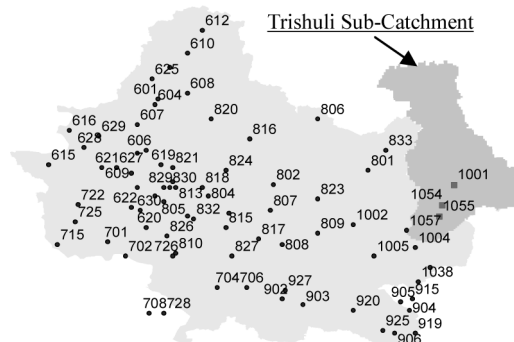

Figure 2: Narayani Basin

2.2 Satellite Precipitation Estimates

There are various types of global satellite data available released by different organizations which can be downloaded freely from the internet. In this paper, TRMM and GSMaP were analyzed. The properties of these satellite products are listed in table below:

Table 1: Detail of satellite based data

Product Name	TRMM-3B42	GSMaP MVK+
Temporal Resolution	3-hourly	1-hourly

Spatial Resolution	0.25 degree	0.1 degree
Data extent	1998 to present	2003 to 2006
PMW data	Y	Y
IR data	Y	Y
Adjusted by gauge	Y	Y
Data format	HDF	Binary
One day Definition (UTC)	22:30-22:30	00:00-00:00

Where, Y – Yes, PMW – Passive Microwave, IR – Infrared

2.3 Data Processing

The global precipitation data needs data processing before it could be used in a comparison with the ground based gauge precipitation data in runoff modeling. Data is processed for point to pixel comparison. Data processing is done using different Python programs made for spatial rainfall assessment (Abdella and Alfredsen, 2010).

2.4 Methods

Daily accumulated precipitation for TRMM and GSMaP product are compared (point to pixel comparison) with ground based precipitation measurements. Visual and statistical methods like scatter plot, correlation coefficient (RR), Nash–Sutcliffe Coefficient of efficiency (R^2), Estimation Bias (EB), Satellite Conditional Probability of Detection (CPOD_S) and Gauge Conditional Probability of Detection (CPOD_G) were used to compare satellite precipitation with gauge precipitation. Gauging stations were omitted for analysis when there were lots of missing data i.e. when ratio of gauge precipitation days and satellite precipitation days is less than 0.8.

3. RESULTS AND DISCUSSION FOR PRECIPITATION COMPARISON

For comparison, days were excluded when precipitation on both satellite and gauge is recorded zero. Also days with missing data either in satellite or in gauging stations was excluded from the comparison.

3.1 Point to Pixel Comparison

Point to pixel comparison is done between point measurement in gauge stations and the areal mean of the satellite estimate of rainfall over the pixel area. So the most reliable result can be expected if the pixel resolution is high. The TRMM data were compared from 2001 to 2007 and GSMaP data were compared from 2003 to 2006 based in their availability.

First the statistical parameters were compared for each gauge stations and presented in Table 2. During comparison one station with index 1054 is excluded for which R^2 is observed very low (-27.26). The results of the comparison are summarized in Table 2. Moreover in both cases the Nash-Sutcliffe efficiency tends to decrease with increased elevation. Figure 3 and Figure 4 show R^2 as a function of elevation for the TRMM and GSMaP respectively.

