Validation of Satellite Rainfall Estimation in the Summer monsoon Dominated Area of the Hindu Kush Himalayan Region

S. R. Bajracharya, M. S. Shrestha, P. K. Mool, and R. Thapa
International Centre for Integrated Mountain Development (ICIMOD), Kathmandu, Nepal

Abstract

Weather events like intensive rainfall causing floods and flash floods result into loss of lives and properties whereas prolong drought can cause decline in agriculture production and loss of vegetation cover. Rainfall affects the lives and economies of majority of people where the populations are dependent on rain water for agriculture. With the existence of large unpopulated rugged terrain with limited number of observation hydro-meteorological stations, accurate rainfall estimation is a challenging task and the spatial distribution of the rain gauge is not sufficient to provide a detail outlook on highly temporal and spatial variable nature of rainfall that may be needed for applied stream flow modeling technique.

In the present paper, the estimated 24 hours rainfall product developed by National Oceanic and Atmospheric Administration (NOAA) in a South Asian domain was validated with the observed rain gauge data on a daily basis for the monsoon period of 2002 to 2004. The result shows maximum negative bias and root mean square error (RMSE) in the heavy rainy days and Satellite Rainfall Estimation (SRE) overestimates the rain before monsoon and in rain shadow area. Qualitatively rainfall events in general match but quantitatively SRE and observed rain gauge product are vast difference. The study provides important input for the improvement of the SRE development algorithms. Further, incorporation of orographic effect in the algorithm is felt necessary before it should be implemented to the stream flow model for flood forecasting.

KEY WORDS: Satellite Rainfall Estimation, flood forecasting, monsoon, validation

1. Introduction

Precipitation is an integral component of the hydrologic cycle, and accurate rainfall estimates are basis for meteorology, hydrology and environmental science and also necessary to improve short-, medium-, long-term weather forecasts and climate prediction (Kamarianakis,Y.et al. 2006). Rainfall affects the lives and economies of almost all people in the HKH region, where a large percentage of the farmers depend on rainwater for agriculture. Excess water causes riverine floods, as well as flash floods and other water-induced disasters. This results in loss of lives and huge economic damages to infrastructures and properties. Over the last 30 years South Asia has seen more than 65,000 deaths, and approximately a billion people have been affected by floods and landslides, accounting for about 33% of all the flood events of Asia (Shrestha and Takara 2007). High rates of poverty and population growth have increased the vulnerability to flood disasters. Flooding poses severe constraints on socioeconomic development, including investments in agriculture, infrastructure, and industrial production. Reliable and timely flood forecasting and warning are some of the most effective non-structural measures to minimize the loss of life and the socioeconomic impacts of floods.

The satellite rainfall estimation (SRE) technique provides information on rainfall occurrence, amount, and distribution over the region. An important technology for rainfall measurement that provides near real-time data, it can be used alongside conventional gauge data. It is expected that applying satellite rainfall estimation into stream flow model will lead to more precise, timely, and accurate flood warnings. With sufficient lead times, people can evacuate to safer places ahead of a disaster, therefore reducing the loss of life and property. The use of satellite-derived quantitative rainfall estimate technology can be crucial for obtaining rainfall patterns to be used to forecast discharge, study hydrological cycles, plan water management, provide flash floods guidance, monitor drought, and plan agriculture in the HKH region. Hence using advanced remote sensing tools and techniques as satellite rainfall estimation (SRE) would provide reliable and timely data to supplement the gauge stations and fill in the data gaps to forecast floods with greater accuracy.

It is therefore important to have an idea of their accuracy and expected error characteristics. This is done by validating the satellite precipitation estimates against ground values from rain gauges. A thorough verification of satellite-based precipitation products should quantify their accuracy in a wide range of weather and climate regimes; give users information on expected errors in the estimates and in which applications they should be used; and help developers understand the strengths and weaknesses of the satellite rainfall algorithms by showing which aspects need the most improvement, monitoring the performance of existing algorithms, and assisting with evaluating algorithm upgrades (Source:http://www.bom.gov.au/bmrc/SatRainVal/validation-intercomparison.html).

The National Oceanic and Atmospheric Administration (NOAA) has developed several satellite-based techniques and algorithms for estimating rainfall that support the weather and flood monitoring activities of United States Agency for International Development (USAID) and United States Geological Survey (USGS). Among them is the system developed at the Climate Prediction Centre (CPC) of NOAA known as the CPC-RFE2.0. CPC-RFE2.0 estimates precipitation for the whole globe on a 0.1° x 0.1° grid with 24 hours temporal resolution. This product has been implemented to the southern Asia region, including the Mekong River Basin, the Hindu Kush Himalayan region, and surrounding areas in May of 2001. The system first combined linearly three satellite estimates, and then merged with rain gauge station data to determine the final product, which significantly increases accuracy by reducing bias and random error compared to individual data sources (Xie and Arkin 1996).

