Journal of Hydrology: Regional Studies 13 (2017) 138–151

Contents lists available at ScienceDirect

Journal of Hydrology: Regional Studies

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

Evaluating the accuracy of Climate Hazard Group (CHG) satellite rainfall estimates for precipitation based drought monitoring in Koshi basin, Nepal

Narayan K. Shrestha[a,*], Faisal M. Qamer[a,*], Diego Pedreros[b], M.S.R. Murthy[a], Shahriar Md. Wahid[a], Mandira Shrestha[a]

[a] International Centre for Integrated Mountain Development (ICIMOD), Nepal
[b] Department of Geography, University of California, Santa Barbara, CA, USA

ARTICLE INFO

Keywords:
Drought
CHG
CHIRP
CHIRPS
Koshi basin
SPI

ABSTRACT

Study region: Koshi basin, Nepal.

Study focus: While rainfall estimates based on satellite measurements are becoming a very attractive option, they are characterized by non-negligible biases. As such, we assessed the accuracy of two satellite products of the Climate Hazard Group (CHG) – (a) a satellite-only Climate Hazards Group InfraRed Precipitation (CHIRP) product, and (b) a CHIRP blended with ground-based station data (CHIRPS) – at a monthly time scale from 1981 to 2010 in the Koshi basin of Nepal using ground-based measurements. A separate analysis was also made for the data set after 1992, as the number of stations used in the blending has significantly reduced since 1992. Next, both CHG data sets were used to calculate one of the most popularly-used precipitation-based drought indicators – the Standardized Precipitation Index (SPI).

New hydrological insights for the study region: The accuracy of the CHG data set was found to be better in low-lying regions, while it was worse in higher-elevation regions. While the CHIRPS data set was better for the whole period, the CHIRP data set was found to be better for the period after 1992. Physiographic region-wise bias correction has improved the accuracy of the CHG products significantly, especially in higher-elevation regions. In terms of SPI values, the two CHG data sets indicated different drought severity when considering the whole period. However, the SPI values, and hence the drought severity were comparable when using the data from after 1992.

1. Introduction

Potential inevitable impacts of climate change and the ever-rising demand of water due to growing industrialization and urbanization have increased drought incidence, frequency, and severity in recent times (Mishra and Singh, 2010). Water resources planners and managers thus need to understand historical drought events and their severity. Based on plausible drought outlook scenarios, they need to identify and develop several measures in order to lessen and mitigate any future adverse effects that drought spells could have on the economy, the environment, and society.

While literature points to several drought-related definitions (Mishra and Singh, 2010; Palmer, 1965; Wilhite, 1992; WMO, 1986), the most common definitions reference a deficiency in rainfall (or precipitation). Hence, it is evident that drought monitoring and

* Corresponding authors.
E-mail addresses: shrestha.narayan@hotmail.com (N.K. Shrestha), Faisal.Qamer@icimod.org (F.M. Qamer).

http://dx.doi.org/10.1016/j.ejrh.2017.08.004
Received 8 February 2017; Received in revised form 14 August 2017; Accepted 16 August 2017
Available online 01 September 2017

early warnings need robust rainfall estimates over space and time (Funk et al., 2014). Such data are hard to obtain using only ground measurements, as it is difficult to install and maintain a network of rain gauge stations providing real-time high-resolution rainfall data. Hence, rainfall estimates based on satellite measurements are becoming a very attractive option, as they provide continuous spatial measurements of the rainfall field; there are several open-source satellite products, too (Hessels, 2015). There exist several satellite-based precipitation products with varying resolutions (Dinku, 2014; Shrestha, 2011). However, these products are often characterized by a 'non-negligible bias' (Xie and Arkin, 1995) and therefore need thorough assessments before use. One such relatively high-resolution satellite product is maintained and distributed by the Climate Hazards Group (CHG). The CHG provides data archives of a quasi-global, gridded 0.05° resolution, precipitation time series from 1981 to the near-present. As such, two CHG products exist: (a) the Climate Hazards Group InfraRed Precipitation (CHIRP) and (b) the CHIRP blended with ground-based station data (CHIRPS). The former data satellite-only products (CHIRP) are available almost immediately (after the end of a pentad: 2nd, 7th, 12th, 17th, 22nd and 27th). The final CHIRPS product is available after some delay − only after the 15th of the following month (Funk et al., 2014). Out of the two CHG products, CHIRPS can be perceived to be more robust in representing the rainfall field, as it is blended with ground measurements.

There have been some efforts in evaluating the accuracy of the CHG products. For instance, Dinku (2014) compared CHIRP and CHIRPS data against other satellite-based products and against ground-based observations, and it was found that the station-only product (CHIRP) over-performed in Ethiopia and Tanzania. The result is contrary to expectations that the station-blended product (CHIRPS), which merges with ground measurements, would match better with the in-situ observations. Also, it has been established that the merging of precipitation observations from several sources (e.g. satellite and in-situ observations) improves the overall quality of the satellite product (Xie and Arkin, 1995). Katsanos et al. (2016) performed validation tests of CHIRPS products all over Cyprus and found good matches with ground-based rain gauge data. Similarly, validation of CHIRPS over Colombia by Pedreros et al. (2014) found higher correlation with ground-based data, especially in drier months. Hessels (2015) carried out cross-comparisons of several open-source satellite products for the Nile basin and found CHIRPS to be one of the best products currently available for hydro-meteorological applications.

This study is carried out with the main objective of assessing the accuracy of two CHG products – the satellite-only product (CHIRP) and station-blended product (CHIRPS) – in representing the rainfall field over one of the major tributary rivers of the Ganges (the Koshi River) within Nepal in view of using them for drought monitoring, assessment, and forecasting of the river basin. Due to Nepal's rugged topography, these products are assessed on the basis of the existing physiographic region (divided based on the elevation range) of Nepal. The effects of bias correction on the raw satellite products is also assessed. Moreover, the CHG products are used to calculate a popularly-used precipitation-based drought index – the Standardized Precipitation Index (SPI). The results of the study will provide bases for decision makers on the usability of these data products for hydro-meteorological studies in general and for drought monitoring in particular.

2. Materials and methods

2.1. The study area

The Koshi River is a major transboundary river and a major sub-basin of the Ganges. The river drains from the northern slopes of the Himalayas in Tibet to the southern slopes of Nepal and a small part of India. The total catchment area is 87,460 km^2 above Kursera, upstream of the confluence with the Ganges, with the highest percentage (45%) lying in Nepal (see Fig. 1). It covers the entire north-south segment of eastern Nepal (Fig. 1, inset map). Based on topography, Nepal is divided into five physiographic regions. The Terai region, in the south, is a low-lying (60300 m; see Table 1) plain area. The Siwalik region is a narrow foothill belt with an elevation of less than 1,000 m, while the Middle Mountain region is the widest strip with an elevation of up to 3,000 m. The High Mountain region generally has steep slopes and deep-cut valleys with an elevation of up to 4,000 m, and the Himal region is in the north and is generally above the snow line (Shrestha et al., 2011).

2.2. Data sources

2.2.1. Ground-based rainfall

Data from a rain gauge network consisting of forty-nine rain gauges which are lying inside the Koshi basin (Nepal part; see Fig. 2) are the source of ground-based rainfall. These stations are being operated and maintained by the Department of Hydrology and Meteorology (DHM), Nepal. Daily precipitation data from 1980 to 2010 have been made available. The rain gauge network density is fairly good in the Middle Mountain, Terai, and High Mountain regions and poor in the Siwalik region. In the Himal region, not a single rain gauge station exists (refer to Table 1 and Fig. 2). These ground-based data were checked for consistency using time series plots and double-mass curve techniques. Some stations have a prolonged period of missing data (refer to Fig. 2). Incomplete, dubious data found during the quality check were discarded from further analyses. As such, one station (named Barhabise, lying in the Middle Mountain region of the northwest part of the Koshi basin) was discarded, as there was significant deviation in the double mass curve which reduced the total number of stations used in the comparison to forty-eight.