Table 2: Comparison of Result for Two Satellite Products

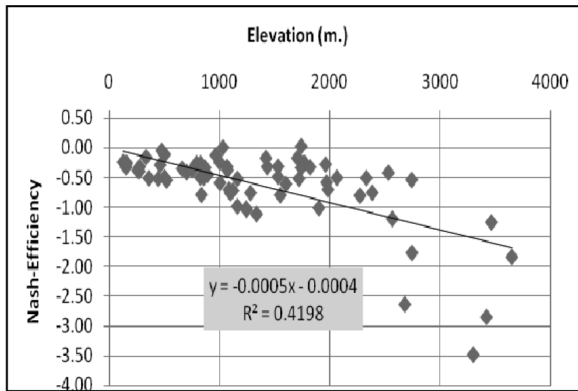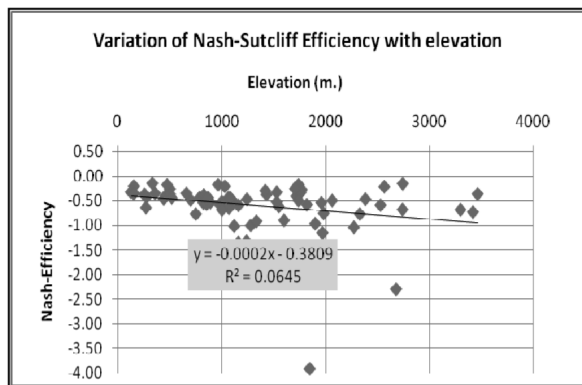| Statistical Parameters	TRMM		GSMaP	
	Max.	Min.	Max.	Min.
R^2	3.48	0.03	0.15	-3.92
MAD (mm/per precipitation day)	30.53	5.09	30.97	17.51
CPOD_S	0.92	0.64	0.90	0.66
CPOD_G	0.77	0.14	0.81	0.11
EB (%)	331.13	-68.44	148.93	-81.81
RR	0.24	00	0.19	00

Figure 3: R² Vs. Elevation (TRMM)

Figure 4: R² Vs. Elevation (GSMaP)

After comparing for individual gauge stations, TRMM and GSMaP data were compared with gauge data as whole. Total of 115419 data sets from 79 stations were analyzed for statistical comparison from year 2003 to 2006. The results of point to pixel comparison for the two different products are listed in Table 3.The result in the table shows similar kind of error statistics except for the estimation bias.

Table 3: Comparison of Satellite Products

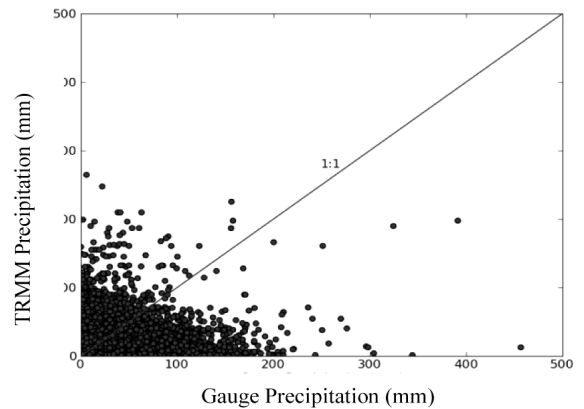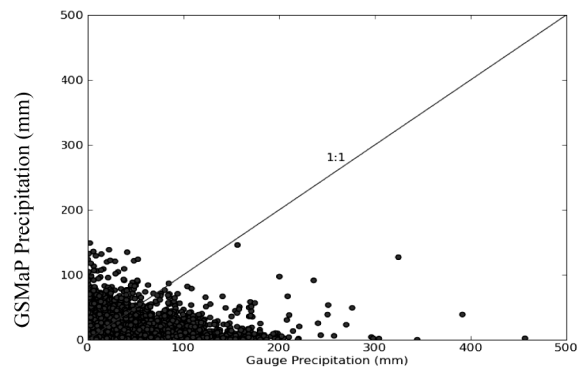| Statistical Parameters	TRMM	GSMaP
R^2	-0.30	-0.33
MAD(mm/perprecipitationday)	17.31	17.25
CPOD_S	0.77	0.75
CPOD_G	0.54	0.51
EB (%)	-28.48	-56.64
RR	0.04	0.02

Figure 5: Scatter Plot for TRMM

Figure 6: Scatter Plot for GSMaP

The negative estimation bias for both precipitation indicates that TRMM and GSMaP both are under estimating the precipitation than gauge precipitation. But the extent of under estimation is almost double for GSMaP as compared to TRMM. This can also be observed in scatter plots where high numbers of dots below the 1:1 line is observed for GsMap (Figure 6) in comparison to TRMM (Figure 5).

3.2 Spatial Variability within a Single Pixel

In point to pixel comparison between gauge and satellite rainfall estimates, any two or more gauging stations which fall within the same pixel will share the same pixel value. For the evaluation of spatial variability within the same pixel, the data with the highest resolution

(GSMaP with resolution 0.1x0.1 degree) was used to identify within pixel variability. A total of six pixels were identified that contains more than one gauging stations. The results of variability within single pixel are shown in Table 4.