This paper contains the results of validation of SRE in summer monsoon dominated area at regional levels. Figure 1 provides an example of the daily satellite rainfall estimate of the HKH region provided by CPC-RFE2.0

2. Working area and data availability

2.1. Working area

The HKH region includes the mountains of South Asia and the Tibetan Plateau. It extends 3,500 km west to east, covering Afghanistan, Pakistan, and the Tibetan Autonomous Region of the People's Republic of China, Nepal,

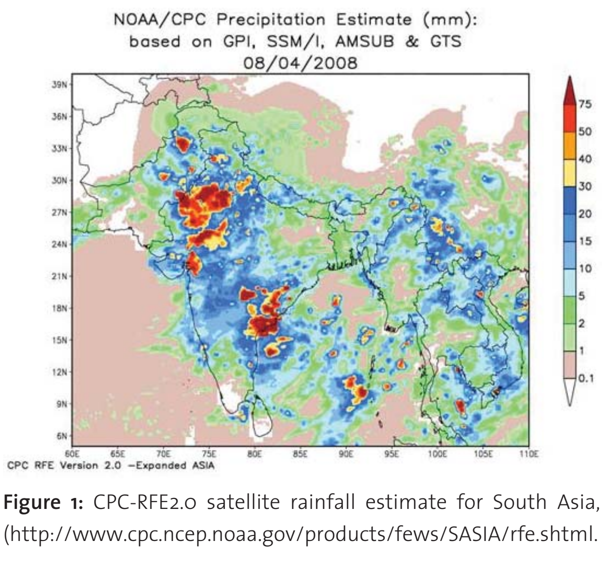

Figure 1: CPC-RFE2.0 satellite rainfall estimate for South Asia, (http://www.cpc.ncep.noaa.gov/products/fews/SASIA/rfe.shtml.

Bhutan, Bangladesh, and Myanmar. This region has a very complex terrain with high elevation and acts as physical and climatic barrier, playing an important role in atmospheric circulation. It forces southwesterly flow to change its direction to easterly during the monsoon season, causing rainfall to decrease from east to west. The rainfall pattern in this region is erratic and the spatial distribution of rainfall depends on the orographic profiles. Likewise in some areas the annual rainfall is 12293 mm (Cherrapunji in India), while in the Trans-Himalayan region (Mustang, Jumla in Nepal) is less than 50 mm. In addition, high intensity, short duration rainfall is common and can deliver more than 500 mm in 24 hours. The term 'Hindu Kush-Himalayan region' is used loosely to describe the area covering all the high mountain chains of Central, South, and Inner Asia, including the Tien Shan, Kun Lun, Pamir, Hindu Kush, Karakoram, Himalaya, and Hengduan mountain ranges; the extensive middle-mountain chains that surround them; and the high-altitude Tibetan Plateau. This region is often called the 'Roof of the World' (Xu.J.et al., 2007). The area extends from about 5 °N – 40 °N and 60 °E – 110 °E, as shown in Fig. 2.

2.2. Data availability

2.2.1. Estimated rainfall data

Three years of 24-hours CPC-RFE gridded rainfall data of 0.1° x 0.1° in Lambert Azimuthal Equal Area projection were obtained from Climate Prediction Center of NOAA (2002 to 2004) for the HKH region. The NOAA CPC data are based on the combination of daily Global Telecommunication System (GTS) rain-gauge data, Advanced Microwave Sounding Unit (AMSU) satellite precipitation estimates, Special Sensor Microwave/Imager (SSM/I) satellite rainfall estimates, and Geostationary Operational Environmental Satellite (GOES) Precipitation Index (GPI) cloud-top infrared (IR) temperature precipitation estimates. The area cover of the South Asian CPC_RFE 2.0 product is shown in Fig. 3.

2.2.2. Observed rainfall data

The used reference daily raingauge data in this study for 2002 to 2004 from 373 rainfall stations in six countries in the HKH region: Bangladesh (53 stations), Bhutan (73), China (38), India (33), Nepal (176), and Pakistan (49). This rain-

