Evaluation of precipitation data sets along the Himalayan front
Cristoff Andermann, Stéphane Bonnet, Richard Gloaguen

To cite this version:

HAL Id: insu-00612265
https://hal-insu.archives-ouvertes.fr/insu-00612265
Submitted on 28 Jan 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Evaluation of precipitation data sets along the Himalayan front

C. Andermann  
Geosciences Rennes, Universite de Rennes 1, Campus de Beaulieu, F-35042 Rennes, France  
(christoff.andermann@univ-rennes1.fr)  
Geosciences Rennes, CNRS/INSU, UMR 6118, Campus de Beaulieu, F-35042 Rennes, France  
Remote Sensing Group, Geology Institute, TU Bergakademie Freiberg, Bernhard-v.-Cotta Str. 2,  
D-09596 Freiberg, Germany

S. Bonnet  
Geosciences Rennes, Universite de Rennes 1, Campus de Beaulieu, F-35042 Rennes, France  
Geosciences Rennes, CNRS/INSU, UMR 6118, Campus de Beaulieu, F-35042 Rennes, France

R. Gloaguen  
Remote Sensing Group, Geology Institute, TU Bergakademie Freiberg, Bernhard-v.-Cotta Str. 2,  
D-09596 Freiberg, Germany

[1] Precipitation is one of the main factors which controls surface processes and landscape morphology. Large orogenic belts such as the Himalayas control precipitation distribution as a result of orographic effects due to their prominent relief. However, precipitation is difficult to monitor because mountain regions are largely inaccessible and therefore not sufficiently covered by ground-based gauge stations. The complexity of orographic effects resulting from the interaction between elevation and climatic processes and the lack of precise meteorological data thus limit our understanding of climatic influence on landscape formation. Therefore, high-quality precipitation observations with good spatiotemporal coverage are needed. Here we evaluate five gridded precipitation data sets derived from remote sensing and interpolation of gauge data with ground-based precipitation measurements. First, we evaluate the bulk error of each data set, then we evaluate the temporal quality of data within five watersheds, and last we compare the spatial performance along seven swath profiles across strike to the Himalayan range in Nepal. Our evaluation shows that the data sets vary significantly along the orographic front and get more consistent toward the adjacent low-relief domains, while bulk errors are largest during monsoon season. In particular, where topographic gradients are important, the resolution of gridded data sets cannot incorporate small-scale spatial changes of precipitation. We show that the data set derived from interpolation of gauge data performs best in the Himalayas. This study gives an overview on the applicability of precipitation data sets within the Himalayan orographic domains where relief has a pronounced impact on precipitation.
1. Introduction

The spatial distribution and the temporal variability of precipitations governs vegetation growth, hydrology and surface mass transport on Earth [e.g., *Istanbulbulluoğlu and Bras*, 2006], whereas precipitation is proposed to be a first-order control on landscape morphology [Tucker and Slingerland, 1997; *Anders et al.*, 2008; *Bonnet*, 2009] as well as on the interplay between climate, erosion and tectonics [Willett, 1999; *Bonnet and Crave*, 2003; Reiners et al., 2003; Whipple and Meade, 2006; Whipple, 2009]. Consequently, precipitation measurements with good spatial and high temporal resolution, recorded over a long time span are crucially needed to better understand the impact of precipitation on landscape [Barros et al., 2006]. This is particularly the case in mountains where local extreme events are much more frequent than in the adjacent flatlands [Wulf et al., 2010].

Mountain topography controls regional precipitation patterns through orographic effects [Roe et al., 2003; Roe, 2005; Bookhagen and Burbank, 2006; Bookhagen and Strecker, 2008]. In mountainous environments, precipitation distribution can also change on short distances and within short periods of time [Anders et al., 2006]. High amplitude rainfall events are often very localized [Nesbitt and Anders, 2009], whereas their impact on landscape forming can be enormous. Landslides, for example, are largely controlled by precipitation intensity and accumulation over time [Gabet et al., 2004; Dahal and Hasegawa, 2008]. Such extreme precipitation events are usually localized and therefore not recorded by widely scattered ground-based meteorological stations (e.g., in Nepal).