Table 4: Variation within Single Pixel

S.N	Stations Pair	Average elevation (masl)	R^2
1	Rampur & Bharatpur	230.5	0.177
2	Sallyan & Pamdur	1080	0.213
3	Darbang & Kuhun	1355	0.002
4	Tatopani & Bega	1506.5	0.001
5	Thamachit & Dhunche	1914.5	0.012
6	Jomsom & Thakmarpha	2655	0.235

The R^2 value indicates that the gauge precipitation measured within a single pixel shows great variability. This effect may be due to orographic and rain shadow effects. After point to pixel comparison of precipitation measurements, the precipitation data were analyzed as input to the LANDPINE model.

4. LANDPINE MODEL

LANDPINE is a distributed rainfall – runoff model developed to study the how changes in landuse influence runoff. The model was developed as part of a 3-year research program, called HYDRA, whose main objective was to investigate how human activities may affect flood regimes in rivers (Rinde, 1998).
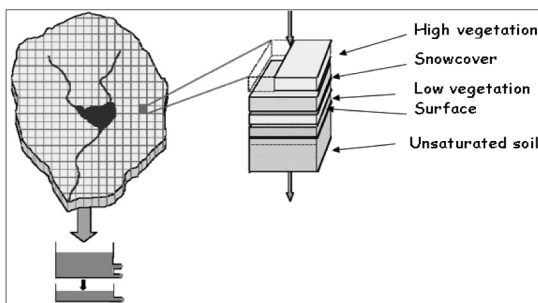

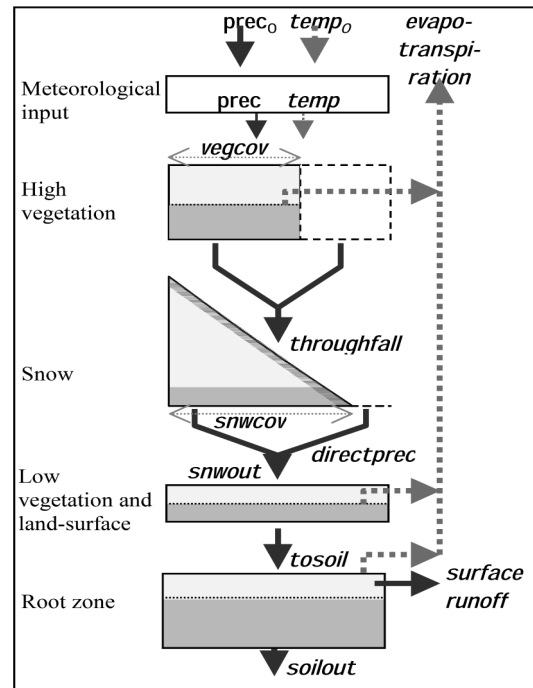
Figure 7: Model structure of LANDPINE (Rinde, 1998)

Figure 8: Representation of Grid-Cell Processes (Rinde, 1998)

Each grid cell in LANDPINE model is categorized into four distinct hydrological zones, high vegetation, snow, low vegetation and soil zone. The different processes in these zones includes interception in high and low vegetation, storage of water on the ground surface, evapotranspiration, accumulation and melting of snow, infiltration, retention of water in the soil and generation of surface runoff and outflow from the soil. These different processes in distinct hydrological zones are described in Figure 8 while Figure 7 represents the general model structure of LANDPINE. The processes are briefly described below:

4.1 Meteorological Input

The model needs precipitation and temperature as input. Data can be imported to the model as time series of point data from a number of stations or as distributed data like radar or satellite products. In case of point measurements, the recorded precipitation is first corrected for

catch deficiencies. Precipitation is defined as rain or snow by the threshold temperature, and a separate correction factor is used if precipitation is in the form of snow. Then spatial interpolation is performed to produce the temperature and precipitation distribution over the catchment. Both temperature and precipitation values are adjusted for variation in elevation over the region. For temperature, dry and wet adiabatic lapse rate is used, and for precipitation an elevation gradient is used.

