Precipitation Retrieval over the Tibetan Plateau from the Geostationary Orbit—Part 2: Precipitation Rates with Elektro-L2 and Insat-3D

Christine Kolbe [1],* , **Boris Thies [1]** , **Nazli Turini [1]** , **Zhiyu Liu [2]** and **Jörg Bendix [1]**

[1] Department of Geography, Laboratory for Climatology and Remote Sensing, Deutschhausstrasse 12, Philipps-Universität Marburg, 35032 Marburg, Germany; boris.thies@geo.uni-marburg.de (B.T.); nazli.turini@geo.uni-marburg.de (N.T.); bendix@mailer.uni-marburg.de (J.B.)

[2] College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China; liuzhiyu2019@mwr.gov.cn

* Correspondence: christine.kolbe@geo.uni-marburg.de; Tel.: +49-(0)6421-28-24270

Abstract: We present the new Precipitation REtrieval covering the TIbetan Plateau (PRETIP) as a feasibility study using the two geostationary (GEO) satellites Elektro-L2 and Insat-3D with reference to the GPM (Global Precipitation Measurement Mission) IMERG (Integrated Multi-satellitE Retrievals for GPM) product. The present study deals with the assignment of the rainfall rate. For precipitation rate assignment, the best-quality precipitation estimates from the gauge calibrated microwave (MW) within the IMERG product were combined with the GEO data by Random Forest (RF) regression. PRETIP was validated with independent MW precipitation information not considered for model training and revealed a good performance on 30 min and 11 km spatio-temporal resolution with a correlation coefficient of R = 0.59 and outperforms the validation of the independent MW precipitation with IMERG's IR only product (R = 0.18). A comparison of PRETIP precipitation rates in 4 km resolution with daily rain gauge measurements from the Chinese Ministry of Water Resources revealed a correlation of R = 0.49. No differences in the performance of PRETIP for various elevation ranges or between the rainy (July, August) and the dry (May, September) season could be found.

Keywords: precipitation retrieval; Insat-3D; Elektro-L2; GPM IMERG; machine learning; random forest

1. Introduction

Precipitation is an essential key variable in the global hydrological cycle. Obtaining dense information about precipitation from gauge observations and radar which would allow an exact quantification of precipitation remains a challenge, especially in remote and high elevated regions, such as the Tibetan Plateau (TiP) [1,2]. Satellite-based precipitation retrievals can provide such information in a high spatial and temporal resolution.

Well known precipitation products are PERSIANN (Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks), CMORPH (Climate Prediction Center MORPHing technique), TMPA (TRMM Multi-satellite Precipitation Analysis) based on TRMM (Tropical Rainfall Measuring Mission), and IMERG (Integrated Multi-satellitE Retrievals for GPM) based on GPM (Global Precipitation Measurement Mission) [3–6]. They retrieve precipitation using single or multispectral data from active or passive instruments onboard of LEO (Low Earth Orbit) and/or GEO (Geostationary Orbit) satellites. The data provided by the satellites are processed with various algorithms that all have specific advantages and disadvantages. All of them are already evaluated over the TiP. Bai & Liu [7] compared five satellited based precipitation products (Climate

