

## STUDY OF GPM-IMERG RAINFALL DATA PRODUCT FOR GANGOTRI GLACIER

Prateek Verma<sup>1,\*</sup>, Sanjay Kumar Ghosh<sup>1</sup>

<sup>1</sup>Geomatics Engineering Group, Department of Civil Engineering, IIT Roorkee, India - vermprateek@gmail.com

Commission V, SS: Natural Resources Management

**KEYWORDS:** GPM constellation satellites, IMERG data, Gangotri glacier, Satellite precipitation estimates, Remote Sensing

### ABSTRACT:

This study presents a comparison of new generation weather observatory satellites Global Precipitation Measurement (GPM) Integrated Multi-satellite Retrievals for GPM (IMERG) rainfall products with field data collected for Gangotri glacier in India. The meteorological analysis of rainfall estimates has been performed on GPM IMERG Final, Late and Early precipitation products available at daily scale with a spatial resolution of  $0.1^\circ \times 0.1^\circ$  for melting season from May to September for the year 2014 and 2015 respectively. The comparison of satellite products with field data was done using correlation coefficient and standard anomaly. The Late run curve showed a high degree of similarity with final run curve while early run showed variation from them. The satellite meteorological data correctly identified non-rainy days with an average of ~86.7%, ~67.5% and ~95% for pre-monsoon, monsoon and post-monsoon season respectively. The rmse for final run data product for 2014 and 2015 are 4.5, 1.23, 1.55, 1.24, 0.8 and 1.14, 7.1, 1.82, 1.15, 1.52 from May to September respectively. Overall, it has been observed that for medium to heavy rainfall final run estimates are close to field data and for light to medium rainfall late run estimates are close. Similar results have been obtained from both datasets for non-rainy days in the study area.

### 1. INTRODUCTION

Precipitation is considered as one of the major driver of land surface hydrology and the most uncertain part of hydrologic cycle as well (Futrell et al., 2005). The availability of freshwater is directly affected by the distribution of precipitation in time and space (Hou et al., 2014). There are challenges in monitoring, modelling and prediction of precipitation as precipitation varies at both temporal and spatial scales (Futrell et al., 2005). The occurrence of extreme precipitation events such as floods, droughts, landslides and hurricanes have significant impact on socio-economic aspect of the region (NRC, 2010). The introduction of satellite remote sensing for study of precipitation and its patterns, has helped in better understanding of weather and climate related research. Tropical Rainfall Measuring Mission (TRMM) mission launched in November 1997 was one such mission where global tropical rainfall dataset was delivered that helped in improved knowledge of cyclone structure and evolution, climate and weather modeling, lightning-storm relationships.

Bookhagen and Burbank (2006), used TRMM precipitation dataset with a spatial resolution of  $5 \text{ km} \times 5 \text{ km}$  for 8 years from 1998 to 2005 to identify spatial distribution of rainfall and relationship between topography, relief and rainfall locations for Himalayas. Analysis shows that there are two discrete bands of high rainfall stretching along the length of the Himalayas at a mean relief of about 1.2 km and 2 km respectively. The results reveal a continuous band of high rainfall near the toe of Greater Himalayas and near the junction of Lesser Himalayas and Indo-Gangetic foreland basin.

Wu et al. (2014), studied the use of real time satellite based precipitation dataset in Global Flood Monitoring System (GFMS) to derive flood-monitoring parameters at a spatial and temporal

resolution of about ~12 km and 3 hours respectively. The accuracy of GFMS was evaluated by running the Dominant River tracing–Routing Integrated with Variable Infiltration Capacity Environment (DRIVE) model for 15 years. The satellite precipitation product was derived, calibrated and tested for streamflow estimation for 1121 river gauges across quasi-global domain using TRMM Multi-satellite Precipitation Analysis (TMPA) dataset. Less accuracy in results was reported for higher latitudes due to larger errors in satellite precipitation input.