4.2 High Vegetation

Five parameters define the high vegetation, the maximum and the minimum seasonal leaf-area index, the specific storage capacity per leaf-area index unit, the vegetation coverage factor, and the average height of vegetation stand. In each time step an interception capacity and a potential evaporation rate is calculated along with actual values for interception and evaporation. In the high vegetation zone, the exchange of air mass is more efficient which results in higher potential evaporation. Actual evaporation from interception in high vegetation is taken as the potential evaporation rate multiplied by the vegetation cover factor. The actual evaporation is also decreased for the precipitation as snow then for rain. Precipitation that exceeds the interception capacity of the high vegetation forms the throughfall to the ground or the snow surface.

4.3 Snow

Precipitation is defined as rain or snow by a threshold temperature. If the air temperature is less than the threshold, throughfall from high vegetation goes to increase the snow layer while if temperature is greater than threshold, the throughfall goes to increase the liquid water content in the snow. If there is no snow present, it passes on to the lower interception storage. A distribution factor is used to specify the snow distribution patter within grid cell. If distribution factor is set to one, the snow pack becomes homogenous. If it is set to 2, the maximum value becomes twice the average value and the minimum vale becomes zero. If it is set greater than 2, only partial snow cover will be simulated. In last two cases, the snow distribution varies linearly within grid cell.

Snow melt is calculated on the basis of actual air temperature, a threshold for melting and a melt factor according to the degree day principle. The slower snow melt in the forested area is considered by a reduction factor. If air temperature is lower than the melt threshold, refreezing of liquid water in the snow is calculated as a refreeze factor times the temperature below the threshold. Snow melt as well as a refreezing is assumed to be homogenous across the snow surface. Melted snow is added to the liquid water content in the snow. A separate parameter specifies the maximum relative amount of such water that can be withheld in the snow. If liquid water crosses the maximum limit in the snow pack then outflow from snow will occur. Throughfall on the snow-free parts contributes directly to the outflow.

4.4 Low Vegetation and Land-Surface

Similar to the high vegetation, interception storage is also computed for the low vegetation. Here, the storage is lumped together with a wetting storage for the land surface. The storage is filled by outflow from the snow routine.

When the storage capacity of low vegetation is exceeded, excess water may infiltrate to ground, or form surface runoff if the infiltration capacity is exceeded. As for the high interception storage, actual evaporation is assumed computed as a fraction of a potential rate. This rate is now however reduced according to the actual evaporation that has already occurred in the high interception storage.

4.5 Root Zone

Any outflow from low vegetation and land-surface will infiltrate to root zone depending upon the infiltration capacity and field capacity of the soil while the excess water will contribute to surface runoff. The amount of infiltration and surface runoff in particular time will depend upon the soil moisture content with respect to the field capacity of the soil which is illustrated in Figure 9. It shows that a larger portion of incoming water will be retained in the zone if the soil is dry than if it is wet. Depletion of soil water can only happen through transpiration until the permanent wilting point of vegetation is reached.
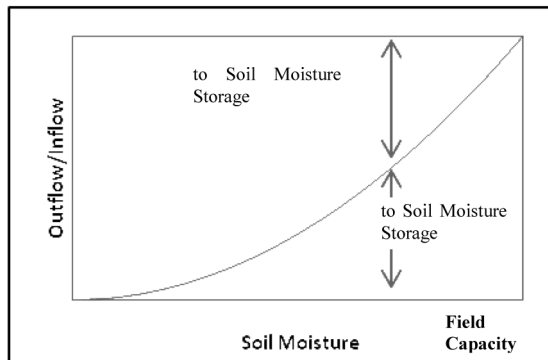

Figure 9: Soil Moisture and Runoff Relationship

4.6 Ground Water and Surface Water

The ground water and surface water in the model is not described in a distributed manner. Instead the average sum of soil and surface runoffs

from all the grid cells are provided as input to a lumped response routine consisting of two linear tanks (Figure 7). The lower outlet of upper tank response to unsaturated soil moisture condition while the upper outlet response to the saturated soil moisture condition during precipitation events. The lower tank in the model represents the flow from ground water reservoir.