2.2.2. Satellite-based rainfall product

Two CHG products (CHIRP and CHIRPS) are the source of satellite-based rainfall data. For CHIRP, the CHG encompasses information from diverse sources. As a first step, IR-derived precipitation estimates from three sources − namely, (a) the Climate

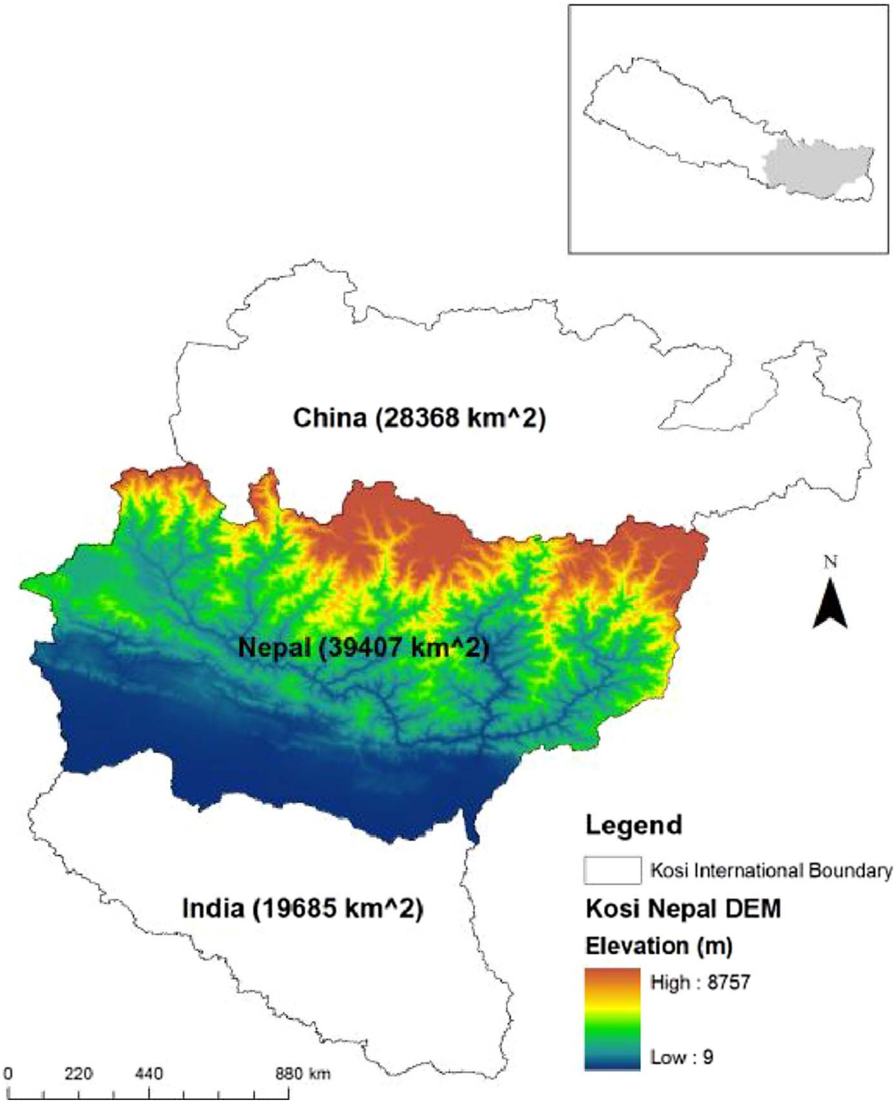


Fig. 1. The Koshi river basin.

Table 1
Raingauge density in the Koshi basin and their sufficiency as per World Meteorological Organization's recommendation.

SN	Physiographic regions	Elevation range[a] (m)	Number of stations	Raingauge density[b]	WMO recommendation[c]
1	Himal	>4000	0	–	100–250 (Ideal Condition) 250–1000 (Difficult Condition)
2	High Mountain	2000–4000	8	850	
3	Middle Mountain	1000–3000	29	500	
4	Siwalik	300–1000	4	1650	600–900 (Ideal Condition) 900–3000 (Difficult Condition)
5	Terai	60–300	8	700	

[a] Taken from Shrestha (2011).
[b] Area of the physiographic regions divided by number of raingauges.
[c] WMO (1994) has recommended raingauge densities as per field conditions (ideal or difficult) for different climate zones.

Prediction Centre (CPC; spatial resolution: 4 km × 4 km; 2000 to present; temporal resolution: 0.5 h), (b) the National Climate Data Center (NCDC; spatial resolution: 8 km × 8 km; 1981–2008; temporal resolution: 3 h), and (c) the Tropical Rainfall Measuring Mission (TRMM; spatial resolution: 0.25° × 0.25°; temporal resolution: 3 h) are expressed as percentages (spatial resolution: 0.05° × 0.05°; temporal resolution: 5 days) as per their long-term climatology. In the next step, they are multiplied by their corresponding long-term climatology – the Climate Hazards Group Precipitation Climatology (CHPclim; time span: 1981–2012) to generate CHIRP (Funk et al., 2015; Funk et al., 2014; Khandu et al., 2016).

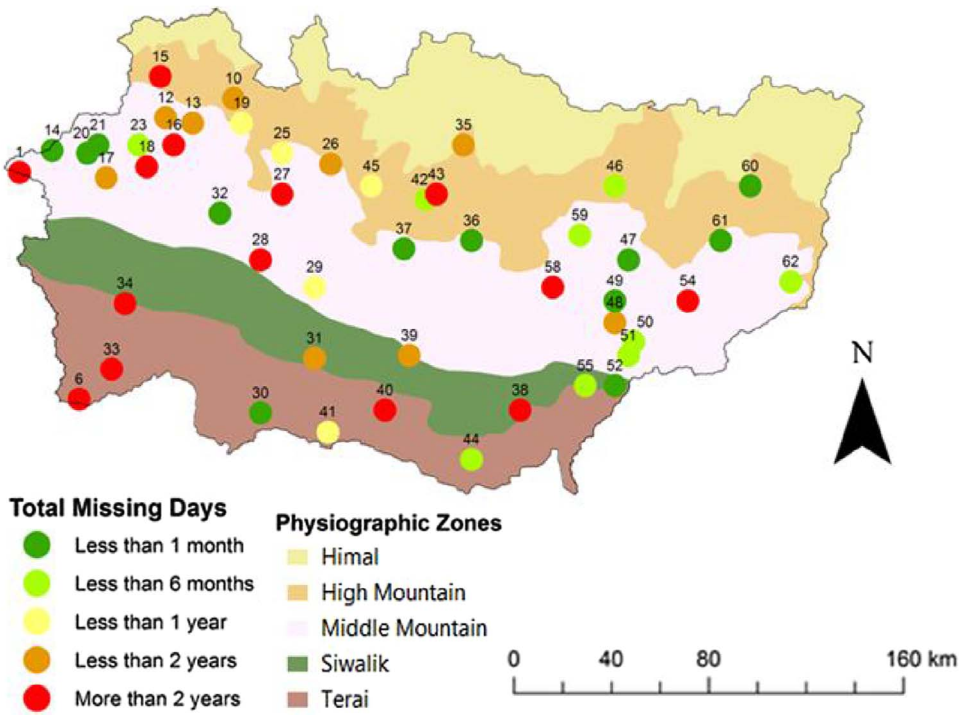


Fig. 2. Koshi basin with available raingauge stations. Also shown are the physiographic divisions of Nepal.

For CHIRPS, CHIRP is blended with station data using modified Inverse Distance Weighing (IDW). As such, two CHIRPS products are available. The preliminary version uses only the daily Global Telecommunication System (GTS) station data; those stations are provided by individual countries for blending and are available on the second day after the end of a pentad. The final CHIRPS uses station data from two more sources: the daily and monthly Global Historical Climate Network (GHCN) and the Global Summary of the Day data set (GSOD). The final CHIRPS data is available only after the 15th of the following month (Funk et al., 2014). In this study, we used both monthly CHIRP and CHIRPS data sets.

In Nepal, the CHG used sixty-five stations before the merging process which drastically reduced the number of stations to less than five after 1992 (see Fig. 3). The reasons for this drastic drop are unknown. Some plausible reasons could be related to problems with data acquisition, quality of data, and so on.

2.3. Gauge- and satellite-based rainfall comparison

The overlapping period of 1981–2010 was used for the comparison of ground-based and satellite-based rainfall products. In general, two approaches have been widely reported while comparing satellite- and ground-based rainfall products. Some researchers (e.g., Shrestha, 2011) used rain gauge networks to estimate spatial rainfall using spatial interpolation methods such as Kriging, IDW, etc., then carried out pixel-to-pixel comparisons. Other researchers (e.g., Shrestha et al., 2013a) extracted rainfall estimates over the

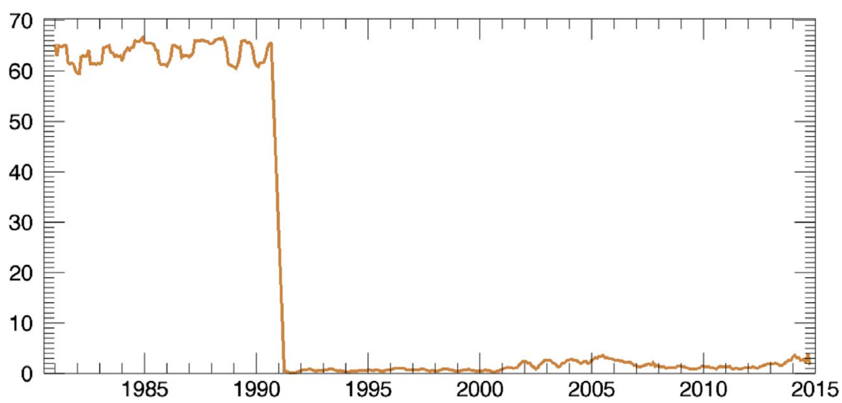


Fig. 3. Evolution of the number of raingauge stations used in the blending process to produce the CHIRPS data from CHIRP data for Nepal.

pixel in which the rain gauge is located and then carried out the comparison. In the second approach, some researchers (e.g., Lack and Fox, 2007) used arithmetic average of four to nine surrounding pixels over the rain gauge stations. In this study, the second approach is adopted because spatial interpolation based on rain gauge data using such sparsely located rain gauges in terrain (like that in Nepal) could be rather error-prone. Also, the fine resolution (0.05° × 0.05°) of the satellite product would mean that the ground-based observation is representative of the areal average precipitation total within the pixel in which the gauge is located.