Sharifi et al. (2016), compares GPM-IMERG final run product, Era-Interim product from European Centre for Medium Range Weather Forecasts (ECMWF) and post real time TRMM and Multi-satellite Precipitation Analysis (TMPA-3B42) against daily precipitation data available through Iran Meteorological Organization agency. The comparison of various satellite rainfall products with meteorological dataset was done for four different locations in Iran covering variable topography and climatic conditions for a period of one year from March 2014 to February 2015. The results indicate better performance of IMERG product in comparison to other rainfall datasets on a daily scale but all three products underestimates precipitation in the study area. It is reported that correlation coefficient between meteorological data and IMERG product is far better to that of TMPA-3B42 and Era-Interim product on a daily scale. The IMERG product was found superior to other products in detection of heavy precipitation ( $>15\text{mm/day}$ ) for all sites.

Sungmin et al. (2017), compared GPM-IMERG version 3 early, late and final run half-hourly rainfall estimates with WegenerNet Feldbach region (WEGN) gauge stations gridded rainfall dataset in Austria. The complete dataset was divided into two seasons (hot and warm) and two rainfall intensity range (high and low) with

---

\* Corresponding author

extended summer season from April to October. The final run product was found to be in best overall agreement with gauge data followed by late and early run products.

The core satellites of GPM has been jointly developed by National Aeronautics and Space Administration (NASA) and Japan Aerospace Exploration Agency (JAXA) and launched on February 28, 2014 at 3:07 am Japan Standard Time. The GPM project is an international mission to measure precipitation at global scale with more accuracy and frequency. With the cooperation of multiple constellation satellites, it will be possible to measure global precipitation about every three hours. GPM constellation is a follow up mission of TRMM satellites, which is designed for significant improvement in observation of rain intensity and its distribution worldwide. The sensors onboard GPM satellites will include the measurement of light intensity precipitation ( $<0.5$  mm/h) and falling snow (Hou et al., 2014).

The other focus of development of this mission is to provide data to user organizations at near real time speed for disaster mitigation such as prediction of flooding, typhoon forecasting and improvement in accuracy of numerical weather prediction (Nio et al., 2014). GPM-IMERG data is produced twice in near-real time mode called as IMERG-Early (with 6 hours latency time) and IMERG-Late (with 18 hours latency time). The late run product provides the advantage of allowing lagging data transmissions that were not available in early run product. The final run product is generated after two and half months of satellite observation with improved accuracy by using climatological adjustment that incorporates gauge data.

This study aims at the assessment of accuracy of different precipitation products available through GPM-IMERG data for a mountainous region present in Indian Himalayan region. The Himalayan region has insufficient spatial and temporal coverage of meteorological sites, which has resulted in limited knowledge about this region. The introduction of satellite remote sensing in the field of precipitation measurement has helped in better understanding of precipitation pattern in Himalayan region.

## 2. STUDY AREA

The study has been carried out for Gangotri glacier, which is one of the largest glaciers of the Himalayas (Verma and Ghosh, 2018). The glacier is situated between  $30^{\circ}43'N$  -  $31^{\circ}01'N$  latitudes and  $79^{\circ}0'E$  -  $79^{\circ}17'E$  longitudes and is the source of River Bhagirathi (Singh et al., 2010). The Gangotri basin is a cluster of many glaciers of which Gangotri glacier is the main trunk. The length and width of Gangotri glacier is 30.20 km and 0.20 – 2.35 km respectively with a total catchment area of about  $556 \text{ km}^2$  (Singh et al., 2010). The location of Gangotri glacier has been shown in Fig. 1 with a star in yellow represents meteorological station located 4 km downstream from snout of the glacier.