5. SIMULATIONS

A LANDPINE setup for Trishuli 3A catchment (Timalsina, 2008) was used to evaluate runoff computations from satellite precipitation. Trishuli is a sub-catchment of the Narayani basin with catchment area of 4542 km². The basic input parameters for the model like elevation data and land use are derived from USGS map while field capacity and soil maps was prepared based on site visits, USGS land use maps and elevation maps. Other data like vegetation type, vegetation cover, leaf area indices etc. were prepared based on USGS vegetation maps. All maps had a resolution of 1 x 1 km². Figure 10 shows some of the distributed input data that were developed for the Trishuli catchment. After preparing these basic parameters, daily gauge precipitation data, satellite precipitation data and temperature data were prepared for simulation. A Python program was used to prepare satellite data for LANDPINE from its raw form.
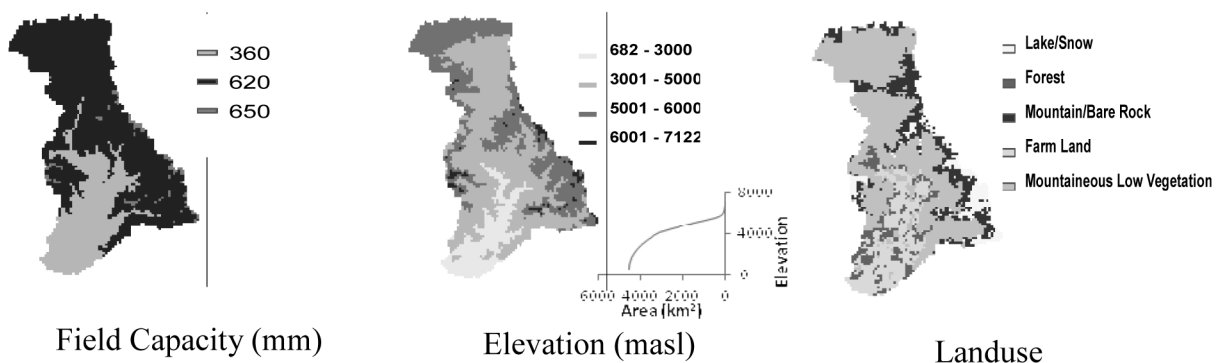

Figure 10: Geographical data for Trishuli Catchment (Data Source: Timalsina, 2008)
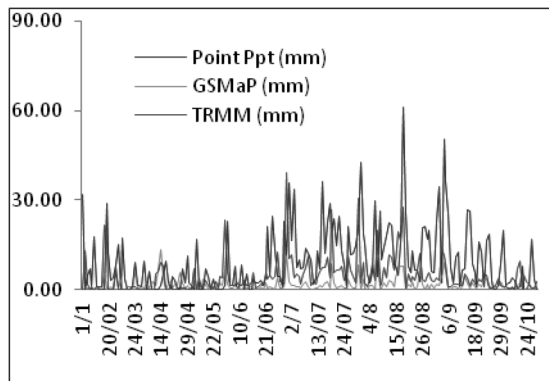
Figure 11: Mean Basin Precipitation for 2003

Three precipitation gauges within the Trishuli catchment with station index no. 1001 (Timure), 1004 (Nuwakot) and 1055 (Dhunche) were used to evaluate point precipitation data while both TRMM and GSMaP data were used to evaluate satellite rainfall products. Based on the data availability, the gauge precipitation and TRMM data were simulated for 2001 to 2004 while GSMaP was simulated for 2003 to 2004. The sample mean daily basin precipitation in 2003 for these three types of data as calculated by the model is presented in Figure 11. The figure shows the clear indication of under estimation by satellite products and this is largest for GSMaP data.