Remotely sensed precipitation data are now available at moderate to high resolution, some over long time spans to allow a reasonable comparison of local precipitation patterns with respect to landscape morphology. Several gridded data sets, with varying temporal and spatial resolution, are available (Figure 1). The measurements are derived from ground and/or satellite observations. Most remotely sensed precipitation data sets are based on multisensor algorithms, merging ground measurements, low-orbiting geostationary satellite observations, and global ground-based gauge databases [e.g., Yatagai et al., 2009]. Rain gauge stations provide highly accurate local information for the point of observation but their spatial representativity is questionable [Tustison et al., 2001], particularly in case of local rainfall gradients such as ridge-valley gradients [Barros et al., 2004; Bhatt and Nakamura, 2005; *Anders et al.*, 2006]. In the Himalayas, precipitations indeed varies between ridges and valleys [Barros et al., 2004], therefore a single rain gauge station does not register variability at the scale of kilometers. Rain gauge data must consequently be compared to their associated pixel value of any remotely measured rainfall information with great caution, especially at high temporal resolution. Gridded data sets (satellite observations or spatial interpolation of gauge data) provide good information on the spatial precipitation distribution, however with potentially large errors within each point of the grid space (pixel), particularly when resolution of the data is larger than the spatial variability of rainfall.

Precipitation measurements from remote platforms are carried out using active precipitation radar (PR), passive microwave radiometer (MWR), such as Tropical Rainfall Measuring Mission (TRMM), Microwave Imager (TMI) and infrared radiometer (IR) sensors [Ushio et al., 2009; Huffman et al., 2007]. The PR sensor is an active precipitation radar, which can record the three-dimensional structure of rainfall distribution [Kummerow et al., 1998]. IR observations are made at the top of the clouds and are therefore indirect measurements [Huffman et al., 2007]. Microwave measurements detect the radiation emitted by the water fraction in the vertical profile of the atmosphere [Kubota et al., 2007]. For all these techniques, differences between ground and remote observed quantities come from the inability to incorporate local conditions in the sensor algorithms. In particular property changes of precipitates, affecting polarization, scatter and absorption, due to slope, snow, ice and orographic effects, are not well accounted for [Vicente et al., 2002]. In general, short-lasting and low precipitation rates as well as frozen precipitates are badly detected from the remote sensors. Hence, at high elevation, where precipitation comes mainly as snow, as light drizzle and during short intense storms remote measurements often underestimate the actual rates. The TRMM satellite system is so far the only platform in operation...
designed specifically for rainfall monitoring from space [Kummerow et al., 2000]. It will be succeeded by the Global Precipitation Measurement Mission (GPM), in 2013.

[6] In this study we compare four gridded spatiotemporal precipitation data sets (Figure 1) and one mean annual compilation of raw TRMM-2B31 data by Bookhagen and Burbank’s [2006] TRMM-2B31 data are not shown here since only one mean layer for 10 years (1997–2007) is available. TRMM-2B31 has a ~4 km spatial resolution and is derived from PR and TMI sensor input.

2. Data and Methods

2.1. Gridded Precipitation Data Sets

[7] We give here a general description of the precipitation data sets tested here (Figure 1). More technical specifications can be found in Text S1.

2.1.1. APHRODITE

[8] APHRO_MA_V1003R1 (Asian Precipitation Highly Resolved Observational Data Integration Towards Evaluation of Water Resources, Monsoon Asia, Version 10, hereafter referred to as APHRODITE) data set is developed by a consortium...
between the Research Institute for Humanity and Nature (RIHN) Japan and the Meteorological Research Institute of Japan Meteorological Agency (MRI/JMA). This consortium develops precipitation products with varying resolution and for several Asian subregions. We used here the latest version daily data set for monsoon Asia (60–155°E and 15°S–55°N) [Yatagai et al., 2009; Xie et al., 2007]. APHRODITE is a distance weighted interpolated data set from precipitation gauge stations. Depending on availability, between 5,000–12,000 stations are considered for the interpolation. Data are available for a statistically robust time span of more than 50 years, 1951–2007. This data set has daily resolution and 0.25° (≈30 km) spatial resolution. The interpolation algorithm incorporates orographic correction of precipitation. Yatagai and Kawamoto [2008] show for the Himalayas that an earlier version of APHRODITE correlates well with monthly active TRMM-PR (2A25) measurements, however they show that TRMM-2A25 considerably underestimates precipitation with respect to APHRODITE.