## 3. DATA USED

GPM is a constellation based satellite mission that carries Dual-frequency Precipitation Radar (DPR), which is jointly developed by National Institute of Information and Communications Technology (NICT) and JAXA and GPM Microwave Imager (GMI) developed by NASA. The advantage of GPM-IMERG sensor with respect to other precipitation satellites is the

measurement of light rain and snowfall along with the heavy rain events (Sharifi et al., 2016). The satellite precipitation dataset is available for download since mid-March 2014 at NASA Goddard Earth Sciences Data and Information Services Center website (GES DISC) at daily time scale. The current version of GPM-IMERG precipitation dataset is Version 5, which is at a spatial resolution of  $0.1^{\circ} \times 0.1^{\circ}$  and a processing level of 3 (Huffman, 2017). The precipitation data is provided for public use at no cost in netcdf format and has been converted into readable format using ArcGIS.

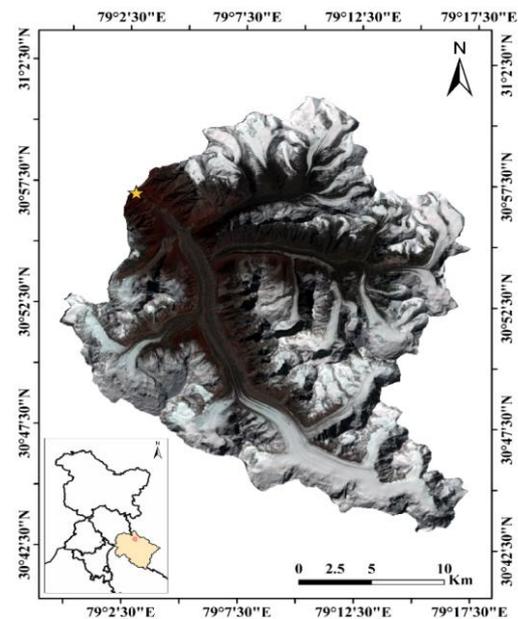


Figure 1. Satellite image of the study area

The observed rainfall data has been acquired by National Institute of Hydrology, Roorkee, India on a daily temporal scale for the melt season of the glacier. The analysis of GPM-IMERG precipitation data and field data has been done for the common time period between the two datasets i.e. for 2014 and 2015 respectively.

## 4. METHODOLOGY

The meteorological analysis of rainfall estimates has been performed on datasets provided by GPM satellites namely GPM IMERG Final, Late and Early Precipitation available at daily scale with a spatial resolution of  $0.1^{\circ} \times 0.1^{\circ}$ . The daily estimate of accumulated precipitation has been derived from the half hourly GPM precipitation data. The daily data is generated by summing valid precipitation retrievals of the day from GPM-IMERG and is given in units of mm. The detailed description of generation of GPM-IMERG satellite precipitation products can be found in Huffman et al., (2015). The precipitation data available in netcdf format is converted into raster format using ArcGIS environment. The months of May and June, July and August, September has been considered under pre-monsoon, monsoon and post-monsoon season respectively. The rainfall events have been classified as light, medium and heavy based on intensity of rainfall for a particular day (Table 1). The methodology followed to achieve results for the desired objective has been shown in Fig. 2.

#### 4.1 Correlation Analysis

It is an important tool to identify degree of association between two datasets and the parameter to observe the degree of closeness between two datasets is called as Pearson correlation coefficient ( $r_p$ ). It is defined as:

$$r_p = \frac{[N \sum xy - (\sum x)(\sum y)]}{\sqrt{[N \sum x^2 - (\sum x)^2][N \sum y^2 - (\sum y)^2]}} \quad (1)$$

where  $\sum x$  is the sum of x scores;  $\sum x^2$  is the sum of squared x scores;  $\sum y$  is the sum of y scores;  $\sum y^2$  is the sum of squared y scores;  $\sum xy$  is the sum of product of paired scores and N is the number of pairs of scores. The value of  $r_p$  ranges between -1 to +1 where positive value indicates positive correlation between two datasets and negative value indicates negative correlation.