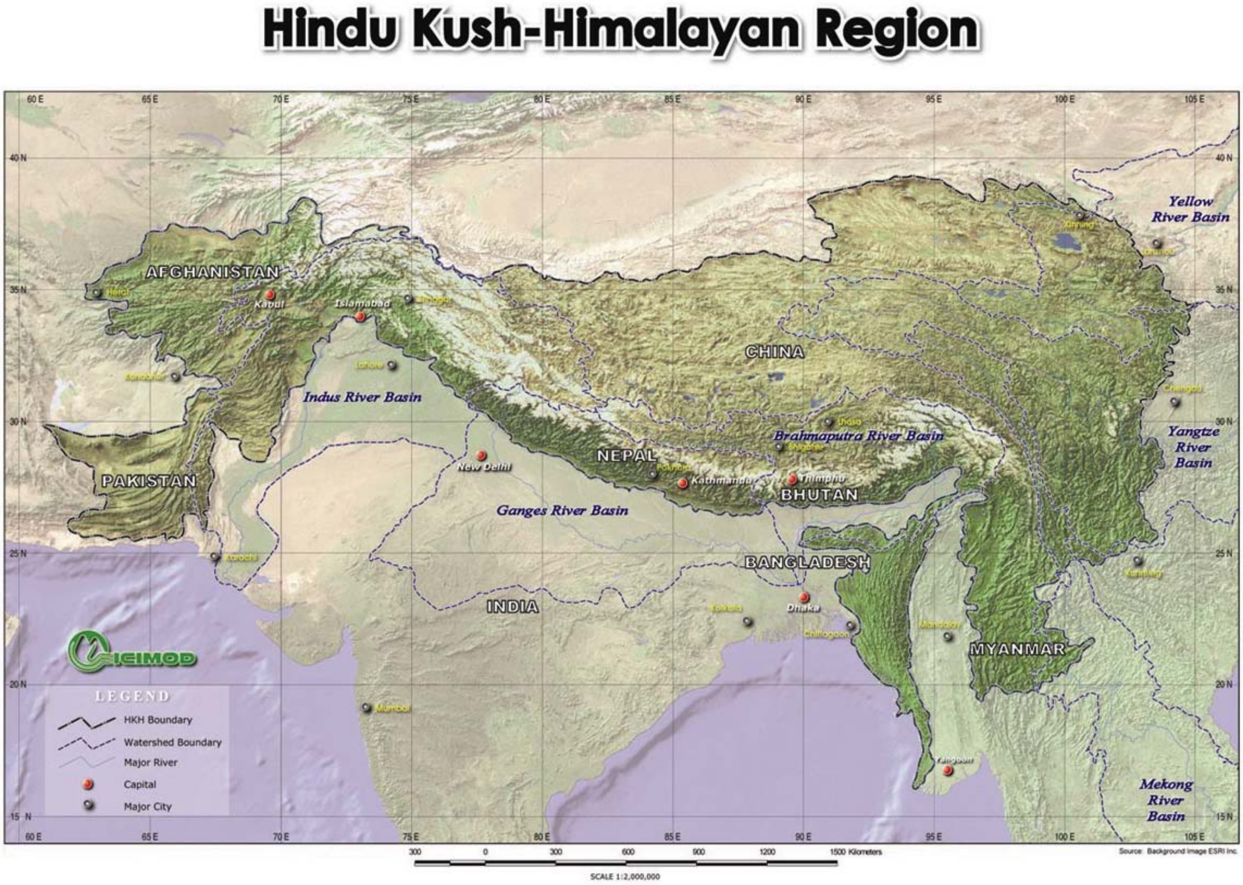

Figure 2: The spatial domain of Hindu Kush-Himalayan Region.

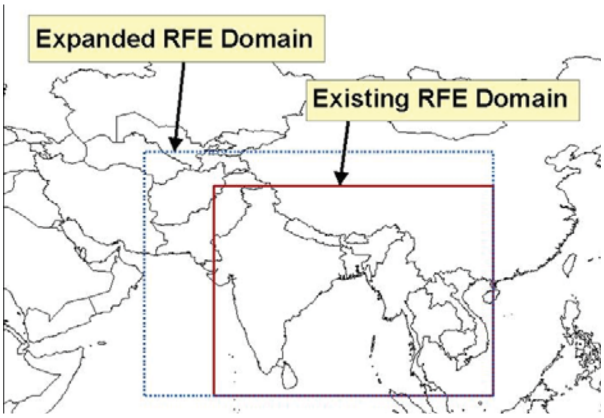

Figure 3: Spatial domain of the CPC-RFE2.0 (new &old domain).

fall station data were, provided by the Bangladesh Water Development Board, Hydro-meteorological Services Division Bhutan, Tibetan Meteorological Bureau, North-East Centre for Environmental Research and Development India, Department of Hydrology and Meteorology Nepal and Pakistan Meteorological Department had been used for validation purposes. The density of the rain gauges varies in each country.

Country or Province	Rainfall Period	Number of Rainfall Stations	Number of GTS Stations
Bangladesh	2002 - 2004	53	10
Bhutan	2002 - 2004	73	-----
China	2002 - 2004	38	38
Himachal Pradesh, India	2002 - 2004	3	-----
North Guwahati, Assam, India	2002 - 2004	30	20
Nepal	2002 - 2004	176	12

Table 1: Summary of Rainfall Stations used for the study.

The density of rainfall stations in Nepal is high compared to other countries in the region. However, the distribution of rainfall stations in Nepal and Bhutan is uneven and very sparse in the northern areas. Most stations are concentrated in urban and middle mountain areas where accessibility is easy. Tab. 1 summarises the rainfall data available for the analysis, listing the amount of station data available country-wise, and also the number of GTS stations. The GTS stations are a subset of the total number of rainfall stations. The geographical distribution of the rainfall