This study is divided into two parts. In the first part, the accuracy of the CHG products (CHIRP and CHIRPS) is assessed against the ground-based rainfall data recorded at a rain gauge network consisting of forty-eight stations (see details on §2.2.1). As Nepal's topography is rather rugged, and there exists a huge variation in elevation (60 m to 8848 m) within the narrow north-south strip of about 160 km (Shrestha et al., 2011), separate analyses in five commonly established physiographic regions (Terai, Siwalik, Middle Mountain, High Mountain and Himal) have been carried out (refer to Fig. 2). Because the Siwalik region contains only four rain gauge stations (Table 1), an accuracy assessment based on just this small number of stations would be rather misleading. Therefore, the Siwalik region has been merged with the Terai region for further analyses.

For comparison, valid pairs are identified when both ground-based and satellite-based daily rainfall depths are higher than 1.0 mm. Such an approach has been widely used for comparison purposes (Shrestha, 2011; Shrestha et al., 2013a; Shrestha et al., 2013b). Next, we estimate the physiographic region-wise Mean Field Bias (MFB). The MFB is calculated as the arithmetic average of the ratios of ground- and satellite-based valid pairs (Eq. (1)). The MFB is sometimes also referred to as a multiplicative bias and is one of the most widely-used methods of satellite-gauge merging techniques (Shrestha et al., 2013a). As such, we group all the valid pairs of a particular physiographic region and then field bias (FB_i) is calculated and averaged to get the MFB of that physiographic region. The satellite-based rainfall data is then corrected using the physiographic region-wise MFB to assess the improvements after bias correction. In order to eliminate the effect of a particular extreme bias value on the MFB, outliers of the individual bias have been defined as values which either exceed the third quartile plus 1.5*inter-quartile range or fall below first quartile minus 1.5*inter-quartile range.

$$MFB = \frac{\sum FB_i}{N} = \frac{1}{N}\sum \frac{R_{gauge,i}}{R_{satellite,i}} \tag{1}$$

Where, MFB = Mean field bias; FB_i is the field bias of i^{th} valid pair; N is the number of valid pairs; $R_{satellite,i}$ is the rainfall estimates from the CHG product of the valid pair i; and $R_{gauge,i}$ is the rainfall recorded in a particular rain gauge for the valid pair i.

2.4. Goodness of fit statistics

To evaluate the accuracy of the satellite-based products when compared against the ground-based rainfall depths, two goodness of fit statistics are used, namely the percentage bias – PBIAS (Eq. (1)) − and the ratio of the root mean squared error to the standard deviation of the observations – RSR (Eq. (2)). Moreover, four performance ratings ('Very Good,' 'Good,' 'Satisfactory,' and 'Unsatisfactory') are used. The qualitative ratings are based on ranges of values of the PBIAS and RSR as proposed by Moriasi et al. (2007). It should, however, be noted that the ranges of goodness of fit statistics values for a particular qualitative rating, as presented in Table 2, are originally designated for streamflow and nutrients and for a monthly time steps. Here, we adopt the same range for the rainfall depths. Further, Moriasi et al. (2007) also suggested the use of the widely-used Nash-Sutcliffe Efficiency (NSE) as another goodness of fit statistic. As the NSE and RSR are related, we opted to use only the RSR in this study.

$$PBIAS = \frac{\sum (R_{gauge} - R_{satellite})}{\sum (R_{gauge})} x100 \tag{2}$$

$$RSR = \frac{\sqrt{\sum (R_{gauge} - R_{satellite})^2}}{\sqrt{\sum (R_{gauge} - \overline{R_{gauge}})^2}} \tag{3}$$

Where, $R_{satellite}$ is the rainfall estimates from the CHG product; R_{gauge} is the rainfall recorded in a particular rain gauge and $\overline{R_{gauge}}$ is the average gauge reading over the considered time span.

Table 2
Range of adopted values of the Percentage of Bias (PBIAS) and Root Mean Squared Error normalized by Standard Deviation (RSR) for a particular qualitative rating.

Performance rating	PBIAS (%)	RSR (−)
Very Good	< ± 15	0–0.5
Good	±15–±30	0.5– 0.6
Satisfactory	±30–±55	0.6–0.7
Unsatisfactory	±55	>0.7

2.5. Standardized precipitation index (SPI)

There exist several drought indices based on different meteorological and hydrological variables. The Standardized Precipitation Index (SPI) is one of the most popular indices based on long-time series of precipitation. Drought monitoring and forecasting based on SPI values is prevalent (e.g., Bayissa et al., 2015; Dahal et al., 2015; Narendra, 2008; Shah and Mishra, 2016; Shah and Mishra, 2015; Tokarczyk and Szalińska, 2014; Wong et al., 2013). While calculating the SPI, the precipitation time series is fitted to a probability distribution (normally the gamma distribution; see Eq. (3)) which is further transformed into the normal distribution so as to get a mean SPI value of zero (McKee et al., 1993). As the SPI needs only one input (precipitation) and can be calculated using various time scales (precipitation summed to time scale of interest), it is being adopted as reference index, as suggested by the WMO (2012).

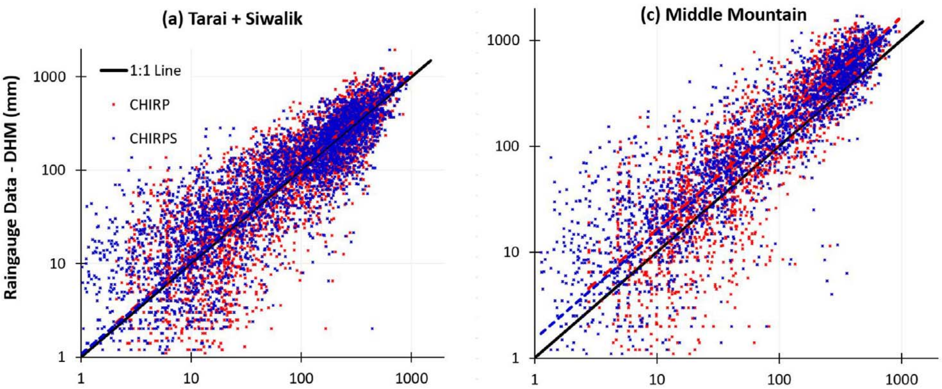
In principal, SPI can be calculated at any time scale depending on the time scale of input and as per its intended use. A study by Szalai et al. (2000) found a time scale of two months to be the most appropriate for hydrological drought, and a time scale of three months for soil moisture related drought. A good time scale for SPI is often subjective, but Mishra and Singh (2010) recommended not to adopt time scales of less than one month (too many zeros at finer time scales may result in skewed distributions) or of more than twenty-four months (fitting the gamma distribution – and, consequently, the parameter estimation – would be error-prone due to a lesser number of data points).

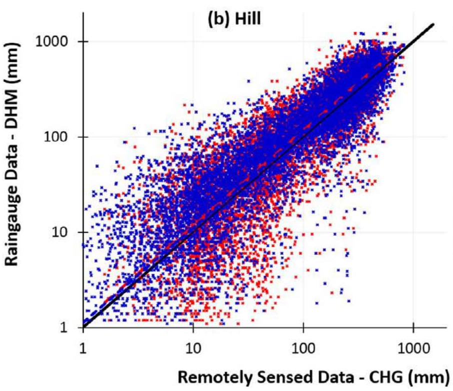
For the second part of the study, the satellite-based CHG products are assessed in representing the SPI (WMO, 2012). There was a need to calculate spatial and temporal variation of the SPI in order to identify the historical droughts' incident and severity patterns. Using the CHG data set (5 km × 5 km) for the Koshi basin would require an SPI calculation for a total of 1416 pixels, which would be practically impossible. Hence, the analysis of SPI values calculated based on the CHG products for the entire Koshi basin required an automated tool which could iterate through all the pixels and calculate the SPI values. For this, we used a freely available executable – the "spi_sl_6.exe", as developed by NMDC (2016) and developed a .NET-based interface to invoke the executable for all the grid cells.