Classes	Rainfall range/day (mm)
Light	< 5
Medium	5 - 10
Heavy	> 10

Table 1. Classification of rainfall events

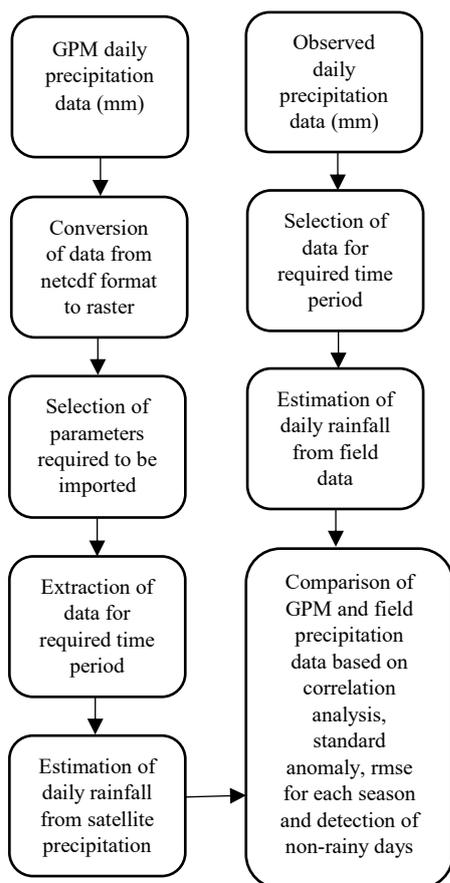


Figure 2. Flowchart of the methodology followed

#### 4.2 Standard Anomaly

The comparison of two datasets can be done using standard anomaly, which is effective in removal of the influence of location in dataset and its spread. It is obtained by dividing the anomaly by its corresponding standard deviation ( $S_x$ ) that is given by (Piyooosh & Ghosh, 2016):

$$z = (x_i - \bar{x}) / S_x \quad (2)$$

where  $x_i - \bar{x}$  is the anomaly, which is the difference between value and its mean in a dataset;  $x_i$  is the value in the dataset and  $\bar{x}$  is the mean of time series dataset.

### 5. RESULTS AND DISCUSSION

#### 5.1 Analysis of GPM-IMERG Early, Late and Final Precipitation product

The three different precipitation products available through GPM-IMERG satellites have been analyzed using correlation coefficient to determine degree of similarity between them. The results reveal the high degree of similarity between Early and Late precipitation products except for pre-monsoon season of 2014 (Table 2). The high value of  $r_p$  indicates that both the products are similar to each other and would provide similar results.

	Pre-monsoon	Monsoon	Post-monsoon
<b>2014</b>	0.51	0.96	0.83
<b>2015</b>	0.96	0.94	0.99

Table 2 Correlation coefficient between GPM-IMERG Early and Late precipitation product

The correlation coefficient has also been estimated for different satellite precipitation products with respect to NIH field data for identification of product that are similar to field data. The high values of  $r_p$  that are marked in bold in Table 3 indicates the closeness of precipitation product with field data for that particular season. As per results in Table 2, late and final run precipitation products show the highest values of  $r_p$  while the early run product shows the highest variation from field data among the three available products. Due to these reasons, early precipitation product has been neglected while late and final run precipitation products have been utilized for further analysis.

	2015			2014		
	Pre-monsoon	Monsoon	Post-monsoon	Pre-monsoon	Monsoon	Post-monsoon
Early	0.54	0.49	0.25	0.50	0.73	<b>0.34</b>
Late	0.61	<b>0.53</b>	<b>0.26</b>	<b>0.51</b>	<b>0.75</b>	<b>0.34</b>
Final	<b>0.67</b>	<b>0.53</b>	<b>0.26</b>	0.50	<b>0.75</b>	<b>0.34</b>

Table 3 Correlation coefficient of GPM-IMERG Early, Late and Final precipitation product with respect to NIH field data

## 5.2 Identification of non-rainy days

The GPM-IMERG data was downloaded and analyzed for the common time period with field data collected from NIH Roorkee dataset at daily interval. The starting date for observed field data in 2014 and 2015 is 20<sup>th</sup> May and 13<sup>th</sup> May respectively, while the final date is common for both the years i.e. 30<sup>th</sup> September. The satellite precipitation dataset showed close agreement with field data in identification of non-rainy days for the melt seasons of 2014 and 2015 respectively (Fig. 3). The satellite meteorological data correctly identified non-rainy days with an average of ~86.7%, ~67.5% and ~95% for pre-monsoon, monsoon and post-monsoon season respectively. The close result between both datasets shows the ability of GPM-IMERG satellites to capture days with light rainfall or no rainfall in mountainous region of Himalayas.