The annual sum for the year 2003 were calculated to 1823.64 mm, 834.70 mm and 324.72 mm for gauge precipitation, TRMM and GSMaP respectively. These differences were also reflected in the variation in the rain correction parameter obtained during calibration. The best results for the simulations were observed when the rain scaling factors for gauge, TRMM and GSMaP were fixed to 1.273, 2.24 and 6.53 respectively. This is in practice just a linear scaling of the precipitation volume. The satellite precipitation data were also simulated using local bias corrections based on the observed precipitation in the region. The bias correction method is generally applied method to adjust precipitation for satellite products. The simulation result was not as good as the result obtained by using single rain scaling factor mentioned above. The simulation results for gauge, TRMM and GSMaP precipitation inputs are presented in Figure 12, 13 and 14 respectively where R^2 represents the Nash-Sutcliffe efficiency and Acc. Diff. represents accumulated difference.
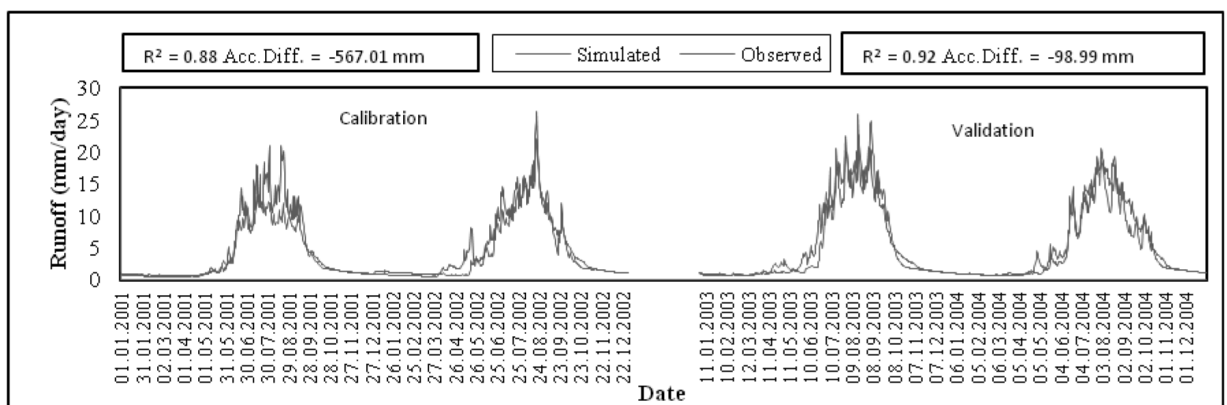

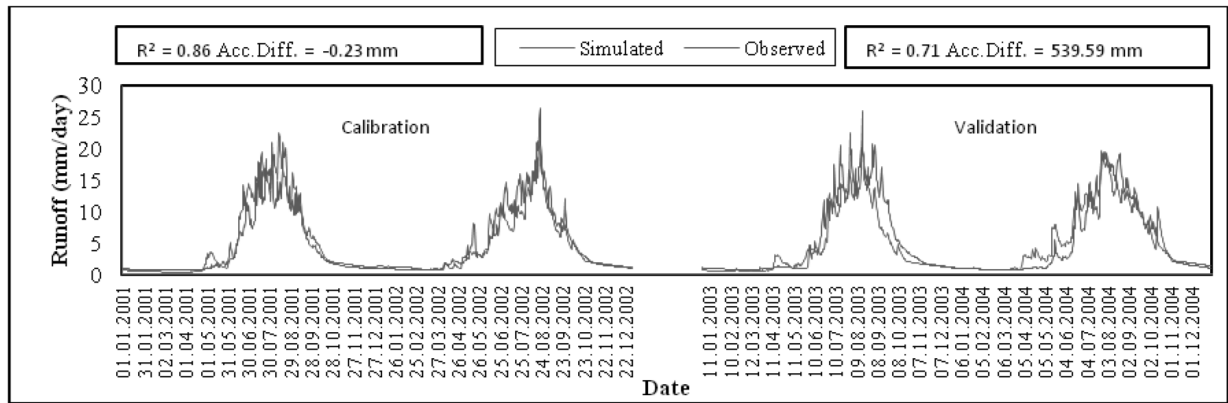Figure 12: Simulation Result for Gauge data
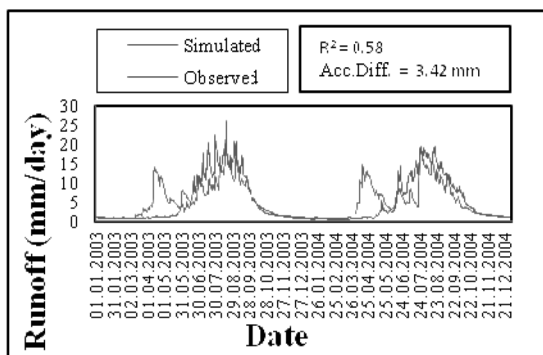
Figure 13: Simulation Result for TRMM data