Hazards Group InfraRed Precipitation with Station data (CHIRPS), CMORPH, PERSIANN climate data record (PERSIANN-CDR), TMPA 3B42, and Multi-Source Weighted-Ensemble Precipitation (MSWEP) in two basins on the TiP and found that the precipitation products improved with gauge adjustment. They relate uncertainties in the precipitation products to precipitation properties, climate conditions, elevation, and topography. The precipitation products displayed similar spatial patterns but large differences in the precipitation amount. Lu & Yong [8] compared IMERG V05 uncalibrated and gauge calibrated precipitation with GSMaP (Global Satellite Mapping of Precipitation) Moving Vector with Kalman Filter (GSMaP-MVK) and gauge-adjusted GSMaP (GSMaP-Gauge). The precipitation products were able to capture spatial patterns of precipitation over the TiP. The authors found gauge-adjusted products to be more consistent with the ground measurements when compared to the satellite products without gauge adjustment. However, they highlight that measuring light precipitation and snow remains a challenge. Gao & Liu [9] evaluated TMPA Analysis 3B42 V6, Real Time version (TMPART), CMORPH, PERSIANN over the TiP for the period 2004–2009. TMPA and CMORPH performed better over the TiP as compared to PERSIANN and TMPART. Like other authors, they stressed the benefit of gauge adjustment of TMPA over TMPART. All of the products better agreed with gauges over humid regions and under-/overestimated light/heavy precipitation. The biases found in TMPART and PERSIANN derive from topography, complex terrain, and surface roughness. Ma et al. [10] evaluated GPM IMERG over the TiP from April 2014 to December 2017 using gauge observations and found that IMERG has a strong ability to capture precipitation and is in a good agreement with the gauges. They separated the TiP according to the Köppen climate classification and stated that IMERG performed everywhere well besides in the class of the arid-desert-cold climate (Bwk). The authors also stressed the potential of IMERG for snowfall detection; however, it still lacks the detection of light and moderate snowfall. Tong et al. [11] compared gauge observations, reanalyses (European Centre for Medium-range Weather Forecasts (ECMWF) reanalysis ERA-40, ERA-Interim) and satellite retrievals (Asian Precipitation—Highly-Resolved Observational Data Integration Towards Evaluation (APHRODITE), TMPA) over the plateau and basin scale of the TiP. The authors conclude that all products are able to detect the large-scale precipitation regime (monsoon precipitation in summer, westerly induced precipitation in winter). Both satellite based precipitation products underestimated the gauge observations from the China Meteorological Administration (CMA). The CMA and APHRODITE displayed similar trends (decreasing/increasing trend in summer/spring and winter).

Singh et al. [12] analyzed heavy precipitation events over India while using two precipitation products derived from Insat-3D which are the Hydro estimte (HE) and the INSAT Multi-Spectral Rainfall Algorithm (IMSRA) as compared to GPM IMERG. Both of the Indian products are available in the half hourly and daily resolution. IMSRA comes with a resolution of 10 km and HE is available at 4 km. HE and IMERG show a similar performance, whereas IMSRA underestimates precipitation. IMERG is found to perform best at very heavy precipitation events. Reddy et al. [13] evaluated multi-satellite precipitation products over India during the monsoon season and found strong biases in the precipitation data sets over orographic regions and strong differences among the satellite products. IMSRA and HE tend to overestimate precipitation over India. GPM based products, such as GSMaP (NRT & moving vector with Kalman filter) and IMERG (V4 real-time and final run), estimate precipitation better as compared to the Insat-3D products. Mitra et al. [14] compared HE and IMSRA with IMERG and daily rain gauge data. The satellite based precipitation products capture precipitation patterns very well regarding the occurrence of precipitation. HE correlates strongly with IMERG; however, it underestimates heavy precipitation at high elevation.

The GPM mission is the latest satellite based precipitation mission from NASA (USA) and JAXA (Japan) as a joint mission centered on the GPM core satellite. The mission started in 2014 and produces precipitation estimates for the spatial coverage from 60°N–60°S. This mission is the successor of TRMM and was established with some improvements. Besides the larger spatial coverage of GPM as compared to TRMM (35°N–35°S), the GPM core satellite was equipped with two radar frequencies

(Ku band and Ka band at 13.6 GHz and 35.5 GHz, respectively) and a multi-channel imager. The GPM mission produces many precipitation data sets and one of them is the Level 3 product IMERG (final run), which uses the GPM core satellite as the reference standard.