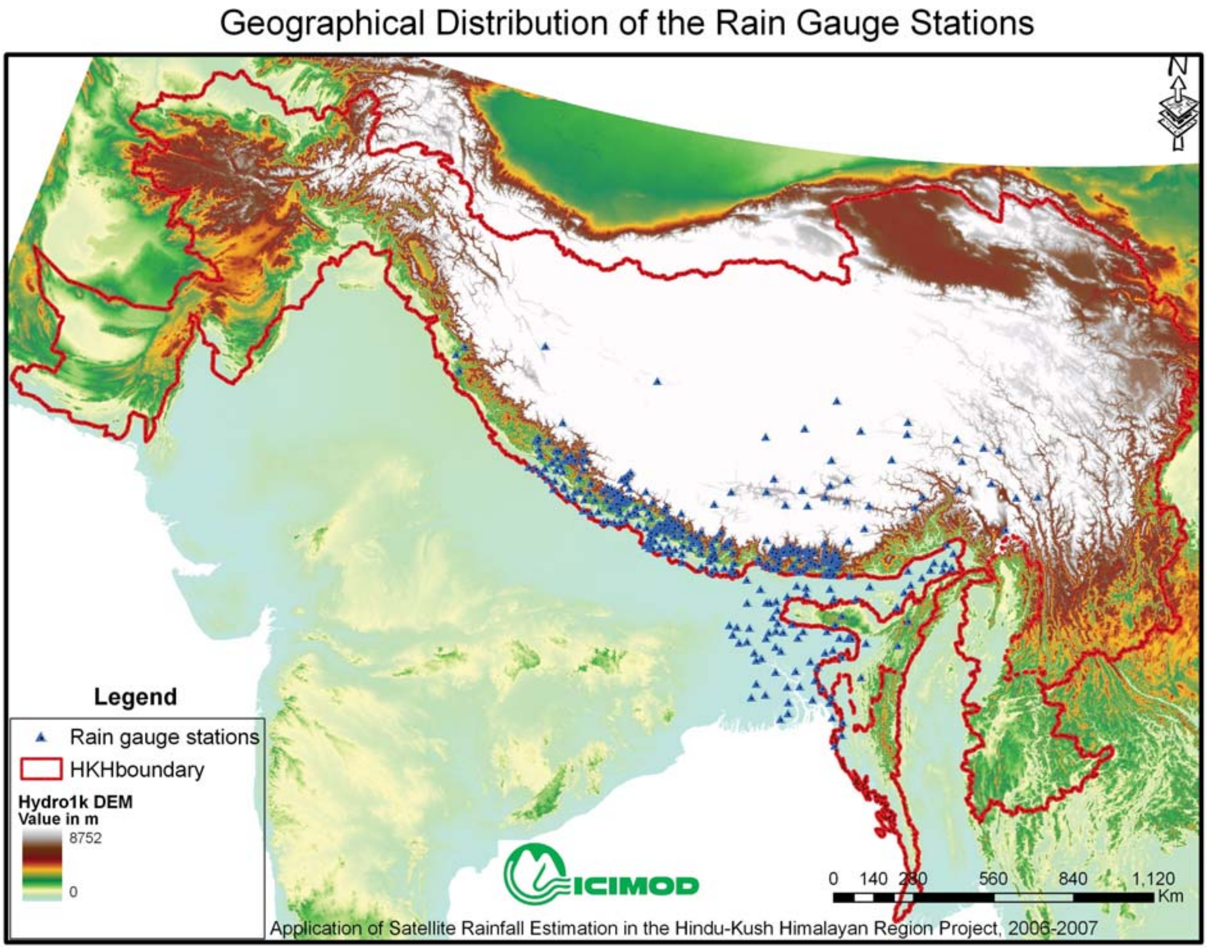

Figure 4: Geographical distributions of the rainfall stations in and adjacent to the HKH region.

stations is shown in Fig. 4.

4. Methodology

The methodology for SRE validation is based on a literature review of validations that have been conducted for similar projects in other regions. The literature reviewed included the validation of satellite estimates for the United States, Africa, South America, and East Asia. In the Ethiopian study the CPC-RFE has been validated using relative percentage error and root mean square error techniques (Ouma et al. 2005). Same as in South America the satellite rainfall estimation has been validated using bias, correlation coefficient, root mean square error, POD, FAR and Skill techniques (Vila et al. 2003). The relationship between the rainfall estimation (RFE) and the in-situ rainfall records has been determined using visual graphical methods.

4.1. Independent rain-gauge data quality control

Nearly all the data came from the hydro-meteorological organizations of the partner countries, where the data were already screened to a certain extent. To ensure consistency, data were further subjected to rigorous quality control by geostatistical analysis in ArcGIS, as shown in Fig. 5. Point rain-gauge data were explored using the spatial data analysis tool. The visualization of rain-gauge data distribution and trends were obtained. Duplicates for each day in the period available were also removed. Rain-gauge data contributing to the GTS were identified and removed because the CPC-RFE2.0 algorithm uses daily GTS rain-gauge data as a primary input. Not removing these data points within the independent gauge dataset would lead to an inaccurate statistical validation with an elevated apparent accuracy. For each day, independent gauge data were compared to the GTS files, incorporated into each daily CPC-RFE2.0 product, and any matching independent gauge data removed. Datasets with less than 6 months of data were also removed to minimize the risk of generating temporal inhomogenities in the interpolated data due to varying station densities. Station information (especially location) were verified where details were available.

4.2. Classification of rainfall regime

The estimated rainfall and gauge-observed rainfall was compared by considering various factors. Since the HKH region is a large area with many spatial and temporal variations of precipitation, it is obvious that such an analysis