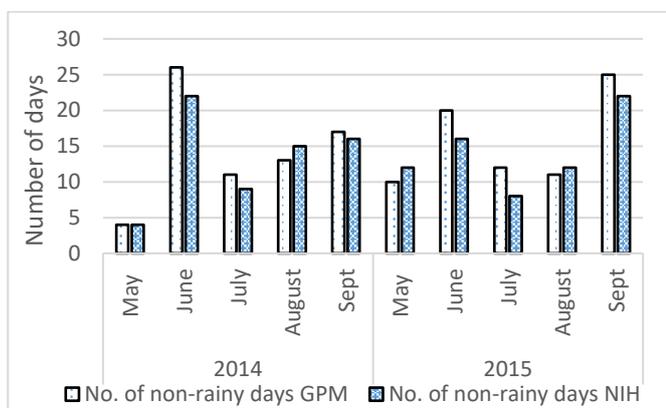


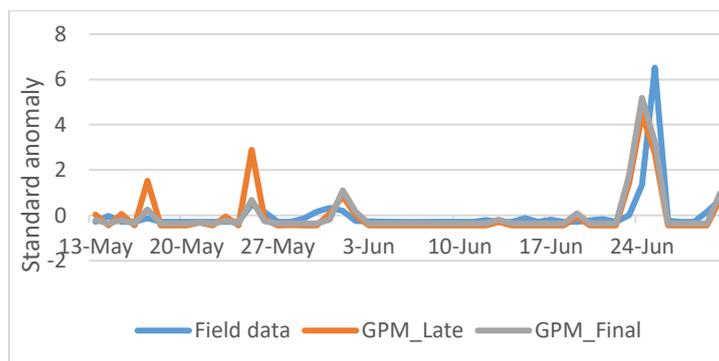
Figure 3. Comparison of GPM-IMERG final run and NIH precipitation data for the identification of non-rainy days

## 5.3 Comparison of GPM-IMERG Late and Final run product with NIH field data

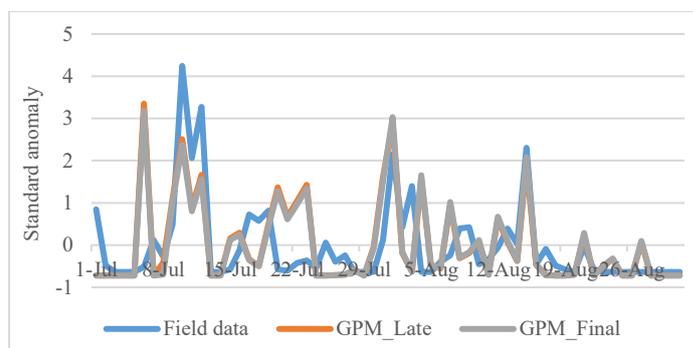
The comparison of GPM-IMERG late and final run data with NIH field data has been done at daily scale for pre-monsoon, monsoon and post-monsoon seasons for 2014 and 2015 respectively (Fig. 4 and 5). The results reveal that late run curve closely follows final run curve in all the seasons for both years except for some events. It is observed in Fig. 4(a) that on 17<sup>th</sup> May and 25<sup>th</sup> May 2015 late run curve shows a lot of variation from final run and field data with overestimation of the rainfall, while on 25<sup>th</sup> June both the satellite products underestimate the heavy rainfall event.