Validation studies that compare various satellite retrievals conclude that GPM IMERG is, so far, the best performing quasi-global satellite based precipitation retrieval which combines MW (Microwave) and IR (InfraRed) data [15–21]. These precipitation estimates are further gauge calibrated while using a global product from the Global Precipitation Climatology Center (GPCC) [22].

The use of MW is the most reliable satellite precipitation retrieval technique [23]. However, onboard of LEO satellites, the MW precipitation estimates suffer from a poor temporal resolution. To overcome this temporal gap, adjusted GEO IR information is added in which texture and patch classification as well as routine updates using MW based precipitation estimates was applied [24,25]. The algorithms of both CMORPH and PERSIANN make use of MW and only one IR band between 10.7 μm–11.5 μm, where the IR wavelength used depends on the GEO satellites used (GOES-8, GOES-10, Meteosat-5, Meteosat-7, GMS-5; [5]). However, using one IR band only leads to uncertainties and inaccuracy in the precipitation retrieval [26,27]. Several studies showed the benefits of the use of multispectral satellite information (IR and/or visible (VIS) bands) [28–31].

The retrieval of precipitation with GEO systems frequently consists of two steps, in which, first, the precipitation area is delineated and second the precipitation rates are retrieved [29,31,32]. With regard to the precipitation area, we recently showed that multispectral IR data from Insat-3D and Elektro-L2 outperformed IMERG's IR only precipitation [33].

In the current paper, we present a new approach for a retrieval method of precipitation rates, where we train models that are based on IR brightness temperatures from the geostationary (GEO) satellites Insat-3D and Elektro-L2 with precipitation from GPM IMERG V05B in order to obtain the rates in a higher spatial resolution (4 km) and in a temporal resolution of 30 min. In detail, we train the machine learning algorithm Random Forest (RF) regression with the gauge calibrated MW precipitation rates from GPM IMERG, which is the best quality product of IMERG. This article is the continuation of the paper by Kolbe et al. [33], who introduced the delineation of the precipitating area in a first step.

The paper is structured, as follows. First, the processing scheme of the precipitation retrieval and the data used in the study are described. For the latter, the methodology applied to retrieve the precipitation rates based on the RF regression models (training and validation) are described. The results from the precipitation area delineation [33] and precipitation rate retrieval are combined. The results section includes the validation of the predicted precipitation rated against IMERG's gauge calibrated MW precipitation. We further show the results from the validation for (1) several elevation ranges and (2) single months. We also compare IMERG's gauge calibrated MW precipitation against IR only precipitation. The comparison of the validation of the predicted precipitation rates using RF and the gauge calibrated MW precipitation with the validation of the IR only precipitation and the gauge calibrated MW precipitation shows if our retrieval outperforms IR only precipitation estimates. Additionally, we depict the results of an independent validation of our product against 28 Chinese rain gauge observations. The result section is followed by a discussion and conclusions.

2. Overview of the Processing Scheme of the Precipitation Rates Retrieval and Data

2.1. Processing Scheme of the Precipitation Rates Retrieval

We retrieve precipitation rates on the TiP using GEO satellite data and an RF machine learning algorithm. Figure 1 provides an overview of the processing scheme of the precipitation retrieval. The preprocessing of the satellite and precipitation data, the preselection of the data according to specific restrictions, the split in training and validation data, as well as the precipitation area delineation (RF classification) was already performed and analyzed. The GEO data were resized from 4 km (3.9 μm, 8 μm, 9.7 μm, 10.8 μm, 11.9 μm) and 8 km (6.8 μm) to 11 km (0.1°) to match the spatial resolution of the IMERG data while using nearest neighbor interpolation. The split in training (80%) and validation (20%)

was done for each scene separately. The training data are used for the feature selection, RF parameter tuning (both on the monthly scale), and for the RF model training (scene scale). For details see [33] and Figure 1 ("classification").