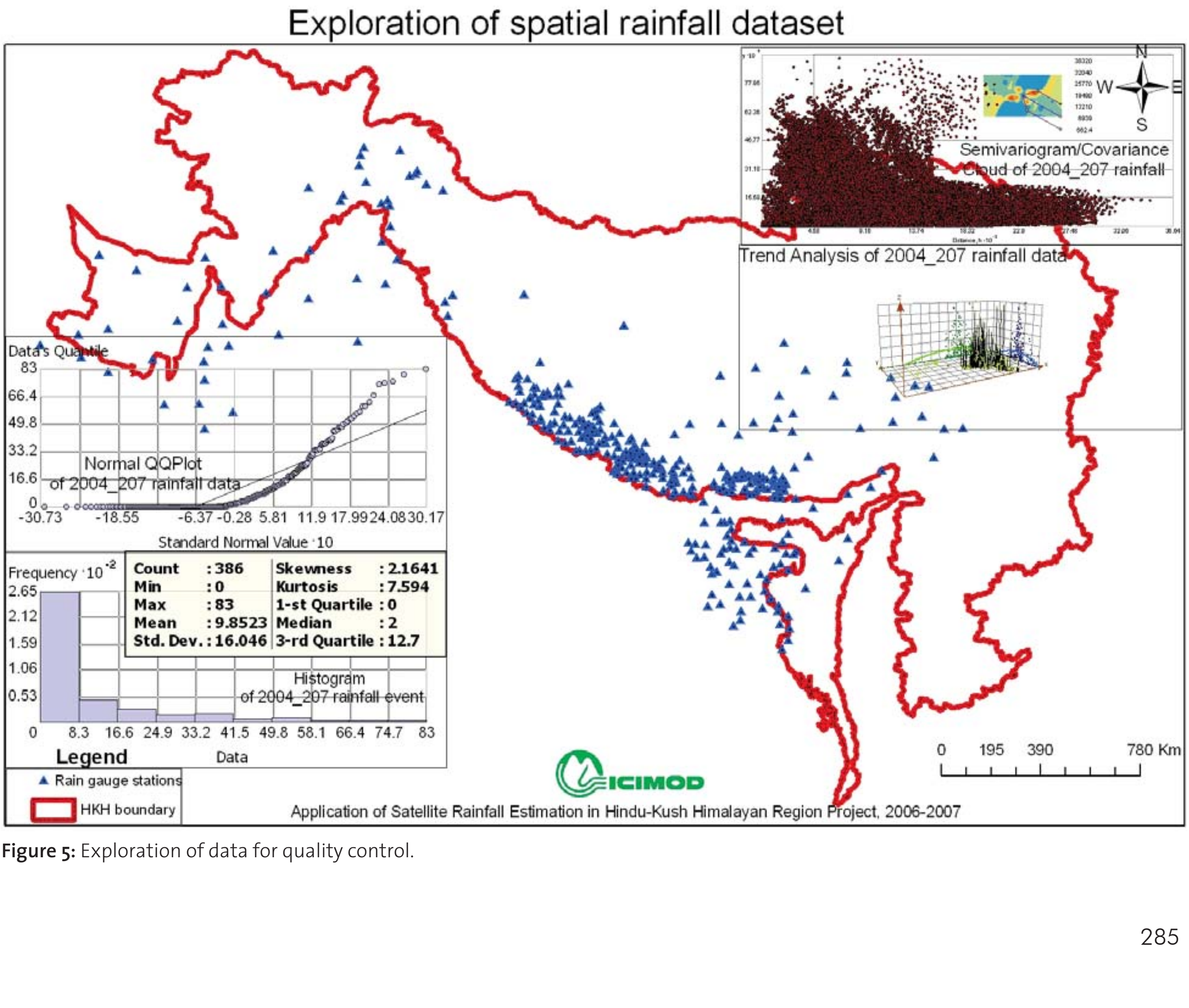

Figure 5: Exploration of data for quality control.

will provide poor results if it consider the region as one unit. Moreover, there is a lot of variation in network density in the region. As the HKH region is influenced differently by summer monsoon and western disturbances in the winter season, for subsequent validation the HKH region was divided into summer- and winter-monsoon dominated areas as shown in Fig. 6. The demarcation of the subregion was done on the basis of the seasonal precipitation charts prepared by IWMI (IWMI 2003).

4.3. Verification methods

In this study the statistical measures used to compare the satellite estimates with the rain-gauge data were taken from the results of the 3rd Algorithm Intercomparison Project of the Global Precipitation Climatology Project and from the book Measuring Precipitation from Space, Advances in Global Change Research 28 (Ebert 1996,2007). The spatial verification methods described here include visual verification, continuous statistics, and categorical statistics.

The verification methodology selected in this study was based on 24-hour accumulated rain-gauge data and satellite-estimated data of 0.1° x 0.1° spatial resolution.

4.3.1. Visual analysis

Visual verification compares maps of satellite estimates and observations. This is one of the most effective verification methods and is subjective. Gridded data of the observation (independent rain-gauge data) and estimated CPC_RFE2.0 data were remapped (overlaid) to the same projection with same colour scale to see the spatial distribution of rainfall (bias map).

4.3.2. Continuous verification statistics

Continuous verification statistics measure the accuracy of a continuous variable such as rain amount or intensity. These are the most commonly used statistics in validating satellite estimates. In the equations to follow Si indicates the satellite estimated value at grid cell or point i, Gi indicates the observed ground rain gauge value at grid cell or point i, and N is the number of observed samples (Ebert