2.1.2. CPC-RFE

CPC-RFE 2.0 (Climate Prediction Center–Rainfall Estimates) is a precipitation product for the South Asian region published by the CPC of National Oceanic and Atmospheric Administration (NOAA), United States Agency for International Development (USAID) and United States Geological Survey (USGS). The product provides real time daily precipitation information with a good spatial resolution of 0.1° (~10 km) for the area 70–110°E and 5–35°N. Data from CPC-RFE are available since May 2001 and continuously updated. RFE2.0 combines 4 different primary products, of which, one is a rain gauge network and three are remotely sensed. The four input products are (1) GTS global gauge network (~1000 stations); (2) GOES Precipitation Index (GPI), a precipitation index derived from Geostationary Operational Environmental Satellites (GEOS) geostationary weather satellites (IR data); (3) Special Sensor Microwave/Imager (SSM/I) observations; and (4) Advanced Microwave Sounding Unit-B (AMSU-B), on board of NOAA-K, -L, -M satellites. In general all data sources have similar large-scale distribution patterns [Xie et al., 2002]. The three satellite products are merged through maximum likelihood estimation methods. In comparing CPC-RFE and ground gauge stations, Shrestha et al. [2008] have run a hydrological model in the Bagamati basin of the middle and lower Nepal Himalayas. They show that CPC-RFE capture the occurrence of rainfall events but considerably underestimate rainfall amounts.

2.1.3. GSMaP

The Global Satellite Mapping of Precipitation, passive microwave radiometer (GSMaP MVK+) data set was developed in order to provide high-
precision and high-resolution global precipitation maps from satellite observations. The project is sponsored by Core Research for Evolutional Science and Technology (CREST) of the Japan Science and Technology Agency (JST), by Japan Aerospace Exploration Agency (JAXA) and the Precipitation Measuring Mission (PMM) Science Team. GSMaP data are a global data set (60°N/S), available since the end of November 2002 and is provided in almost real time (with a ∼10 month data gap in 2007). The data have 0.1° (∼10 km) spatial resolution and 1 h temporal. The project aims to develop an advanced microwave radiometer algorithm based on a deterministic rain-retrieval algorithm and the production of precise high-resolution global precipitation maps [Ushio et al., 2009; Kubota et al., 2007]. The data incorporate MWR measures from TRMM-TMI, SSM/I, Advanced Microwave Scanning Radiometer–EOS on board of AQUA satellite (AMSR-E), AMSU-B and IR. Because of its high spatiotemporal resolution this data set is potentially the most interesting for analyzing climatic influence on surface processes and the links between rainfall distribution and topography. Dinku et al. [2009] compared GSMaP MVK+ with several other satellite derived precipitation data sets and gauge stations over the whole of Colombia. They report that GSMaP MVK+ underestimates precipitation in mountains, where the topography is complex.

2.1.4. TRMM-3B42

[11] The Tropical Rainfall Measuring Mission (TRMM) is a joint collaboration between JAXA and the United States of America National Aeronautics and Space Administration (NASA). TRMM-3B42, is a global multisatellite precipitation analysis data set. It combines several instruments, has a 0.25° (∼30 km) spatial and 3 h temporal resolution, and is available within a global belt, 50°N/S latitude [Huffman et al., 2007; Kummerow et al., 2000]. Basically it is a set of MWR estimates from TRMM-TMI, SSM/I, AMSR-E and AMSU-B, whereas missing pixels are filled with IR observations compiled by CPC [Huffman et al., 2007]. The data are corrected with the monthly field ratios between TRMM-3B43 (monthly compiled version of 3B42) and gauge stations. TRMM-3B42 has been applied successfully for measuring precipitation patterns in many studies on a global [Tian and Peters-Lidard, 2010] and local scale [Kamal-Heikman et al., 2007; Bookhagen and Strecker, 2008; Bookhagen, 2010]. However, underestimation in mountainous regions, in particular with high snowfall contribution [Kamal-Heikman et al., 2007], has been reported.

2.1.5. TRMM-2B31

[12] Bookhagen and Burbank [2006] have developed their own precipitation compilation from primary TRMM-2B31 orbital data (not gridded). Despite a common platform and name, TRMM-2B31 and TRMM-3B42 data set do not use the same sensor (except for TMI). TRMM-2B31 is principally derived from the active PR sensor, found only on board of the TRMM satellite [Bookhagen and Burbank, 2006]. Here TRMM-TMI is used to fill unobserved areas. TRMM-2B31 data have a spatial resolution of 0.05° (∼4 km), one of the finest grid size available at the moment. However its temporal resolution is 1 month, averaged over several years. The PR sensor makes one or two snapshots of the Earth surface per day (depending on the latitude). Therefore, measurements are infrequent and have to be averaged over a long time span (here: 11 years, 1997–2007) to provide reliable rainfall data [Bookhagen and Burbank, 2006]. This data set has been successfully applied for measuring precipitation patterns in the Andes [Bookhagen and Strecker, 2008] and the Himalayas [Bookhagen and Burbank, 2006, 2010].