Figure 14: Simulation Result for GSMaP

6. CONCLUSION AND DISCUSSION

From the result of point to pixel comparison of TRMM to gauge data, it is observed that for most of the stations Nash-Sutcliffe efficiency is negative. The correlation study of Nash efficiency with the elevation shows that there is slight improvement of Nash efficiency when the elevation decreases. The reason behind the low performance at high elevation may be due to orographic and rain shadow effects in high hills within small pixel area. The negative estimation bias is mostly due to under estimation by the TRMM product as seen in the scatter plot. Similar conclusions were also drawn by Islam (2009) for Nepal.

Similarly, the result for GSMaP shows a slightly higher Nash efficiency than TRMM. But there is

a higher negative value for estimation bias due to even higher under estimation of precipitation than TRMM.

The inter comparison of two satellite precipitation product were also done for the period 2003 to 2006. This shows that the TRMM has slightly better result than the GSMaP product. The estimation bias for GSMaP is almost double of the TRMM but in negative side (-56.64% and -28.48% respectively) which indicates higher under estimation by GSMaP product than by TRMM product. The higher under estimation of precipitation by GSMaP is also concluded in the recent study by Shrestha (2011) where she found that the precipitation is underestimated by 48% based on spatial comparison with gauge data for whole Nepal.

After PO-PI comparison, satellite precipitation estimates were used for LANDPINE runoff simulation for Trishuli catchment with and without bias correction for each product. TRMM shows the best simulation result without any bias correction but with a linear scaling of input precipitation for calibration and validation period. This indicates that the TRMM data is underestimated by certain factor but with a similar pattern of precipitation as recoded in gauge stations. Similarly, for GSMaP good simulation result was obtained with a scaling factor of 6.53 for precipitation which shows the

even larger under estimation of rainfall by the GSMaP precipitation product. Simulated runoff patterns were very similar to the observed one except for the months of May and April when relatively high precipitation were observed. This is because the precipitation recorded by GSMaP during May and April were on the higher side in comparison to gauge and TRMM precipitation records (Figure 10).

Finally, the result suggests that TRMM could be used with a linear scaling but without any bias correction for runoff simulation in hydrological models. Hence the results indicate that it could be used for runoff simulation in catchments where there is no detailed precipitation data as long as the scaling factor can be derived. This is a promising result, but more studies in different catchments are needed to strengthen this conclusion.

REFERENCES

Abdella, Y. S. (2005). Application and Comparison of a Distributed Hydrological Model for a Tropical and a Temperate Catchment. Trondheim, Norwegian University of Science and Technology.

Abdella, Y.S. and K. Alfredsen (2010). A GIS toolset for automated processing and analysis of radar precipitation data. Computers and Geosciences 2010 36(4):422-429

Barros, A. P. (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.

Islam, D. M. N. (2009). Analysis of TRMM data in Monitoring Rainfall over Mountainous Regions, SAARC Meteorological Research centre (SMRC). *SMRC Report No. 27.*

Tian, Y., C. D. Peters-Lidard, et al. (2010). Evaluation of GSMaP Precipitation Estimates over the Contiguous United States. Journal of Hydrometeorology 11(2): 566-574.

Timalsina, N. P. (2008). Hydrological modelling of the Upper Trishuli 3A Cathchment in Nepal – Comparison of Lumped and Distributed models. Department of Hydraulic and Environmental Engineering, NTNU.

Rinde, T. (1999). Landpine, a hydrologic model for simulation of how landuse changes inflect runoff (Note) In Norwegian: Landpine. En hydrologisk modell for simulering av arealbruksendringers innvirkning på avrenningsforhold.

SHRESTHA, M.S. (2011). Bias-Adjustment of Satellite-Based Rainfall Estimates over the Central Himalayas of Nepal for Flood Prediction Department of Civil and Earth Resources Engineering, Kyoto University, Japan. *Phd Dissertation.*

Shrestha M.S, Artan G.A, Bajracharya S.R, Sharma R.R (2008). Using satellite-based rainfall estimates for streamflow modelling: Bagmati Basin. Journal of Flood Risk Management **1**(2): 89-99.