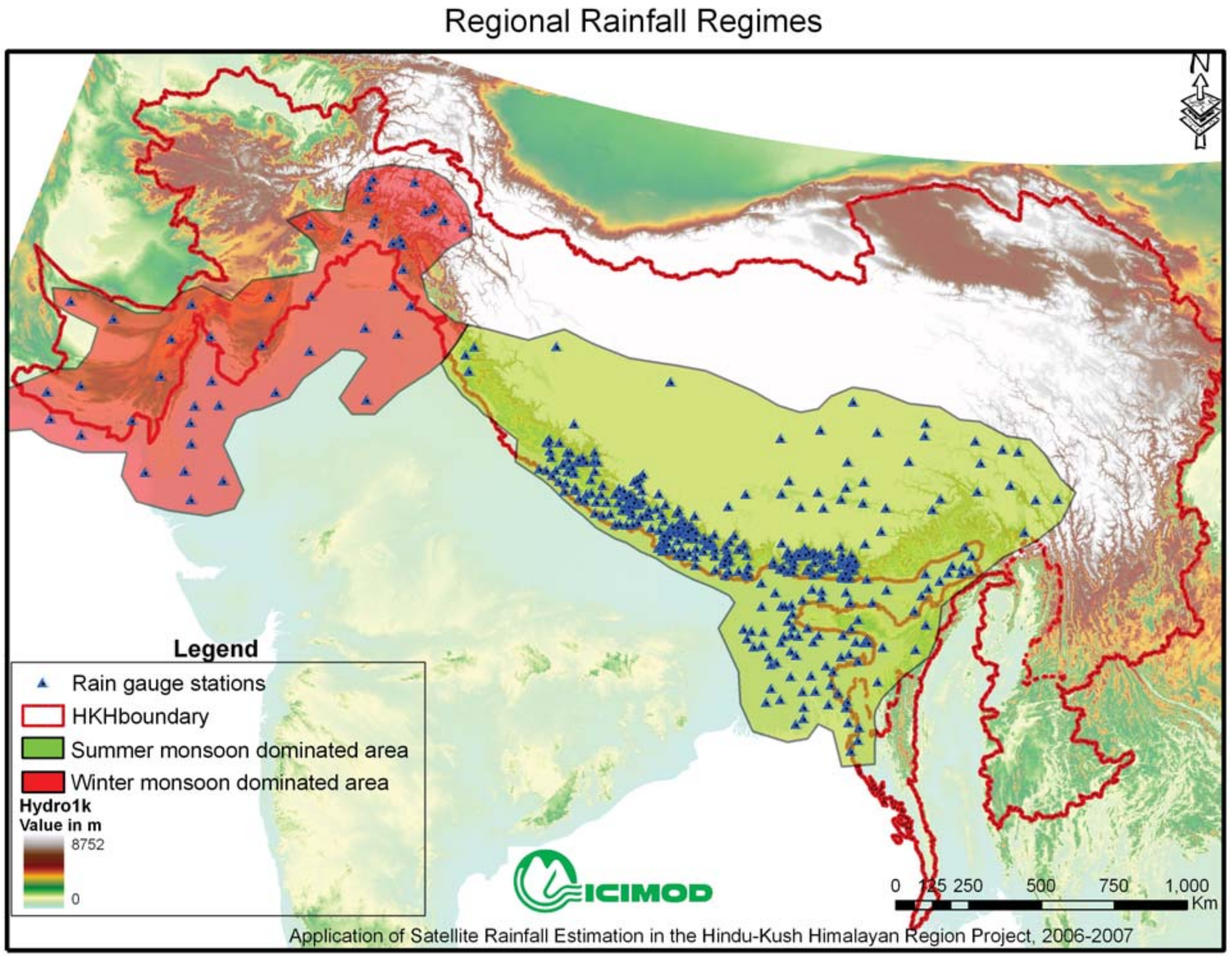

Figure 6: Classification of Hindu Kush-Himalayan region on the basis of summer and winter seasonal rainfall.

2007).

Mean error= $$\frac{1}{N}\sum_{i=1}^{N}(Si-Gi)$$

Mean absolute error= $$\frac{1}{N}\sum_{i=1}^{N}\left|Si-Gi\right|$$

Root mean square error= $$\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(Si-Gi\right)^{2}}$$

The correlation coefficient (r)=

$$\frac{\sum_{i=1}^{N}\left(Si-\overline{S}\right)\left(Gi-\overline{G}\right)}{\sqrt{\sum_{i=1}^{N}\left(Si-\overline{S}\right)^{2}}\sqrt{\sum_{i=1}^{N}\left(Gi-\overline{G}\right)^{2}}}$$

Percentage error (PE) =

$$\frac{Estimated-Observed}{Observed}\times 100\%$$

Skill = $$1-\frac{1}{N}\sum_{i=1}^{N}\frac{\left|Gi-Si\right|}{\left(Gi-Si\right)}$$

4.3.3. Categorical verification statistics

Categorical verification statistics measure the correspondence between the estimated and observed occurrence of events. Most are based on a 2 x 2 contingency table of yes/no events, such as rain/no rain, as shown in Tab. 2. The off-diagonal elements in the table characterize the error. The elements in the table (hits, misses, etc.) give the joint distribution of events, while the elements below and to the right (observed yes, observed no, etc.) are called the marginal distributions (Ebert 2007).

		Estimated Rainfall (SRE)	
		No Rain (No)	**Rain (Yes)**
Observed Rainfall	**No Rain (No)**	Q1 (Correct negatives)	Q2 (False alarms)
(ground rain gauge)	**Rain (Yes)**	Q3 (Misses)	Q4 (Hits)

Table 2: 2 x 2 Contingency table.

Probability of detection (POD): POD =

$$\frac{Q4}{Q3+Q4} \quad \text{or} \quad \frac{hits}{hits+misses} =$$

False alarm ratio (FAR): =

$$\frac{Q2}{Q2+Q4} \quad \text{or} \quad \frac{false\ alarms}{hits+false\ alarms}$$

5. Results

The main purpose of this study is to validate the CPC-RFE2.0 satellite rainfall estimates provided by Climatic Prediction Centre of NOAA for the HKH region so as to determine their operational viability whether it could be used in flood forecasting by feeding the product into Stream flow Model to know the precise discharge of a particular area and the preciseness of the algorithm according to sensitive of topographic influences, topography variation and topography effect. As it is concerning the flood scenario, the current validation has been conducted selecting heavy and light rainfall days with in summer monsoon period taken from the observed rain-gauge table of 2002, 2003, and 2004 but there is no exact threshold and demarcation to separate heavy and light amount rainfall; it is a subjective judgment based on different literature (Source: http://en.wikipedia.org/wiki/Rainfall).