2.2. Rain Gauge Data

[13] We compared precipitation estimates from each product with ground measurements derived from the 55 rain gauge stations located in Nepal (54) and on the Tibetan Plateau (1): Figure 2 (see also Text S1). Most of the data are obtained from the Department of Hydrology and Meteorology Nepal DHM (∼30 years data, numbers 1–51, Text S1). Three high elevation stations (3560, 4260 and 5050 m above sea level (asl), numbers 53–55) in the Khumbu–Everest region are kindly provided by the Ev-K2-CNR, Pyramid-SHARE project, while the station for the Tibetan Plateau (number 52) comes from the LocClim FAO database (http://www.fao.org/nr/climpag/pub/en0201_en.asp). We also compared the precipitation data sets with ground information from rain gauge stations in five watersheds located in Pakistan, India, Nepal (2) and Yunnan/China (Figure 2) maintained within the People and Resource Dynamics Project (PARDYP) program, realized by International Centre for Integrated Mountain Development (ICIMOD) between 1997 and 2006. Note that the 51 gauge stations provided by DHM have been used to generate the APHRODITE data set so there
is a dependency problem between APHRODITE and these gauge station data. They might have also served to calibrate any of the other data sets. However, this is not the case for all the precipitation gauge data of the five PARDYP watersheds, as well as the three Pyramid‐SHARE stations (53–55), which provide independent data that have not been used to generate or calibrate any of these data sets. We presume the uncertainties of all precipitation gauge measurements to be \( \leq 10\% \).

2.3. Bulk Validation of Data

To give an overview of the bulk error of each data set at the scale of Nepal, we compared each data set with all rain gauge stations on a monthly scale. For this purpose, we subtracted the monthly accumulated ground measurements of each station from the corresponding monthly sum of each data set and average the difference considering all the stations. This comparison was carried out considering only the data for the years 2003 and 2004. Because some of the 51 DHM stations might have been used to calibrate or generate some of the products (in particular APHRODITE), we also evaluate the bulk error by considering only the stations that have not been used in the calibration or generation of products (the three Pyramid‐SHARE stations and stations within the five PARDYP watersheds). Because gauge data are not always available for the same period, we sampled precipitation and rain gauge data for months of common availability between 1997 and 2006 and calculated then the monthly bulk error of each product.

2.4. Calculation of Basin‐Wide Precipitation in Five Selected Watersheds

We compared the precipitation data sets with ground information from rain gauge stations in the five watersheds maintained within the PARDYP program (Figure 2). The five relatively small watersheds (15–111 km\(^2\), Table 1) have been equipped with several measuring devices to obtain meteorological, hydrological and erosional parameters. In each watershed, data are available for 4 to 12 rainfall stations (Table 1) for 5 to 10 years, providing a very good data set of ground truth information to calibrate remote sensing information [Andermann et al., 2010]. The station elevation distribution is homogeneous for Bhetagad, Jhikhu Khola and Yarsha Khola basin (Figure 3), whereas in Xizhuang basin a large part between 2200 and 3000 m asl is not covered by stations. The higher part of the Hillkot basin (elevation >1800 m asl) is not covered by rain gauge stations (Figure 3). In each catchment, we interpolated (nearest neighbor interpolation technique) the available gauge data to a mean basin‐wide value. The mean basin‐wide value was then extracted from each data set. Since TRMM‐2B31 data do not exist with high temporal resolution (daily nor monthly) it was not included in the comparison of products here.

2.5. Calculation of Precipitation Along Swath Profiles

We compared the precipitation data sets during the years of common availability of all data sets, 2003 and 2004 (Figure 1), along seven swath profiles perpendicular to the Himalayan front (Figure 2). Each swath profile is 60 km wide and 650 km long. Precipitation and elevation along one profile represent the average over the swath width of the profile. Topography information is derived from the Shuttle Radar Topographic Mission (SRTM) Version 4 (A. Javis et al., Hole‐filled seamless SRTM data V4, International Centre for Tropical Agriculture (CIAT), 2008, available from http://srtm.csi.cgiar.org) with a spatial resolution of...
∼90 m. Along the profiles, precipitation totals were compiled annually as well as for monsoon (June to September) and for the nonmonsoon season (October–May).

3. Results

[17] We present the results of the evaluation of precipitation data sets, first the bulk difference between fully independent and semi-independent gauge data and precipitation data sets. Second, we present results within the five PARDY catchments, allowing us to test the temporal quality of data, from annual to daily scale. Last, we present results along the swath profiles perpendicular to the Himalayan chain and examine the spatial variations of data with regard to elevation.