3. Results and discussion

3.1. Accuracy assessment results

Fig. 4 shows scatter plots of monthly rainfall depths for the whole period (1981–2010) between ground-based (DHM) and satellite-based (CHG) products at different physiographic regions. In the low-lying regions (Terai and Siwalik), the accuracy of the CHG





Statistics						
Region	Number of Valid Pairs		Slope of linear fit		R^2 of linear fit	
	CHIRP	CHIRPS	CHIRP	CHIRPS	CHIRP	CHIRPS
Terai + Siwalik	3461	3446	1.11	1.07	0.66	0.69
Middle Mountain	7892	7859	1.16	1.13	0.68	0.73
Higher Mountain	2307	2304	1.65	1.53	0.73	0.73

Fig. 4. Scatter plot between the satellite (CHG) products and ground-based raingauge (DHM) data for the whole period (1981–2010). Also shown are the linear fit (trendline: y = mx, y: CHG data and x: DHM data). Red colored markers and line are for the CHIRP data set while blue color for the CHIRPS.

products outperformed all other regions, as the slope of the linear fit has been found to be nearly 1. However, the linear fit resulted in the lowest R^2 value in low-lying regions, which indicates that although the average bias between the two sources of rainfall was the lowest, the one-to-one matching between the valid pairs was rather worse. At the Middle Mountain region, the accuracy of CHG products has been found to be fairly comparable to the low-lying regions, but considerably lower scattering has duly been reflected in a relatively higher R^2 value of the linear fit between the two sources of rainfall.

Out of the two CHG products, the station-blended satellite product (CHIRPS) performed slightly better than satellite-only product (CHIRP) across all the physiographic regions, as indicated by rather consistent values of near-optimal slope and higher R^2 values. A systematic underestimation of rainfall depths by the CHG products has been observed, as all linear fit lines were above the 1:1 line (slope higher than 1). The case was most severe in the High Mountain region, also as indicated by the higher value of the slope (1.53 for CHIRPS and 1.65 for CHIRP). Such a serious underestimation of rainfall depths by satellite products in complex terrain, such as that in the High Mountain region, has also been observed in other parts of the world – over Nepal using NOAA CPC_RFE2.0 satellite estimates (Shrestha et al., 2013a), in the Central United States using several satellite products (AghaKouchak et al., 2011), in Bhutan using the station-blended (CHIRPS) product (Khandu et al., 2016), in Mozambique, also using the station-blended (CHIRPS) product (Toté et al., 2015), in a Himalayan basin of Bhutan using the Tropical Rainfall Measuring Mission Multi-Satellite Precipitation Analysis (TMPA) product (Xue et al., 2013), over all of India using other satellite products (Prakash et al., 2014), and in several parts of Northern India using several other satellite products (Shah and Mishra, 2016).

However, other researchers (e.g., Katsanos et al., 2016) have found mixed results – systematic underestimation during early decades (the 1980s) and systematic overestimation over the last decade when using the station-blended product (CHIRPS) in Cyprus. In our case, owing to the fact that the R^2 value at the High Mountain region was rather high, it is fair to state that although the CHG products seriously underestimate the rainfall depths, the one-to-one matching of valid pairs are rather better. It then implies that these products could be substantially improved by applying multiplicative correction factors such as the MFB.

In terms of mean PBIAS value, CHG products best matched the ground measurements in the Middle Mountain region, but higher scattering (as indicated by the length of the error bars in the box plots; see Fig. 5) mean that higher variability exists. Moreover, lower mean PBIAS could be the result of a compensating effect of negative and positive biases amongst the stations clustered in the region. The latter argument can also be complemented by slightly higher RSR values at the region as compared to the low-lying regions (Terai and Siwalik). In line with the behavior as seen in Fig. 4, the higher values of PBIAS and RSR further indicate the rather poor accuracy of both CHG products at the High Mountain region.

The findings of this study are consistent with the observation of Shrestha et al. (2011). The better matching at the low-lying regions (Tarai and Siwalik) was expected, as the elevation variation in these regions is rather low and so orography plays a less important role. As the Middle and High Mountain regions lie in high elevations, orography would play a rather important role. Consequently, CHG products tend to be less accurate in these regions of the Koshi basin.

In terms of the qualitative ratings calculated based on the ranges of PBIAS and RSR values (refer to Table 2), the accuracy of both CHG products is fairly comparable (see Fig. 5). However, the station-blended product (CHIRPS) was found to be slightly better at representing the ground-based rainfall depths as compared to the satellite-only product (CHIRP). This is somehow expected, as the gauge-satellite merging technique, as employed in the blended product, would improve the accuracy of the product. Out of the forty-eight stations taken into consideration, only one station (#65) resulted in the 'Very Good' matching of the satellite-only (CHIRP) data set when compared against the ground measurements, while in case of the blended product (CHIRPS), two more stations resulted in a 'Very Good' accuracy. In the low-lying regions (Terai and Siwalik), the rating was found to be slightly better, as most of the stations had the 'Good' rating. The qualitative ratings in Middle Mountain region varied from 'Very Good' to 'Unsatisfactory'. In the High Mountain region, the matching of gauge and satellite products was found to be the worst, as most of the stations lying in that region resulted in the 'Unsatisfactory' rating.

From the analysis, it can be stated that, out of the two CHG products, the station-blended product (CHIRPS) was found to be slightly better at all the physiographic regions of the Koshi basin and that the accuracy of the both products at the low-lying regions (Terai and Siwalik) was the highest. At the High Mountain region, the performance was found to be the worst. Moreover, the CHG products systematically underestimated the rainfall depths at all the physiographic regions when compared against the ground measurements.

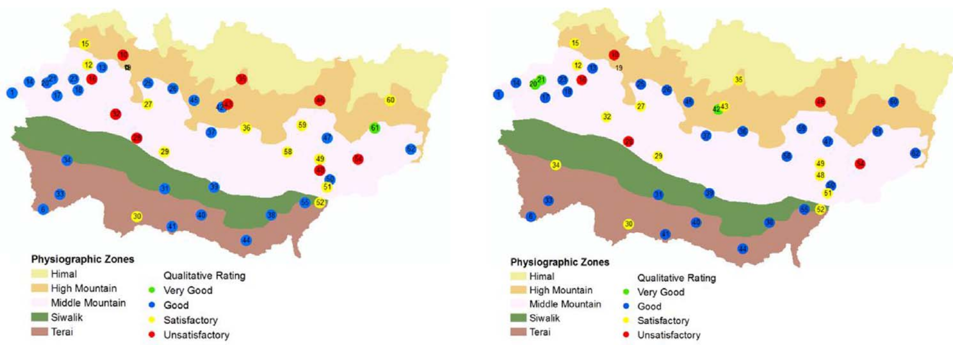


Fig. 5. Qualitative ratings of the CHG products: CHIRP (Left) and CHIRPS (Right) as compared with the ground-based raingauge data.

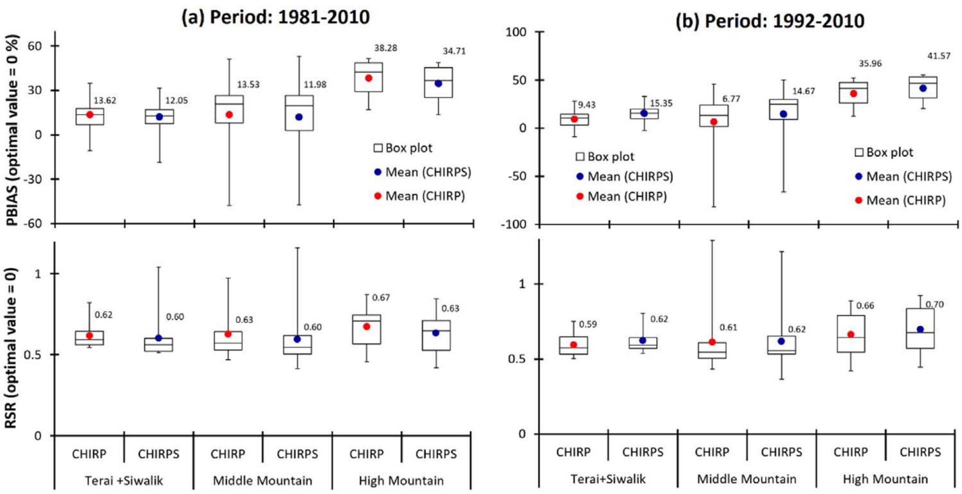


Fig. 6. Box plots of the Percentage of Bias (PBIAS) and Root Mean Squared Error normalized by Standard Deviation (RSR) values calculated for the different physiographic regions using the data set of (a) the whole period (1981–2010) and (b) the period after 1992 (1992–2010).

Since the number of rain gauge stations incorporated in the blending process has drastically reduced from 1992 onwards, a need was therefore felt for a separate accuracy assessment of the CHG data set splitting into two − one considering the whole time span (1981–2010) and other for the time span after 1992 (1992–2010). The box plot (Fig. 6b) of the PBIAS and RSR, calculated for each valid pair while comparing the ground- and satellite-based products for a period after 1992 (1992–2010), indicated quite a reverse scenario from that observed for the whole data set (1981–2010; refer to Fig. 6a). For the data set of the whole period, the blended product (CHIRPS) was found to be better; for the data set after 1992, the satellite-only product (CHIRP) was found to be better as indicated by the lower values of the PBIAS and RSR at all physiographic regions of the Koshi basin. However, a general trend of the systematic underestimation of rainfall depths by both CHG products still exists.