For the monsoon season of 2015, late run curve completely follows the final run curve but shows variation from field data at few events e.g. 6<sup>th</sup>, 21<sup>st</sup>, 22<sup>nd</sup>, 23<sup>rd</sup> July and 4<sup>th</sup> August 2015. Both the satellite products overestimate the light rainfall event on these days as evident from field data. The late run underestimates and final run overestimates the heavy rainfall event on 10<sup>th</sup>, 11<sup>th</sup> and 12<sup>th</sup> July 2015 as shown in Fig. 4(b). The curves in post-monsoon season of 2015 shows variation on 2 days i.e. 18<sup>th</sup> and 22<sup>nd</sup> September (Fig. 4c). The field data shows 18<sup>th</sup> September as heavy rainfall event and 22<sup>nd</sup> September as light rainfall event. The satellite products predict opposite values on both days by underestimating the heavy rainfall day and overestimating light rainfall day.

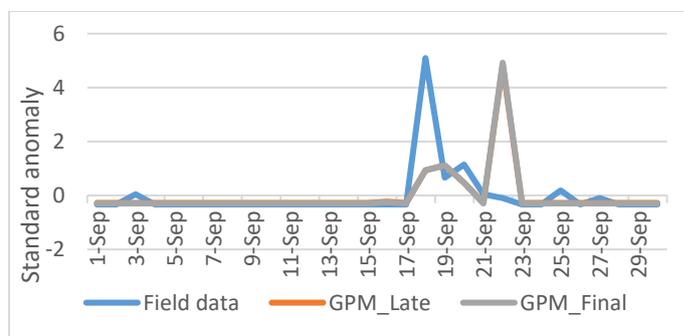
The late and final run curves show variation from field data on 14<sup>th</sup>, 18<sup>th</sup> and 25<sup>th</sup> June in pre-monsoon season of 2014 as shown in Fig. 5(a). The field data shows light rainfall days on first two dates and medium intensity rainfall on third date, while satellite data underestimates the rainfall on first date, overestimates on second date and shows no rainfall at all on third day. The curves of satellite products show difference from field data on few occasions in monsoon season of 2014 as shown in Fig. 5(b) i.e. 22<sup>nd</sup> July, 2<sup>nd</sup>, 4<sup>th</sup> and 6<sup>th</sup> August. The field data reveals light rainfall events on 22<sup>nd</sup> July and 6<sup>th</sup> August while satellite products show heavy rainfall events on these dates. The heavy rainfall events are shown on 2<sup>nd</sup> and 4<sup>th</sup> August by field data but satellite products underestimate on both these dates.



(a)

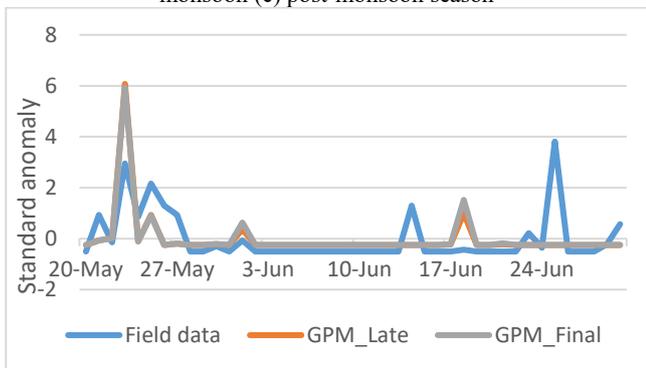


(b)

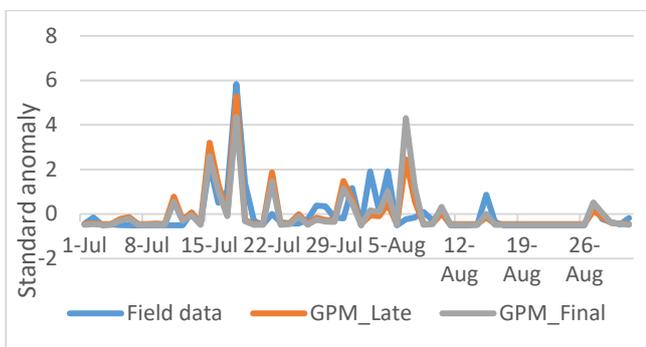


(c)