3.1. Bulk Error and Comparison of Products

[18] The annual bulk comparison between products and fully independent gauge data (Pyramid-SHARE and PARDYP), which have not been used to generate nor calibrate any product, including APHRODITE shows that CPC-RFE, TRMM3B42 and GSMaP underestimate rain gauge data up to 400 mm yr\(^{-1}\), while TRMM-2B31 considerably overestimates (∼600 mm/yr) the independent data set (Figure 4a). Precipitation is mainly underestimated during monsoon season (CPC-RFE 52 mm/yr, TRMM3B42 77 mm/yr, GSMaP 130 mm/yr), while APHRODITE (maximal 12 mm/yr) does not significantly differ from gauge data, whatever the season considered. Because of its temporal resolution, TRMM-2B31 cannot be compared to rain gauge

![Figure 4](image-url)
Results from the bulk error analysis are also presented within the five PARDYP watersheds. In contrast to observations made by Bookhagen and Burbank [2010] for the Himalayas, the hydrologic underestimation of 0.98 (APHRODITE) and of 0.82 (TRMM-3B42) underestimates precipitation whereas TRMM-2B31 overestimates gauge data, but to a lesser degree than with independent data (328 mm/yr).

Results from the bulk error analysis are also reflected in the intercomparison of the data sets in map view (Figure 5). The available gridded products show contrasting patterns of annual precipitation along the Himalayas. All products show high precipitation rates along the mountain chain, but with different amplitudes and patterns. Most products show the westward decrease of precipitation already described by Bookhagen and Burbank [2006] as well as the two along-strike rainfall peaks they also described. They are clearly expressed by TRMM-2B31 because of the higher spatial resolution of this product. The comparison of the remote products with respect to APHRODITE shows that all products differ considerably in detail (Figures 5f, 5g, 5h, and 5i). It illustrates an overestimation of precipitation by TRMM-2B31 with respect to APHRODITE, possibly because APHRODITE cannot depict small-scale changes because of its moderate resolution (0.25°), but also because peak precipitation rates by TRMM-2B31 are likely overestimated, as the authors acknowledge themselves [Bookhagen and Burbank, 2010]. When compared to APHRODITE, remotely sensed products CPC-RFE, GSMaP and TRMM3B42 significantly underestimate precipitation in the eastern and central part of the Himalayas (Figures 5g, 5h, and 5i), while CPC-RFE (Figure 5g) tends to underestimate precipitation in the Western part. All data sets show similar low precipitation rates (<0.5 m/yr) on the Tibetan Plateau and moderate ones (0.5–1 m/yr) in the Indian foreland.

3.2. Comparison of Data Within the Five PARDYP Watersheds

Within the five PARDYP watersheds, precipitation estimates by APHRODITE and TRMM-3B42 fit measurements derived from ground gauge stations, both at the monthly (Figures 6 and 7) and annual scales (Text S2). The correlation coefficient between monthly precipitations derived from gauge data and data sets is 0.87 for APHRODITE and 0.69 for TRMM-3B42 when one considers the five catchments all together (Figure 7). The best correlation is found in the Jhikhu Khola catchment with r² of 0.98 (APHRODITE) and of 0.82 (TRMM-3B42). APHRODITE always fit very well the monthly precipitation derived from gauge data (correlation coefficient between 0.83 and 0.98) except in the Hillkot catchment (Figure 6) where it gives higher estimates than the interpolated gauge data during monsoon season. This is likely the consequence of the lack of gauge stations at high elevations in this basin (Figure 3). Indeed, Bhatt and Nakamura [2005] and Barros et al. [2004] report strong ridge-valley gradients on a basin scale in the Himalayan front. If we assume an orographic gradient, with lower precipitation in the valley bottom than close to the ridges, the absence of stations at high elevation in the Hillkot catchment will result in an underestimation of mean basin-wide precipitation. Note that APHRODITE also correlates very well with precipitation derived from gauge stations at the daily scale in the Jhikhu Khola catchment (Figure 8). This correlation is not observed with the other data sets nor in the other basins. Monthly precipitation derived from TRMM-3B42 usually correlates well with gauge data (correlation coefficient between 0.78 and 0.84; Text S2) except again in the Hillkot catchment, likely for the same reason as discussed above (correlation coefficient of 0.52). Overall, CPC-RFE and GSMaP data do not match the ground information at the annual and monthly scale (Figure 6).