While pair-wise comparison of ground- and satellite-based rainfall products for the whole period (1981–2010) show the blended product (CHIRPS) as being the better product across all physiographic regions, further analysis on a seasonal scale point to a different scenario. Even for the whole period, the station-blended product (CHIRPS) outperformed the satellite-only product (CHIRP) in the summer season (JJAS, Table 3) only. For all other seasons, the satellite-only product (CHIRP) better matched the ground measurement across all the physiographic regions. Moreover, for the data set after 1992 (1992–2010), the satellite-only product (CHIRP) has consistently outperformed the station-blended product (CHIRPS) in all seasons and at all physiographic regions. This highlights the importance of incorporating a sufficient number of stations in the blending process, as the number of stations used in the gauge-satellite rainfall depths after 1992 has reduced drastically (Fig. 3). Furthermore, results confirm that there exists a systematic underestimation of rainfall depths in both satellite products as compared with the ground measurements, especially at the High Mountain region of the Koshi basin, irrespective of the seasons (see Table 3).

Table 3
Average rainfall recorded in the DHM (ground-based) and CHG (satellite only – CHIRP and station blended – CHIRPS) at different physiographic regions (Terai + Siwalik, Middle and High Mountain) and in different seasons for the whole data set (1981–2010) and for the data set after 1992 (1992–2010).

Whole Data Set (1981–2010)															
Regions →	Terai + Siwalik					Middle Mountain					High Mountain				
Sources↓	DJF	MAM	JJAS	ON	Yearly	DJF	MAM	JJAS	ON	Yearly	DJF	MAM	JJAS	ON	Yearly
DHM	29	187	1183	81	1480	45	248	1271	72	1636	57	334	1944	112	2447
CHIRPS	14	140	1083	45	1281	22	180	1101	50	1353	26	179	1237	60	1502
CHIRP	17	143	1051	47	1258	27	185	1068	50	1329	33	184	1150	61	1428
Data Set after 1992 (1992–2010)															
Regions →	Terai + Siwalik					Middle Mountain					High Mountain				
Sources↓	DJF	MAM	JJAS	ON	Yearly	DJF	MAM	JJAS	ON	Yearly	DJF	MAM	JJAS	ON	Yearly
DHM	27	187	1197	89	1499	42	241	1246	69	1599	55	344	1882	114	2395
CHIRPS	13	132	1056	45	1247	19	166	1035	44	1264	22	160	1095	49	1326
CHIRP	17	147	1115	55	1333	25	190	1111	55	1382	33	187	1165	65	1450

DJF: December-January-February; MAM: March-April-May; JJAS: June-July-August-September; ON: October-November.

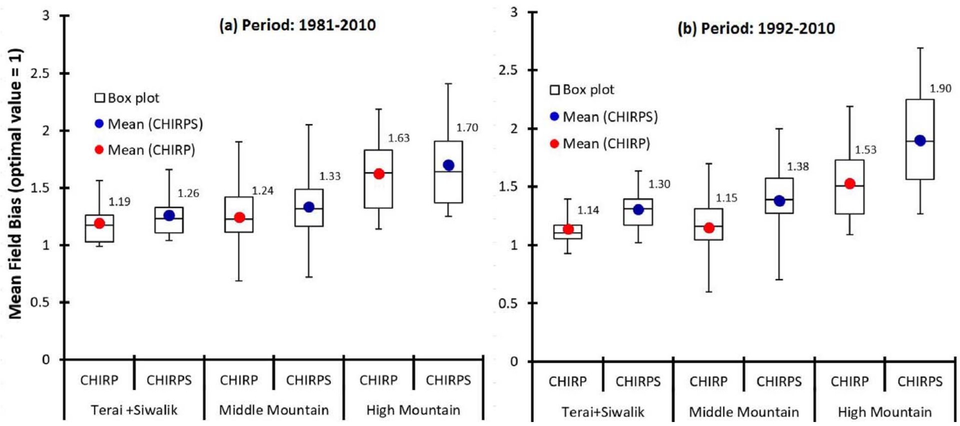


Fig. 7. Box plots of the Mean Field Bias (MFB) values of the whole data set (1981–2010, Left) and data after 1992 (1992–2010, right) for the different physiographic regions.

From the above analysis, it is clear that the satellite-only product (CHIRP) is a better data set than the blended product (CHIRPS) at all physiographic regions except for during the summer (JJAS) season when considering the whole data set (1981–2010). For data sets after 1992 (1992–2010), the satellite-only product (CHIRP) is found to be better. Hence, at least for our case, blending the station data with the satellite product has not improved the accuracy of the CHG product.

3.2. Mean field bias (MFB) correction results

The physiographic region-wise MFB values calculated as an average of the individual field bias of all the valid pairs have depicted a scenario almost similar to that observed before (see Fig. 7). On average, the MFB values are relatively lower in the low-lying regions (Terai and Siwalik). The serious underestimation of rainfall depths by the satellite products in the High Mountain region has been reflected in rather higher values of the MFB (up to 1.9). While the MFB calculations for the Middle Mountain region are comparable to those for the low-lying regions, a higher variation has been observed for the former, as indicated by the extent of error bars. The MFB values calculated for all physiographic regions has been found to be consistently above 1.0, which further confirms the underestimation of rainfall depths by both CHG products. Out of the two CHG products, the satellite-only (CHIRP) data set resulted in lower MFB values for both time spans and at all physiographic regions, again highlighting the superiority of the satellite-only product (CHIRP) over the blended product (CHIRPS).

The effect of physiographic region-wise MFB correction on the rainfall estimates of both CHG products has been clearly reflected in the probability of exceedance plots (see Fig. 8). As can be seen in the figure, the effect is more pronounced especially in the rainfall

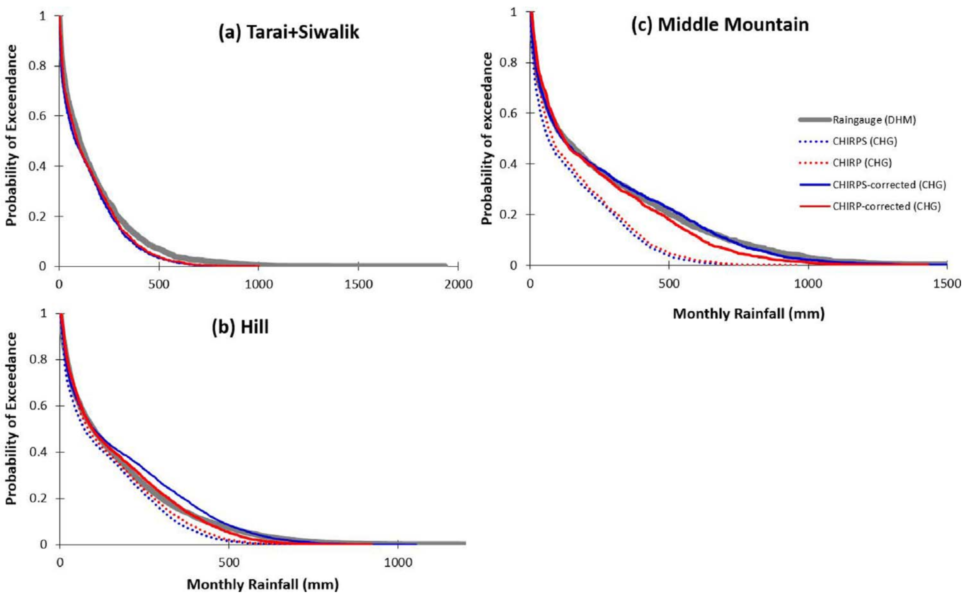


Fig. 8. Effect of the Mean Field Bias (MFB) correction on the CHG products in terms of the probability of exceedance plots for the data set after 1992 (1992–2010).

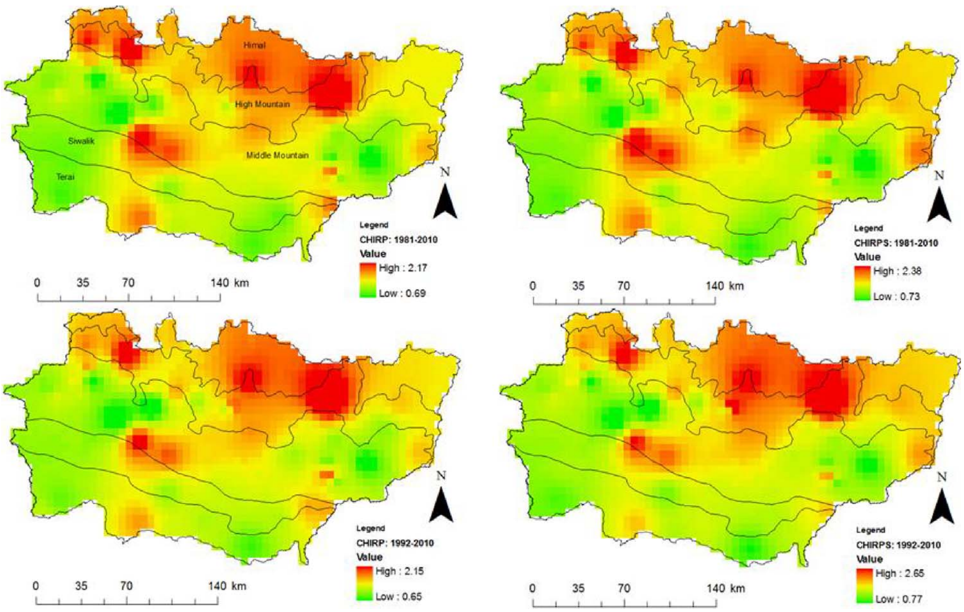


Fig. 9. Spatial interpolation of the Mean Field Bias (MFB) values in the Koshi basin; Top left: CHIRP (1981–2010); Bottom Left: CHIRP (1992–2010); Top Right: CHIRPS (1981–2010); and Bottom Right: CHIRPS (1992–2010).