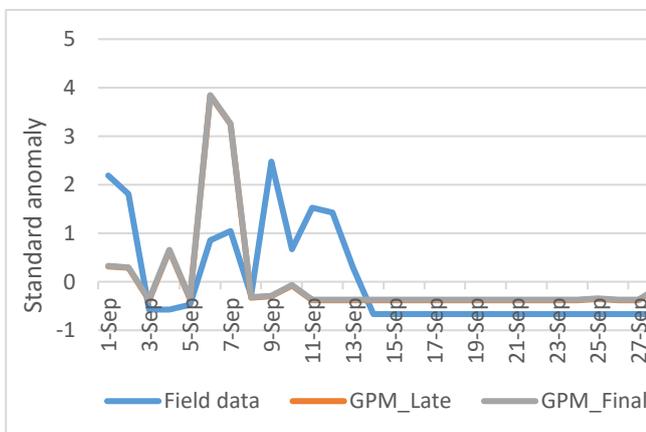
Figure 4. Comparison of NIH field data with GPM-IMERG late and final run precipitation data for 2015 (a) pre-monsoon (b) monsoon (c) post-monsoon season



(a)



(b)



(c)

Figure 5. Comparison of NIH field data with GPM-IMERG late and final run precipitation data for 2014 (a) pre-monsoon (b) monsoon (c) post-monsoon season

The post-monsoon season of 2014 shows variation on first few days of the month between field data and satellite rainfall products as evident from Fig. 5(c). The satellite products overestimates rainfall intensity on 4<sup>th</sup>, 6<sup>th</sup> and 7<sup>th</sup> September while underestimates the rainfall on 1<sup>st</sup>, 2<sup>nd</sup>, 9<sup>th</sup>, 10<sup>th</sup>, 11<sup>th</sup> and 12<sup>th</sup> September. The assessment of final run satellite precipitation product with respect to field data was done by estimating root mean square error (rmse) as shown in Table 4. The low rmse values indicate closeness of

final run product with field data after removing the days that are predicting large over or underestimation of rainfall events. The month of May in 2014 and June in 2015 shows a large rmse value indicating greater variation from field data.

	May	June	July	August	September
2014	4.5	1.23	1.55	1.24	0.8
2015	1.14	7.1	1.82	1.15	1.52

Table 4 Monthly values of rmse between final run product and field data

## 6. CONCLUSION

In this study, the GPM-IMERG satellite precipitation products were evaluated at daily scale using field data from meteorological station for Gangotri glacier in Himalayas. The satellite products available in the form of Early, Late and Final run were assessed for 2 melt seasons of 2014 and 2015. The results reveal daily curves of final and late run datasets are showing more closeness to field data as compared to early run product. The late run curve closely follows final run curve with some variation on few occasions. Since, the final run satellite product uses gauge data to provide research level products but the necessity of field data is felt to assess the accuracy of the product. There is need of improvement in satellite precipitation products for areas with variable topography and complex weather system. Overall, it has been observed that for medium to heavy rainfall final run data product estimates are close to field data and for less to medium rainfall late run data product estimates are good. Similar results have been obtained from both datasets for non-rainy days in the study area.

## ACKNOWLEDGEMENT

The authors would like to thank Director (NIH Roorkee), Dr. Manohar Arora (Head of Snow and Glacier Laboratory, NIH Roorkee) and all members of NIH for sharing their field observatory data taken under extreme climatic conditions. The authors would also like to thank the team of GPM constellation satellites NASA and JAXA for providing satellite precipitation data under open source policy. The first author would also like to thank MHRD, Government of India for providing necessary financial help for carrying out the research.