The analysis was done by pixel to pixel comparison with the whole HKH region as one homogenous region and also by partitioning the region into the summer monsoon dominated area. Comparing RFE and observed rain-gauge data, qualitatively rainfall events generally match but quantitatively there are differences. No-rainfall days also match satisfactorily. The correlation is high in maximum rainfall days and low in minimum rainfall days. There is negative bias in maximum rainfall days and positive bias in minimum rainfall days. While in rain shadow areas of Trans-Himalaya RFE has a positive bias with overestimation of the rainfall. The results indicate that during intense rainfall, the CPC-RFE2.0 underestimates the rainfall with high negative percentage error. The categorical verification statistics show good rain/no-rain discrimination with POD above 0.8 in heavy rain falls FAR below 0.02. In low-rainfall days the POD is below 0.5. RMSE is high in maximum rainfall days and minimum in low rainfall days. All the analysis and results based on the summer monsoon dominated and HKH region as one homogeneous unit are shown in Fig. 7 and summarized in Tab. 3. Rainfall occurrence is underestimated by about half and more than half in monsoon during heavy and moderate rainfall. The results show that the CPC-RFE2.0 algorithm delineates the rainy area with less accuracy of rainfall intensity during heavy rainfall. This algorithm fails to capture monsoon depression, monsoon break, monsoon trough, and orographic conditions.

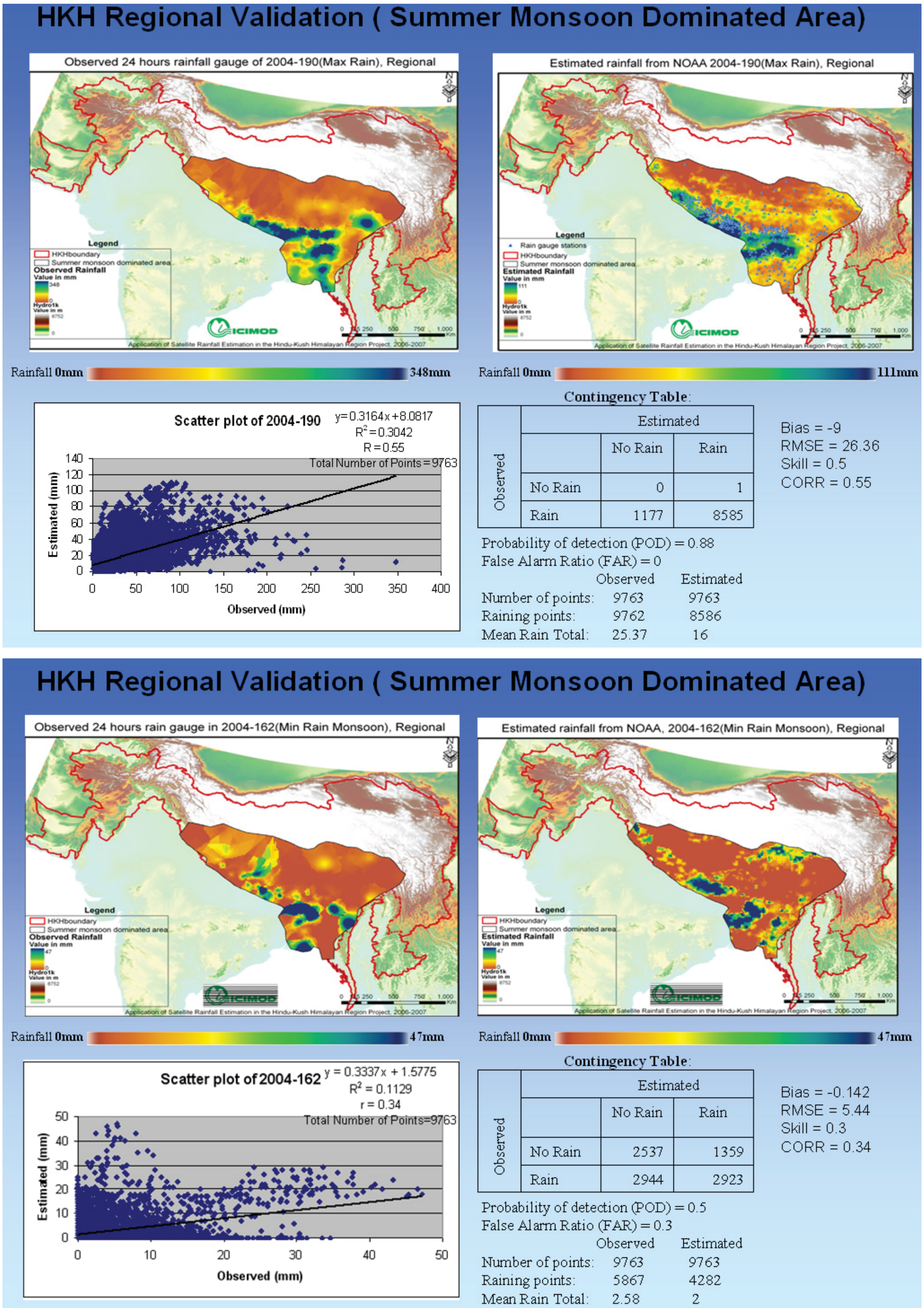

Figure 7: Regional validation maps of the summer-monsoon dominated area for 2004_190 and 2004_162.

		Continuous verification statistics					Categorical verification statistics	
Description	Pixel No	Bias (mm)	Correlation	RMSE (mm)	% error	Skill	POD	FAR
Summer-monsoon Dominated Area								
2002_203 (Heavy rain)	9466	-5.0	0.72	23.42	-24.27	0.85	0.80	0.02
2002_243 (Light rain)	9763	-0.74	0.01	5.62	-28.46	0.30	0.38	0.49
2003_190 (Heavy rain)	9763	-7.0	0.54	22.20	-40.65	0.98	0.85	0.01
2003_153 (Light rain)	9763	-0.62	0.01	5.96	-31.20	0.41	0.48	0.42
2004_190 (Heavy rain)	9763	-9.0	0.55	26.36	-35.47	0.50	0.88	0.00
2004_162 (Light rain)	9763	-0.14	0.34	5.44	-5.50	0.30	0.50	0.30

Table 3 : Statistical summary of regional validation.