From IMERG we used the gauge calibrated MW precipitation data set for training (80% of precipitating pixels from each scene) and validation (20% of precipitating pixels from each scene). We did not choose all of the data for modeling and removed data with regard to some conditions that we defined to improve the modeling based on additional data provided by GPM IMERG and Insat-3D (see "preselection" in Figure 1). Besides the precipitation information, GPM IMERG provides a quality index for the gauge calibrated MW precipitation and the overflight time of each sensor which contributes to the GPM IMERG product. Only data with an index > 0.9 contain data from the half-hour MW swath. Therefore, we restricted ourselves to data that only contain precipitation with (1) a quality index > 0.9 in order to only use reliable precipitation data. In addition, (2) we only used cloudy data defined by the cloud mask from Insat-3D and (3) limited ourselves to data that does not exceed the time difference of seven minutes between the overflight of the sensors contributing to IMERG and Insat-3D. We set the maximum threshold of seven minutes to make sure that precipitation events can be correctly captured by the GEO satellites' brightness temperatures. For more information about the satellite and precipitation data and its preprocessing (Elektro-L2, Insat-3D, GPM IMERG), please refer to [6,25,33,34] and Figure 1. To retrieve precipitation over the TiP, we used two different undersampling techniques to balance the uneven distribution of low and high precipitation rates to enhance the RF modeling. The distribution of precipitation over the TiP is diverse (very arid in the west of the TiP, wet in the south and east of the TiP). The model would not be able to correctly map precipitation without any sampling technique. For more information about sampling techniques, refer to [29,35].

2.2. Data

We used four spectral IR bands from Insat-3D (3.9 μm, 6.8 μm, 10.8 μm, 11.9 μm) [36–39], and combined these with two additional bands from Elektro-L2 (8 μm, 9.7 μm) [40]. We used the combination of both satellites to extend the spectral range. The bands from Elektro-L2 provide additional information regarding the cloud phase, which is helpful for the retrieval of precipitation. Based on the bands of Insat-3D and Elektro-L2, we calculate band differences, one band geostatistical texture features (variogram, rodogram, madogram), and two band geostatistical texture features (cross variogram, pseudo cross variogram). Insat-3D also provides the sun azimuth angle and the solar zenith angle. Based on Insat-3D, a cloud mask is also available (clear, cloudy, probably clear, probably cloudy, cold space). The cloud mask is used to exclude all data which is not defined as cloudy. Removing non cloudy pixels reduces the amount of data and is essential for the models to find possible precipitation fields. The GEO satellite data and the cloud mask are available for the study period May–September 2017 every 30 min in a 4 km resolution.

We used the brightness temperatures from the GEO satellites, the sun azimuth angle, the solar zenith angle, the band differences, and texture features as predictors for the RF models [33,41]. The gauge calibrated MW precipitation served as target variable for the RF regression models. We validated our new precipitation product against the independent part of the gauge calibrated MW precipitation, which was not used for training (validation data).