In contrast to observations made by Anders et al. [2006] and Kamal-Heikman et al. [2007], using remote precipitation measurements (TRMM/PR and TRMM-3b42), our annual precipitation estimates from APHRODITE or from interpolated rain gauge data exceed annual water discharge recorded at the catchment outlet of the five PARDYP watersheds. In Jhikhu Khola catchment for example, annual precipitation measured by APHRODITE, TRMM-3B42 and by gauge stations (~1400 mm/yr) is roughly 3.5 times as high as the annual specific discharge (~400 mm/yr) recorded at the basin outlet [Merz, 2004]. As pointed out by Bookhagen and Burbank [2010] for the Himalayas, the hydrologic budget is only correct when evapotranspiration and snow and glacier melt processes are taken into account. High snowfall contribution on the Tibetan Plateau is difficult to detect by remote sensors and is shown to lead to considerable underestimation of basin-wide water budgets [Kamal-Heikman et al., 2007].
Figure 5. Mean annual precipitation distribution of the five tested precipitation data sets for their common availability (2003 and 2004, TRMM-2B31 1997–2007): (a) APHRODITE, (b) TRMM-2B31 [Bookhagen and Burbank, 2006], (c) CPC-RFE, (d) GSMaP, and (e) TRMM-3B42. Figures 5f–5i illustrate the differences between the data sets in respect to APHRODITE (APHRODITE-data set): (f) APHRODITE versus TRMM-2B31, (g) APHRODITE versus CPC-RFE2.0, (h) APHRODITE versus GSMaP, and (i) APHRODITE versus TRMM-3B42.
Figure 6. Monthly mean basin-wide precipitation rates from gridded precipitation data and basin-wide interpolated rain gauge stations for the five watersheds. Gray shading represents the range of interpolated gauge data. The upper and lower limits of the range represent the minimal and maximal monthly sum of precipitation rates.
3.3. Comparison of Data Along the Seven Swath Profiles

Evaluating the different data sets along swath profiles has the advantage to investigate precipitation distribution as a function of elevation. Swath profiles show mean precipitation over the swath width and average over local variations. Therefore, gridded data sets with the same resolution (e.g., TRMM-3B42 and APHRODITE) should match quantitatively. This is however not the case in our results (Figure 9).

While the middle hills and the foreland are easily accessible, the High Himalayas are only sparsely covered with stations and the station elevation might not reflect the surroundings [Bhatt and Nakamura, 2005]. However, most existing rain gauge data indicate high precipitation rates in the Lesser Himalaya and a decrease at higher elevation, in the Higher Himalayas and on the Tibetan Plateau. Along most profiles (Figure 9 and Text S3), gauge data consequently document the orographic effect across the Himalayas, despite possible problems of point data vs. spatial data. Due to inaccessibility, gauge stations are generally situated in valleys, especially within the high Himalayan range (Text S3). Note that most stations used here are situated at mean swath elevation in the transition between the Indian Lowlands and the mountain front (e.g., stations 16, 17, 18, 19, and 20 along profile 3: Figure 9a and Text S3), whereas stations in the high mountain front are situated at minimum elevation of the swath profile (e.g., stations 12, 13, 14, and 15 along profile 3: Figure 9a and Text S3). Therefore, along profile 3 it is not clear if the decreasing trend defined by stations 18 and 14 is due to elevation or stations positioning. In contrast, thanks to the three Pyramid-SHARE stations (numbers 53–55), nearly all stations in profile 5 (Figure 9d) are situated close to mean elevation (Text S3). Here, the strong decrease in precipitation rates between stations 31 and 55, above 3000 m asl is likely the consequence of the orographic effects. Locally, the annual difference between two neighboring stations is remarkable, e.g., between stations 46 and 48 (profile 7, Text S3) or between stations 38 and 44 (profile 6, Figure 9g). In both cases the stations are almost at the same latitude but at different elevation, so they likely record different local annual precipitation variations linked to orography.

For the data sets evaluated here, all seven swath profiles (Figure 9 and Text S3) illustrate the orographic effect of the Himalayan chain, on annual scale as well as during monsoon season (May–October). Depending on the data set, the amplitude of the orographically induced rainfall peak is more or less pronounced. The orographic influence during nonmonsoon season is much weaker, as already observed by Bookhagen and Burbank [2010]. Overall, all data sets are more or less consistent during nonmonsoon season.

Figure 7. Correlation between monthly basin-wide precipitation rates from gridded data and basin-wide interpolation from station data. Data from all five watersheds is plotted here.