## REFERENCES

- Bookhagen, B. and Burbank, D.W., 2006. Topography, relief, and TRMM-derived rainfall variations along the Himalaya. *Geophysical Research Letters*, 33(8).
- Futrell, J.H., Gephart, R.E., Kabat-Lensch, E., McKnight, D.M., Pyrtle, A., Schimel, J.P., Smyth, R.L., Skole, D.L. and Gephart, J.M., 2005. *Water: Challenges at the intersection of human and natural systems* (No. PNWD-3597). PNNL (Pacific Northwest National Laboratory (PNNL), Richland, WA (United States)).
- Huffman, G., 2017. *GPM IMERG Final Precipitation L3 1 day 0.1 degree x 0.1 degree V05*, Edited by Andrey Savtchenko, Greenbelt,

- MD, Goddard Earth Sciences Data and Information Services Center (GES DISC), Accessed: July 2018, 10.5067/GPM/IMERGDF/DAY/05. precipitation and a coupled land surface and routing model. *Water Resources Research*, 50(3), pp.2693-2717.
- Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, K., Joyce, R., Kidd, C., Nelkin, E. J., and Xie, P., 2015. Algorithm Theoretical Basis Document (ATBD) Version 4.5. NASA Global Precipitation Measurement (GPM) Integrated Multi-satellite Retrievals for GPM (IMERG), [https://pmm.nasa.gov/sites/default/files/document\\_files/IMERG\\_ATBD\\_V-4-5.pdf](https://pmm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V-4-5.pdf).
- Hou, A.Y., Kakar, R.K., Neeck, S., Azarbarzin, A.A., Kummerow, C.D., Kojima, M., Oki, R., Nakamura, K. and Iguchi, T., 2014. The global precipitation measurement mission. *Bulletin of the American Meteorological Society*, 95(5), pp.701-722.
- National Research Council, 2010. *When weather matters: Science and services to meet critical societal needs*. National Academies Press.
- Nio, T., Takeshima, T., Fuda, M., Fujita, Y., Saito, N. and Konishi, T., 2014, July. JAXA GPM mission operation system overview and GPM data distribution. In *Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International* (pp. 3746-3749). IEEE.
- Piyoosh, A.K. and Ghosh, S.K., 2016. A comparative assessment of temperature data from different sources for Dehradun, Uttarakhand, India. *Journal of Meteorological Research*, 30(6), pp.1019-1032.
- Sharifi, E., Steinacker, R. and Saghafian, B., 2016. Assessment of GPM-IMERG and other precipitation products against gauge data under different topographic and climatic conditions in Iran: Preliminary results. *Remote Sensing*, 8(2), p.135.
- Singh, P., Haritashya, U.K. and Kumar, N., 2008. Modelling and estimation of different components of streamflow for Gangotri Glacier basin, Himalayas/Modélisation et estimation des différentes composantes de l'écoulement fluvial du bassin du Glacier Gangotri, Himalaya. *Hydrological sciences journal*, 53(2), pp.309-322.
- Singh, P., Haritashya, U.K. and Kumar, N., 2007. Meteorological study for Gangotri Glacier and its comparison with other high altitude meteorological stations in central Himalayan region. *Hydrology Research*, 38(1), pp.59-77.
- Sungmin, O., Foelsche, U., Kirchengast, G., Fuchsberger, J., Tan, J. and Petersen, W.A., 2017. Evaluation of GPM IMERG Early, Late, and Final rainfall estimates using WegenerNet gauge data in southeastern Austria. *Hydrology and Earth System Sciences*, 21(12), pp.6559-6572.
- Verma, P. and Kumar Ghosh, S., 2018. Classification of glacial lakes using integrated approach of DFPS technique and gradient analysis using Sentinel 2A data. *Geocarto International*, pp.1-14.
- Wu, H., Adler, R.F., Tian, Y., Huffman, G.J., Li, H. and Wang, J., 2014. Real-time global flood estimation using satellite-based