In addition, we compared the gauge calibrated MW precipitation with the IR only precipitation from IMERG. This was done to demonstrate the potential improvement of multispectral information used in PRETIP over the use of the single 10.8 μm IR band in IR only, both with reference to the gauge calibrated MW precipitation.
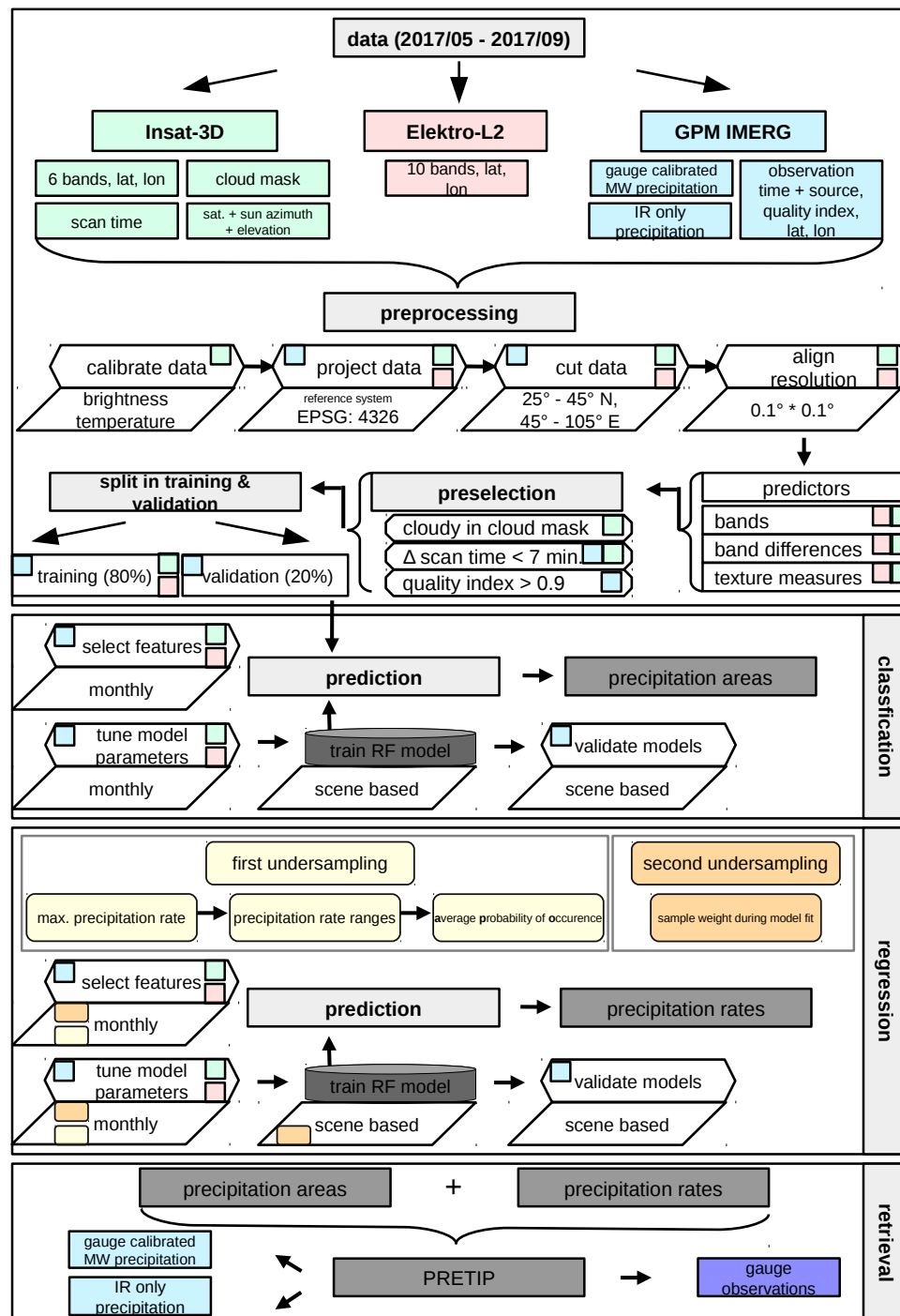
Figure 1. Processing scheme for the precipitation rates retrieval based on Elektro-L2, Insat-3D, and the gauge calibrated microwave (MW) precipitation from GPM IMERG. RF = Random Forest. PRETIP = Precipitation Retrieval covering the Tibetan Plateau. Modified and extended after Kolbe et al. [33].