Figure 8. Daily correlation (Jhikhu Khola, Nepal) between APHRODITE data and basin-wide mean interpolated precipitation rates.
Figure 9. Precipitation swath profiles (a, b, and c) 3, (d, e, and f) 5, and (g, h, and i) 6 (further profiles are given in Text S3). Mean annual precipitation is given in Figures 9a, 9d, and 9g for each precipitation data set and swath profile, plotted against elevation from north to south. Shaded curves represent maximum, minimum, and mean values along each swath profile. The gridded data are for the years of common availability (2003 and 2004), while TRMM-2B31 [Bookhagen and Burbank, 2006] is for a 10 year time span (1997–2007). Gauge data are plotted as mean values for both time spans. Error bars show the long-term (~30 years) maxima and minima. Figures 9b, 9e, and 9h represent the monsoon season (June–September) only. Figures 9c, 9f, and 9i represent the nonmonsoon season (October–May).
Along the Himalayan front, annual precipitation estimates at a given location along the profiles are always maximum for TRMM-2B31 data and minimum for GSMaP, the difference being often as large as two- to threefold (Figures 9a, 9d, and 9g). Between these two extreme data sets, APHRODITE usually shows higher estimates than TRMM-3B42 and CPC-RFE data, except for profiles 1 and 2 (Text S3) where the latter delivers the highest values. The three Pyramid-SHARE project stations (independent to APHRODITE) fit very well with the APHRODITE data, on an annual as well as a seasonal scale (Figures 9d–9f). A striking feature is the large difference between TRMM-2B31 data set and all the other data sets when considering mean and maximum values, as well as the high frequency spatial variations of precipitations documented. This is a direct consequence of the high spatial resolution of this data set. Overall, most of the gauge stations usually plot near the mean value of the TRMM-2B31 product, except for some local examples (e.g., stations 27 and 29 in profiles 4; stations 12 and 13 in profile 3; stations 31 and 34 in profile 5; stations 38, 39 and 41 in profile 6; see Figure 9 and Text S3). The extreme values of the TRMM-2B31 product always exceed gauge values (e.g., extremely high precipitation values (>7000 mm/yr) for few pixels on profile 3, Figure 9a), as already noticed by Bookhagen and Burbank [2010] and Anders et al. [2006]. These extreme values likely overestimate real precipitation rates. Note that Bookhagen and Burbank [2006] used a network of rainfall stations to calibrate their TRMM-2B31 data; however, most of the stations we used here have not been used in their work. In general the swath profiles reflect the findings of the bulk error estimation (Figure 4).

Bookhagen and Burbank [2010, 2006] demonstrated that precipitation profiles across the Himalayas mimic the topography. When the topography steadily increases from the Indian Lowland to the Tibetan Plateau, rainfall distribution is characterized by a single high peak of rainfall at elevation ~0.9 km. Conversely, for a two-stepped increase of topography, rainfall distribution shows two peaks of lower amplitude at ~0.9 km and 2.1 km (e.g., Figures 5b and 9g). This bimodal distribution of rainfall is visible in most of the TRMM-2B31 profiles we show here (e.g., profiles 1, 4, and 6). This particular distribution is also well depicted by the APHRODITE product in most cases (e.g., profiles 1, 4, and 6) but with a lower amplitude. However, this is usually not depicted by the other products, except in profile 1 (Text S3). Overall, all the products show the increase of precipitation rates at the front of the Himalayas. TRMM-3B42, GSMaP and CPC-RFE, however, do not describe correctly the precipitation distribution at elevations higher than 1 km. This result highlights the difficulty of remote sensing techniques to capture precipitation in areas with strong orographic effects. The direct comparison of TRMM-2B31 and TRMM-3B42 in the Andes already shows that TRMM-3B42 cannot detect the local orographic precipitation maxima, due to its moderate spatial resolution [Bookhagen and Strecker, 2008].

4. Discussion and Conclusion

We show that existing gridded precipitation data sets as well as published sources [Bookhagen and Burbank, 2006] display large differences along the Himalayan orographic front. With the exception of CPC-RFE, all measurements correspond in low-relief landscapes (Indian Lowlands, Tibetan Plateau) and during nonmonsoon season. On the basis of comparison with independent ground observations (Figure 4a) we show that most remote products underestimate precipitation during monsoon season at the annual and monthly scale whereas TRMM-2B31 [Bookhagen and Burbank, 2006] overestimates precipitation at the annual scale. These problems of precipitation estimation are likely due to remote techniques and calibration procedures. They do not concern APHRODITE data, a product processed from gauge stations, which gives the best precipitation estimates when compared to independent ground observations (Figure 4a). However, the lack of stations at high elevation limits the accuracy of this data set.