estimates at the High Mountain region. In the Middle Mountain region, the effect of MFB correction on the station-blended product (CHIRPS) seems to be bit strong as the corrected CHIRPS rainfall depths are now slightly overestimated as compared to the ground measurements. From the analysis, it can be stated that even a simple multiplicative correction such as the physiographic region-wise MFB correction could improve the accuracy of CHG products quite significantly. Hence, it is recommended that the bias of the CHG products are assessed and, if needed, corrected before using them for any further applications. It could then be argued that the usefulness of CHG products would be limited, as bias correction requires the availability of high-resolution ground truth data from a dense rain gauge network. For regions with limited ground truth data (thereby limiting a robust accuracy assessment and correction, if required) CHG products should be used with caution.

From our analysis, it is also clear that the bias between the CHG product and ground measurements varied between the gauges, and this shows the need for strong bias corrections at the High Mountain region. In view of our aim of deriving precipitation based at spatial scales of the CHG product (5 km × 5 km), it was felt that a single correction factor for a physiographic region would be a simplification, especially when there are gauges in a single physiographic region which show both underestimation and overestimation of CHG products (see Fig. 7). Hence, the individual MFBs of each rain gauge have been preserved and have been used to create spatially varied MFB correction layers. Fig. 9 shows spatially interpolated MFB values using the Inverse Distance Weighing (IDW) in the ArcGIS for all satellite products and time spans in the spatial scale of the Koshi basin. This layer would then be used to correct the CHG products, and then for the precipitation-based drought index map of the Koshi basin.

3.3. Standardized precipitation index (SPI) results

In an attempt to explore how the differences in rainfall depths between ground- and satellite-based data would be reflected in SPI values, an analysis was made at a station named Kathmandu Airport (station #21; see Fig. 2). This station is purposely chosen; it has no missing data during the entire period (1981–2010), and the data is consistent with surrounding stations' data sets. One-month SPI values were calculated for this purpose.

The time series plot of one-month SPI values calculated based on ground-based (DHM) and both CHG-based (CHIRP and CHIRP) rainfall depths, and using the whole period (1981–2010), have been presented in Fig. 10 (top). For clarity purposes, the plot of the last decade (2000–2010) has been shown in the figure. As already depicted, a zero value of the SPI indicates the normal condition, positive values of the SPI indicate wetter conditions (more rainfall observed at that particular time when compared against long-term statistics), and negative values of the SPI indicate drought conditions (deficit rainfall at the particular time when compared against long-term statistics). The severity of drought conditions is indicated by different levels of the SPI values, as suggested by WMO (2012). SPI values of less than −1.0 mean a moderately dry condition, values between −1.0 and −2.0 mean a severely dry condition, and values of more than −2.0 mean an extremely dry condition. This particular station experiences several instances of dry and wet spells as indicated by frequent troughs and peaks in the SPI time series (see Fig. 10). A well-known winter drought during 2008–09 (MoAC et al., 2009) was also evident in the plot where the SPI values have nearly crossed the extreme drought limit (SPI: −2.0). Similarly, some reported summer droughts, such as in the years 2004, 2005, and 2009 (Dahal et al., 2015), as well as during the winter drought of 2005-06, have also been reflected in the SPI values.

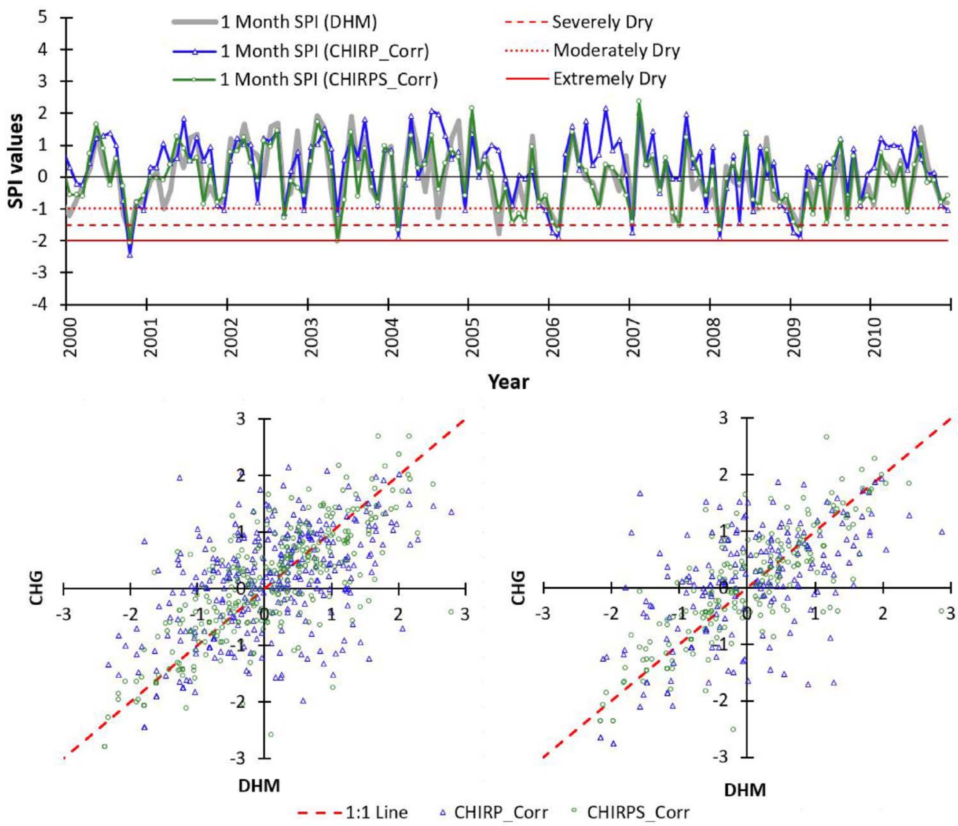


Fig. 10. Time series plot of 1-month SPI values at the Kathmandu Airport, calculated using ground-based (DHM) and satellite-based CHG products (CHIRP and CHIRPS – MFB corrected) using the whole data set (1981–2010). Time series plot of last decade is shown for clarity (Top). Scatter plots of the 1-month SPI values using the whole data set (1981–2010, Bottom Left) and data using the data set after 1992 (1992–2010, Bottom Right).

While the statistics calculated by comparing the valid pairs between ground- and satellite-based rainfall depths at the considered station were found to be comparable (CHIRP: PBIAS = −1%, RSR = 0.5 and CHIRPS: PBIAS = −2% and RSR = 0.5), the differences have been translated into SPI values in a slightly exaggerated manner. As can be seen in time series plot (Fig. 10, top), the station-blended product (CHIRPS) based SPI values have matched slightly better (R^2 = 0.61) with the ground-based SPI values than the satellite-only product (CHIRP) based SPI values (R^2 = 0.33) have. This is also evident on the scatter plots (bottom left for the period 1981–2010 and bottom right for the period 1992–2010), as the SPI values based on satellite-only (CHIRP) data sets showed marked scattering compared to the CHIRPS-based SPI values. Hence, even though the difference in the rainfall depths is small, the difference could be significant in terms of SPI. This behavior could be related to the fitting (parameter estimation) of the Gamma distribution of individual data sets for the calculation of the SPI values. Further study is needed to explore the sensitivity of the parameters of the distribution in different data sets.

The SPI results for the entire Koshi basin also showed different results when using the two CHG data sets (see Fig. 11 for the whole period of 1981–2010 and Fig. 12 for the period of 1992–2010). The summer (JJA, three-months SPI) drought of 1992 has been calculated and used for this purpose. The pixel-by-pixel comparison of the SPI values for the whole data set (1992–2010) showed that the satellite-only product (CHIRP) based SPI values tend to underestimate the severity of drought which is also clear from the scatter plot having rather mild R^2 values of 0.49; also, almost all the SPI values are lying above the 1:1 line. While the station-blended product (CHIRPS) based drought rating indicates 'Extremely Dry' (SPI values < -2.0) for almost the entire basin; the satellite-only product (CHIRP) based rating is a bit milder. The SPI values calculated using the data set after 1992 (1992–2010), however, show rather comparable SPI values (see Fig. 12). This is also evident in the scatter plot (see Fig. 12), as data points are around the 1:1 line, and the R^2 value (0.68) has improved significantly, too. It should, however, be noted that the SPI calculations for the period after 1992 (1992–2010) have been based on the time series data consisting of only eighteen years. While WMO (2012) suggested using a non-missing time series data of at least twenty to thirty years, some researchers (e.g., Guttman, 1994) have argued that a longer period (fifty to sixty years or more) is "optimal and preferred". Hence, the SPI values calculated for the period after 1992 should be used with caution, as studies (e.g., WMO, 2012) suggested that SPI calculations using shorter time series might "weaken the confidence". However, in our case, comparing the one-month SPI values calculated using the whole data set (1981–2010) and the data set after 1992 (1992–2010) at a pixel containing the Kathmandu Airport rain gauge, show rather minimal differences. For both products (CHIRP and CHIRPS), the R^2 values calculated against the one-month SPI values estimated using the data sets of different lengths is found to be >0.91 (see Fig. 13).