6. Conclusion

Measuring rainfall from space appears to be an effective, viable means to estimate regional precipitation in the HKH, where sharing data still remains a challenge. These products form a backbone for hydrological and agricultural modelling, and the key platform to maintain routine observation along inaccessible areas. Though the CPC-RFE2.0 estimates are obtained after merging four data sources, including the GTS data from ground stations, very few stations from the GTS seem to have been included due to the timely availability of data from ground stations. This suggests that incorporating additional gauged ground-station data could yield better results, and hence further work should be carried out to enhance the RFE validation and improve the estimates.

The current validation has been conducted selecting heavy and light rainfall days for monsoon season. In general the results indicate the CPC-RFE2.0 provides reasonable rainfall estimates over the HKH region but needs to be improved before it can be implemented for operational flood forecasting in the region. In the analysis 373 raingauge stations were used to conduct summer monsoon dominated area provided by Bangladesh, Bhutan, China, India and Nepal and partitioning the area according to monsoon influence. The quality of the estimates was assessed by comparing bias, Root Mean Square Error (RMSE), correlation coefficient, probability of detection (POD) and False Alarm Ratio (FAR) values for each set of validation data. Some error statistics showed that the RFE estimates were better when the whole region was divided into the winter and summer monsoon dominated areas. In the summer monsoon dominated areas the RFE estimates were low compared to the observed indicating a negative bias. There is a need to investigate gaps and shortfalls particularly related to mountain orographic processes for further applications and usefulness in the region.

Acknowledgements

Author is grateful to the USAID/OFDA for funding this work. My special thanks go to A. Sezin Tokar, Hydrometeorological Hazard Advisor, USAID/OFDA for her help and valuable suggestions. I would like to express gratefulness to Dr. Ralph Ferraro, Dr. Daniel A. Vila and Dr. Chris Kidd for their valuable suggestion, comments, guidance and special thanks go to Dr. F. Joseph Turk from Naval Research Laboratory, Monterey, USA for wonderful book. The work would not have been possible without the active support and participation of the partner countries. Therefore I would like to thank all the partners from Bangladesh, Bhutan, China, India, Nepal, and Pakistan for providing data and for their valuable contributions. Last but not least my heartfelt thanks go to Mr. Basanta Shrestha, Head, Mountain Environmental Natural Resources information system (MENRIS) for his encouragement and continuous support on every step of this work and given an opportunity to work on this field.

References

EBERT, EE., 1996: Result of the 3rd Algorithm Intercomparison Project (AIP-3) of the Global Precipitation Climatology Project (GPCP). Bureau of Meteorology Centre, Report No 55, Melbourne, Australia.

EBERT, EE., 2007: Methods for verifying satellite precipitation estimates Measuring precipitation from space: EURAINSAT and the future, Advances in Global change research vol28, pp.345-356.

IWMI, 2003: World Water and Climate Atlas

KAMARIANAKIS, Y., FEIDAS, H., KOKOLATOS, G., CHRYSOULAKIS, N., KARATZIAS, V.: Validating remotely sensed rainfall e weighted regression.

KAMARIANAKIS, Y., FEIDAS, H., KOKOLATOS, G., CHRYSOULAKIS, N., KARATZIAS, V., 2006: Comparing rainfall estimates derived from rain gauges and satellite images at the eastern Mediterranean region. Paper presented at the 9th AGILE Conference on Geographic Information Science, Visegrad, Hungary,pp 15-20.

KIDD C., 2004: Validation of satellite rainfall estimates over the mid-latitudes. 2nd IPWG working group workshop, Monterey, CA. 25-28 October 2004, pp205-215.

OUMA, GO., CHANE, MUTHAMA, NJ., 2005: Validation of satellite derived rainfall estimates: the Ethiopian case study. Paper presented at the 7th Kenya Meteorological SocietyWorkshop on Meteorological Research, Applications and Services, 17-21 Ocotber 2005, Nairobi, Kenya.

SHRESTHA, MS., TAKARA, 2007: Impacts of Floods in South Asia. Paper presented at the meeting of the Asia and Oceanic Geological Society (AOGS), Bangkok, 31 July – 3 August, 2007.

VILA, D., INES, V., 2002: Some experiences on satellite rainfall estimation over South America. Proceeding of 1st International Precipitation working Group Workshop, Madrid, Spain, 23-27.

VILA, D., SCOFIED, R., KULIGOWSKI, R., DAVENPORT, J., 2003: Satellite Rainfall Estimation Over South America: Evaluation Of Two Major Events. NOAA Technical Report NESDIS Nº 114, U.S. Department of Commerce.

VILA, D., LIMA, A., 2004: Satellite Rainfall Estimation Over South America: The Hydroestimator Technique. Anales del 14th International Conference on Clouds and Precipitation, Bologna, Italia.

XIE, P., ARKIN, PA., 1996: Analysis of global monthly precipitation using gauge observations, satellite estimates and numerical model predictions. Journal of Climate 9: 840-858.

XU, J., SHRESTHA, A., VAIDYA, R., ERIKSSON, M., HEWITT, K., 2007: The Melting Himalayas Regional Challenges and Local Impacts of Climate Change on Mountain Ecosystems and Livelihoods. ICIMOD Technical paper: The Melting Himalaya, pp.1-12.

Correspondence to:

SAGAR RATNA BAJRACHARYA

International Centre for Integrated Mountain Development (ICIMOD)

G.P.O. Box 3226, Kathmandu, Nepal

e-mail: sagbajracharya@icimod.org