Most of the rain gauges used to calibrate CPC-RFE, GSMaP and TRMM-3B42 data are derived from the GTS network with reportedly poor spatial coverage in the Himalayas [Yatagai and Kawamoto, 2008], which might explain the underestimation of precipitation during monsoon season. Additionally, Yatagai and Kawamoto [2008] report that the GTS database includes erroneous entries in the Himalayan region, where 0 mm precipitation values were reported instead of missing values, thus resulting in underestimating precipitation. In the case of APHRODITE, up to 4.5 times as many stations as GTS existed for the interpolation space [Yatagai and Kawamoto, 2008]. Many difficulties in the estimation of precipitation from space may also arise from remote techniques themselves. Remote sensors cannot determine accurately snowfall which is the major
Our study also highlights some of the difficulties in evaluating remote precipitation products using rainfall gauge data. In the Himalayas, precipitation varies according to a wide range of spatial scales, from small-scale ridge-valley gradients [Barros et al., 2004; Bhatt and Nakamura, 2005; Anders et al., 2006; Craddock et al., 2007] to large-scale orographic effects over the whole mountain [Anders et al., 2006; Bookhagen and Burbank, 2006]. Ideally, the validation of any remote sensed product from gauge stations is only possible if the resolution of the products is sufficient to take into account the scale of spatial variability of precipitation. Hence, the coarse resolution of all products (Figure 1) introduce an inescapable problem in the validation procedure. For example, gauge stations in valleys within the High Himalayas (e.g., profile 3; see Text S3) likely introduce a bias in the reference value of precipitation they provide. We show that several stations, covering full elevation range, are necessary to represent the climatic situation correctly and to validate remotely sensed products. We also show that basin-wide precipitation measurement deduced from gauge data are significantly altered if elevation is not sufficiently covered by gauge stations (Hillkot watershed, Figure 6). The current gauge network within the High Himalayas is generally not sufficient to characterize orographic precipitation phenomena correctly.

Along the Himalayan range, several precipitation products are of potential interest depending on the problem addressed. TRMM-2B31 is a good product when one wants to investigate rainfall patterns. This is of significant interest for example to understand the topographic influence on rain distribution. The use of this product is however, limited by its temporal resolution, which is not adequate to investigate, for example, event-scale processes. As also observed by Bookhagen and Burbank [2006], the rainfall peaks of TRMM-2B31 data along the Himalayan front are generally overestimated. TRMM-2B31 describes correctly the large-scale orographic rainfall distribution along the Himalayan front. This distribution is poorly depicted by all other data sets based on remote sensing techniques. As in the work by Tian and Peters-Lidard [2010], our study shows that sensor algorithms for mountainous regions, where terrain changes on short distances and orography influences precipitation, must be improved.

As observed in the five small watersheds studied here, APHRODITE (and to a lesser extent TRMM-3B42) deliver good temporal variability, both on annual and monthly scale. In some cases, for example in the Jhikhu Khola catchment, which is located in the lower middle mountains with a low relief (~1400 m), even daily precipitation estimates by APHRODITE are representative (Figure 8). Because of its long availability (>30 years), coupled with its good temporal resolution, the APHRODITE product is appropriate to track above-threshold events driven by precipitation (e.g., landslides thresholds [Gabet et al., 2004; Dahal and Hasegawa, 2008]) as well as for hydrological budget and discharge analysis. It can be applied for hydrological budget and discharge analysis of large basins (>2000 km² ~ 2 pixels). If higher temporal resolution than APHRODITE is needed, then the TRMM-3B42 data, with its 3-hourly resolution, could also be exploited for relative analysis, even if not tested here. Accurate precipitation data in an active mountain belt such as the Himalayas are essential for a real understanding of the potential couplings between climate, erosion and tectonics processes as well as for hazard mitigation.

Acknowledgments

We would like to thank the German academic exchange service DAAD and the French German university DFH for funding the Ph.D. thesis of C. Andermann. Pyramid Ev-K2-CNR–SHARE project for meteorological data. We would like to thank Madhav Dhakal and Isabelle Provodilli from ICIMOD Kathmandu for their outstanding support, Jianchi Xu and Ma Xing for readily sharing information on Xizhuan watershed, as well as B. P. Kothiyari from GBPHEP India for providing data on Bhetagad watershed. Many thanks go also to Bodo Bookhagen for providing us with a mean annual data set of TRMM-2B31. We would like to thank Jagadish Karmacharya from DHM Nepal for his great support to select and obtain gauge data. We thank Alain Crave for helpful discussions and Guillaume Dupont-Nivet for proofreading. We would also like to thank two anonymous reviewers and the Editor of G-cubed, Louis Derry, whose comments and suggestions have considerably improved this manuscript.
References