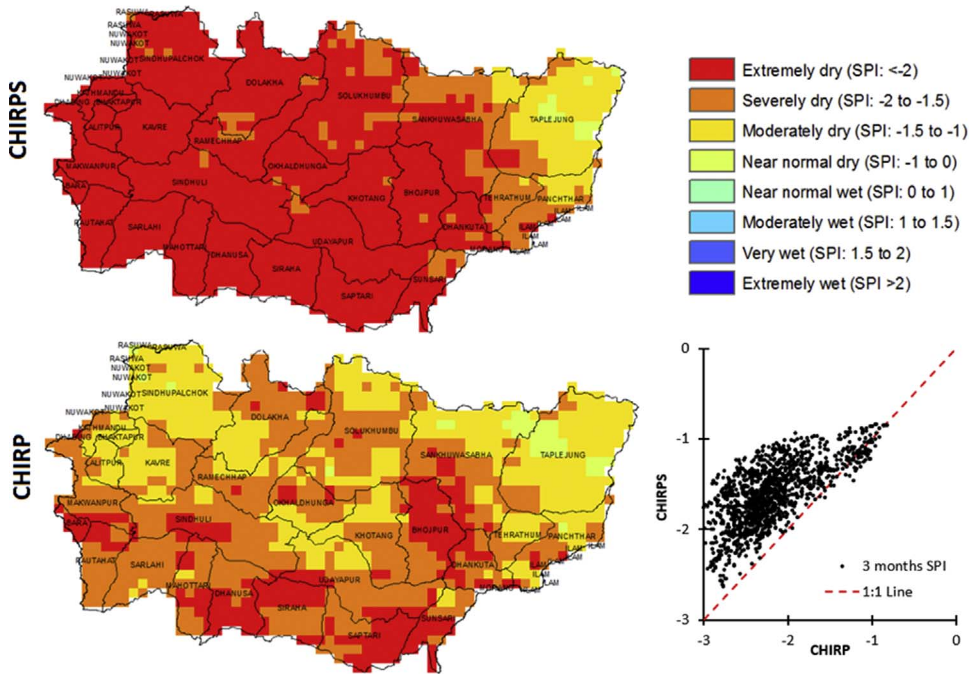


Fig. 11. Summer (JJA) drought of 1992 represented as 3-months SPI over the Koshi basin (Nepal part only) calculated using both CHG data set of the whole (1981–2010) period. Also shown is the scatter plot of SPI values calculated from the two different CHG products.

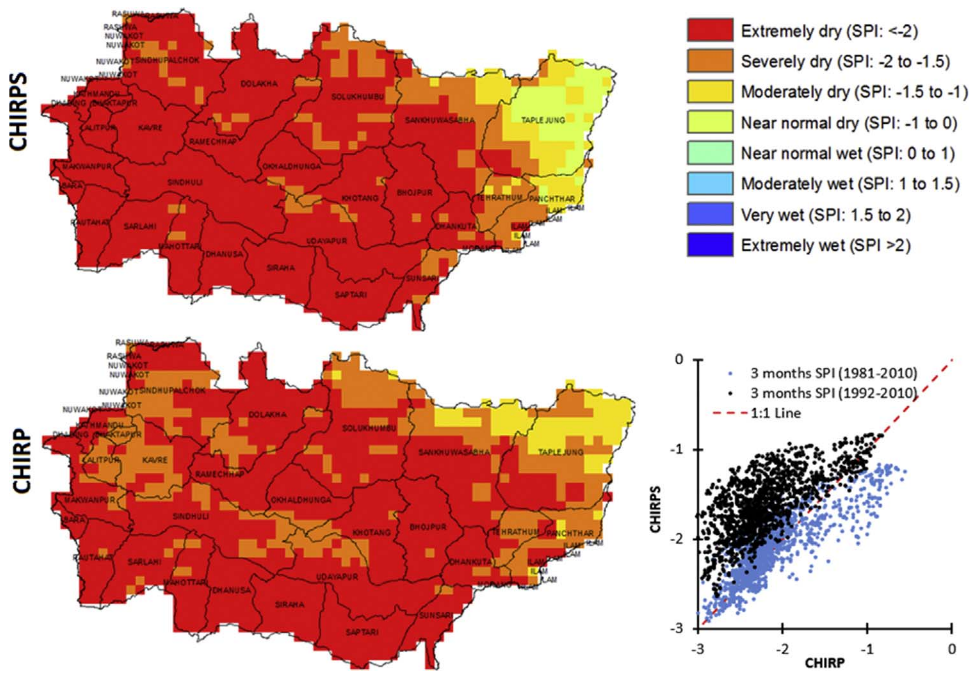


Fig. 12. Summer (JJA) drought of 1992 represented as 3-months SPI over the Koshi basin (Nepal part only) calculated using both CHG data set of the period after 1992 (1992–2010) period. Also shown is the scatter plot of SPI values calculated from the two different CHG products.

From the analysis, it is apparent that the SPI values calculated based on the satellite-only product (CHIRP) and station-blended product (CHIRPS) are almost comparable when using the data set after 1992 (1992–2010). Hence, despite the limitations of using a shorter time frame (less than thirty years) while calculating the SPI values, the satellite-only CHG product (CHIRP), which is readily available, can be used for precipitation-based drought monitoring, assessment, and forecasting of the Koshi basin.

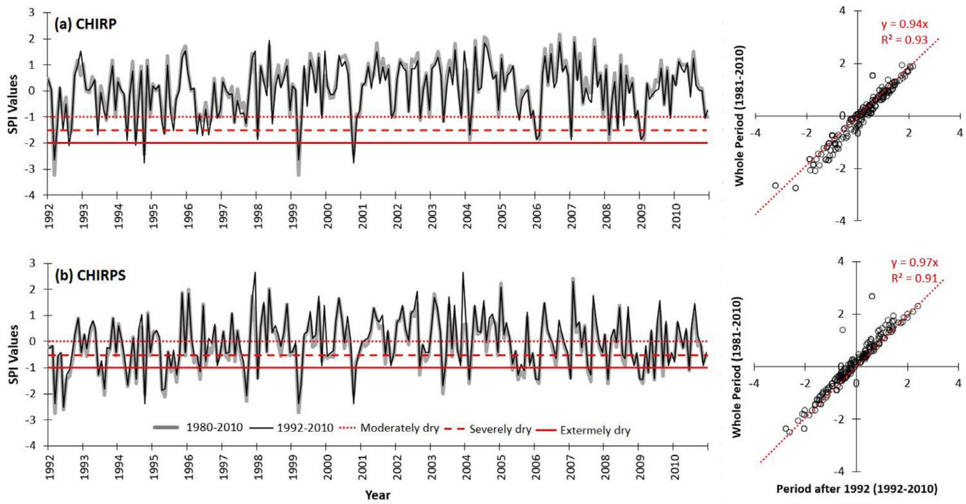


Fig. 13. Time series and scatter plots of 1-month SPI values calculated considering the whole data set (1981–2010) and considering the data set after 1992 (1992–2010) of MFB corrected (a) satellite only product (CHIRP) and (b) station blended product (CHIRPS) at a pixel above the Kathmandu Airport raingauge station.

4. Conclusions and recommendations

An assessment was made to check the accuracy of two Climate Hazard Group (CHG) satellite products – namely, the satellite-only (CHIRP) and station-blended (CHIRPS) products – based on a rain gauge network consisting of forty-eight rain gauge stations operated and maintained by the Department of Hydrology and Meteorology (DHM), Nepal over the entire Koshi basin (Nepal part). Analyses were based on monthly rainfall depths over a time span of thirty years (1981–2010). Separate analyses were made for the data set after 1992, as the number of stations used in the blending process has significantly reduced since 1992.

From the analysis, it was found that the rainfall estimates from both CHG data sets were fairly comparable to ground-based measurements. The station-blended product (CHIRPS) was marginally better while considering the whole period (1981–2010). However, for the period after 1992 (1992–2010), the satellite-only product (CHIRP) was better. This was confirmed by the comparison statistics at different physiographic regions of the basin and from the Mean Field Bias (MFB) values. In general, both CHG products tend to underestimate the rainfall depths systematically as compared to the ground-based measurements in all seasons. The accuracy of the CHG data sets were found to be better in the low-lying regions while it was the worst in higher elevation regions. In mid-lands, accuracy of the data varied from station to station while average statistical parameters showed comparable results as those in the low-lying areas. Physiographic region wise MFB correction seemed to improve the accuracy of the CHG products quite significantly, especially in the higher elevation region.

Analyses made at a station where the accuracy of both CHG products was almost the same showed that even the minimum difference in rainfall depths tends to amplify when using the CHG products to calculate the Standardized Precipitation Index (SPI). This could be related to the parameter estimation of the gamma distribution that the SPI employs. This case indeed needs further investigation. Finally, the SPI values calculated over the entire Koshi basin showed that the SPI values based on the two CHG data sets were fairly comparable, especially when using data after 1992.

It is therefore concluded that CHG data sets could be used after correction and that, especially after 1992, the satellite-only product (CHIRP), which is more readily-available than the station-blended product (CHIRPS), can be used for precipitation-based drought monitoring of the Koshi basin.

Acknowledgements

This study was part of the Koshi Basin Programme of ICIMOD funded by the Department of Foreign Affairs and Trade (DFAT) of Australia. The core datasets were produced under the SERVIR Hindu Kush-Himalaya (HKH) Programme supported by the United States Agency for International Development (USAID) and National Aeronautics and Space Administration (NASA). We gratefully acknowledge the support from Department of Hydrology and Meteorology, Nepal for providing observed station data. The authors also gratefully acknowledge the support of core donors of ICIMOD: the governments of Afghanistan, Australia, Austria, Bangladesh, Bhutan, China, India, Myanmar, Nepal, Norway, Pakistan, Switzerland, and the UK. The views expressed are those of the authors and do not necessarily reflect that of ICIMOD.

References

AghaKouchak, A., Behrangi, A., Sorooshian, S., Hsu Amitai, K.E., 2011. Evaluation of satellite-retrieved extreme precipitation rates across the central United States. J. Geophys. Res.: Atmospheresn 116 (D2) (/a-n/a).

Bayissa, Y.A., Moges, S.A., Xuan, Y., Van Andel, S.J., Maskey, S., Solomatine, D.P., Griensven, A.V., Tadesse, T., 2015. Spatio-temporal assessment of meteorological drought under the influence of varying record length: the case of Upper Blue Nile Basin, Ethiopia. Hydrol. Sci. J. 60 (11), 1927–1942.

Dahal, P., Shrestha, N.S., Shrestha, M.L., Krakauer, N.Y., Panthi, J., Pradhanang, S.M., Jha, A., Lakhankar, T., 2015. Drought risk assessment in central Nepal: temporal and spatial analysis. Nat. Hazards 80 (3), 1913–1932.

Dinku, T., 2014. Validation of the CHIRPS satellite rainfall estimate. In: Proceedings of the 7th International Precipitation Working Group (IPWG) Workshop. Tsukuba, Japan, 17–21 November 2014. International Precipitation Working Group (IPWG).

Funk, C.C., Peterson, P.J., Landsfeld, M.F., Pedreros, D.H., Verdin, J.P., Rowland, J.D., Romero, B.E., Husak, G.J., Michaelsen, J.C., Verdin, A.P., 2014. A quasi-global precipitation time series for drought monitoring. Survey. U. S. G.

Funk, C., Verdin, A., Michaelsen, J., Peterson, P., Pedreros, D., Husak, G., 2015. A global satellite assisted precipitation climatology. Earth Syst. Sci. Data Discuss. 8 (1), 401–425.

Guttman, N.B., 1994. On the sensitivity of sample l moments to sample size. J. Clim. 7, 1026–1029.

Hessels, T.M., 2015. Comparison and Validation of Several Open Access Remotely Sensed Rainfall Products for the Nile Basin. Delft University of Technology, the Netherlands.

Katsanos, D., Retalis, A., Michaelides, S., 2016. Validation of a high-resolution precipitation database (CHIRPS) over Cyprus for a 30-year period. Atmos. Res. 169 (Part B), 459–464.

Khandu, wange, J.L., Forootan, E., 2016. An evaluation of high-resolution gridded precipitation products over Bhutan (1998–2012). Int. J. Climatol. 36 (3), 1067–1087.

Lack, S.A., Fox, N.I., 2007. An examination of the effect of wind-drift on radar-derived surface rainfall estimations. Atmos. Res. 85 (2), 217–229.

McKee, T.B., Doesken, N.J., Kleist, J., 1993. The relationship of drought frequency and duration to time scales. In: Proceedings of the 8th Conference on Applied Climatology. 17–22 January 1993, Anaheim, California. American Meteorological Society.

Mishra, A.K., Singh, V.P., 2010. A review of drought concepts. J. Hydrol. 391 (1-2), 202–216.

MoAC, WFP, FAO, 2009. 2008/09 winter drought in Nepal: crop and food security assessment. Joint Assessment Report.

Moriasi, D.N., Arnold, J.G., Van Liew, M.W., Binger, R.L., Harmel, R.D., Veith, T., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. ASABE 50 (3), 885–900.

NMDC, 2016. Program to Calculate Standardized Precipitation Index. National Drought Mitigation Center, Colorado State University, USA.

Narendra, B.H., 2008. Drought Monitoring Using Rainfall Data and Spatial Soil Moisture Modelling. Gadjah Mada University.

Palmer, W.C., 1965. Meteorological Drought. U.S. Department of Commerce Weather Bureau, pp. 58.

Pedreros, D.H., Rojas, A., Funk, C., Peterson, P., Landsfeld, M.F., Husak, G.J., 2014. The Use of CHIRPS to Analyze Historical Rainfall in Colombia.

Prakash, S., Sathiyamoorthy, V., Mahesh, C., Gairola, R.M., 2014. An evaluation of high-resolution multisatellite rainfall products over the Indian monsoon region. Int. J. Remote Sens. 35 (9), 3018–3035.

Shah, R.D., Mishra, V., 2015. Seasonal drought prediction in India. In: Proceedings of the American Geophysical Union. Fall Meeting 2015, San Francisco, California, USA. American Geophysical Union.

Shah, H.L., Mishra, V., 2016. Uncertainty and bias in satellite-based precipitation estimates over indian sub-continental basins: implications for real-time streamflow simulation and flood prediction. J. Hydrometeorol.

Shrestha, M., Takara, K., Kubota, T., Bajracharya, S.R., 2011. Verification of GSMaP rainfall estimates over the central Himalayas. J. Jpn Soc. Civil Eng. Ser. B1 (Hydraul. Eng.) 67 (4) (6).

Shrestha, M.S., Rajbhandari, R., Bajracharya, S.R., 2013a. Validation of NOAA CPC_RFE Satellite-based Rainfall Estimates in the Central Himalayas. International Center for Integrated Mountain Development (ICIMOD).

Shrestha, N.K., Goormans, T., Willems, P., 2013b. Evaluating the accuracy of C- and X-band weather radars and their application for stream flow simulation. J. Hydroinf. 15 (4), 1121–1136.

Shrestha, M.S., 2011. Bias-Adjustment of Satellite-Based Rainfall Estimates over the Central Himalayas of Nepal for Flood Prediction. Kyoto University, Japan.

Szalai, S., Szinell, C., Zoboki, J., 2000. Early Warning Systems for Drought Preparedness and Drought Management. WMO, pp. 161–176.

Tokarczyk, T., Szalińska, W., 2014. Combined analysis of precipitation and water deficit for drought hazard assessment. Hydrol. Sci. J. 59 (9), 1675–1689.

Toté, C., Patricio, D., Boogaard, H., van der Wijngaart, R., Tarnavsky, E., Funk, C., 2015. Evaluation of satellite rainfall estimates for drought and flood monitoring in Mozambique. Remote Sens. 7 (2), 1758.

WMO, 1986. Report on Drought and Countries Affected by Drought During 1974–1985. World Meteorological Organization, Geneva, pp. 118.

WMO, 1994. Observing the World's Environment: weather, climate and water. WMO Report No. 796. World Meteorological Organization, Geneva.

WMO, 2012. Standardized Precipitation Index User Guide. World Meteorological Organization, Geneva.

Wilhite, D.A., 1992. Preparing for Drought: A Guidebook for Developing Countries, Climate Unit. United Nations Environment Program, Nairobi, Kenya.

Wong, G., van Lanen, H.A.J., Torfs, P.J.J.F., 2013. Probabilistic analysis of hydrological drought characteristics using meteorological drought. Hydrol. Sci. J. 58 (2), 253–270.

Xie, P., Arkin, P.A., 1995. An intercomparison of gauge observations and satellite estimates of monthly precipitation. J. Appl. Meteorol. 34 (5), 1143–1160.

Xue, X., Hong, Y., Limaye, A.S., Gourley, J.J., Huffman, G.J., Khan, S.I., Dorji, C., Chen, S., 2013. Statistical and hydrological evaluation of TRMM-based multi-satellite precipitation analysis over the Wangchu Basin of Bhutan: are the latest satellite precipitation products 3B42V7 ready for use in ungauged basins? J. Hydrol. 499, 91–99.