We also validated our product against 28 daily gauge measurements from the Chinese Ministry of Water Resources. The gauge stations are evenly distributed around the south central TiP (see Figure 2). These data are available without any gaps and are completely independent from GPM IMERG to which most state weather services contribute.
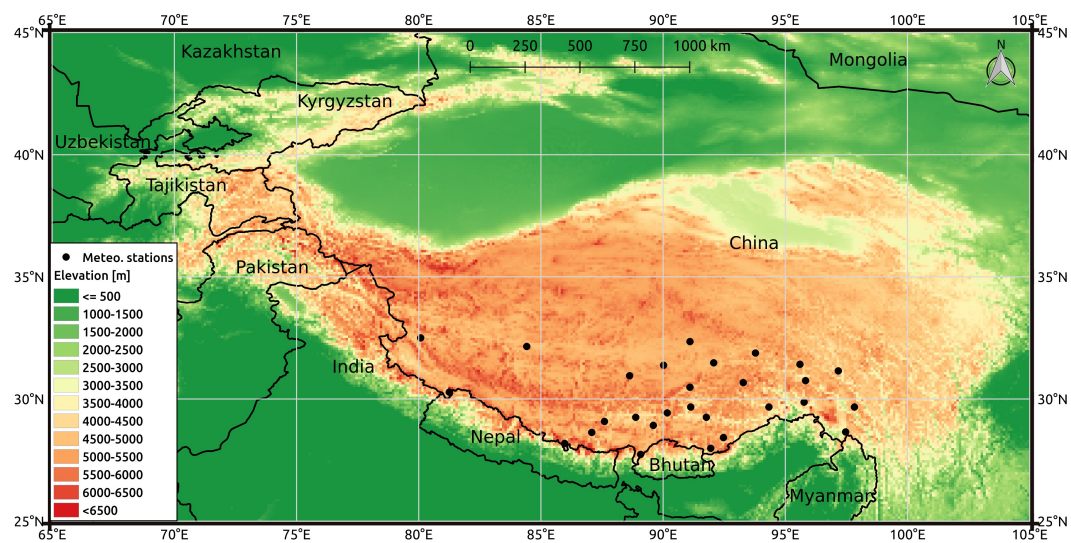
Figure 2. Study area including the distribution of 28 IMERG-independent gauge stations provided by the Chinese Ministry of Water Resources. The digital elevation model is taken from GTOPO30 [42].

3. Random Forest Model Training and Validation

3.1. Random Forest Regression

RF was first introduced by Breiman [43,44]. We used the scikit-learn package for implementing the RF model for precipitation rate assignment [45]. We only considered precipitating pixels (>0 mm/h) for feature selection, model tuning, and model training. The precipitating pixels were defined by the gauge calibrated MW precipitation. High precipitation rates are known to be very uncertain and they can lead to artifacts in the precipitation retrieval. In general, the maximum precipitation rate is set to 50 mm/h in IMERG V05B; however, due to correction processes, this rate can increase. We removed all of the pixels with precipitation rates higher than 50 mm/h, since we consider them to be unreliable (Huffman, 2019, personal communication).

Regarding the approximately 90 predictors (bands, band differences, geostatistical texture features, solar zenith angle, and sun azimuth angle), we performed a feature selection based on a recursive backward feature elimination on the monthly scale. The monthly scale accounts for the seasonality of the different precipitation regimes on the TiP. Figure 3 displays the most important features for the precipitation rates retrieval on average for all months based on the models' feature importance. The RF parameter tuning was also done on the monthly scale. Here, we tuned the number of trees (n-estimator) from 700–1400 in steps by 100 and the number of features to consider the best split (max-features) from 2 to 8 by 2. In both procedures, we randomly sampled 10% of each month where only precipitating pixels were considered and repeated the feature selection and tuning 50 times to account for stability. This resulted in a number of trees between 900 and 1300, and in two features to consider the best split on the monthly scale.

The features that are defined by the feature selection and the numbers for the parameters defined by model tuning were used to train and apply RF models for each scene separately. In addition, we also trained daily models with these features and numbers of tuning parameters. The daily trained models were used for precipitation rate assignment for those Insat-3D and Elektro-L2 scenes, where no MW precipitation is available for model training on a scene base. For an overview, please refer to the part "regression" in the processing scheme presented in Figure 1.
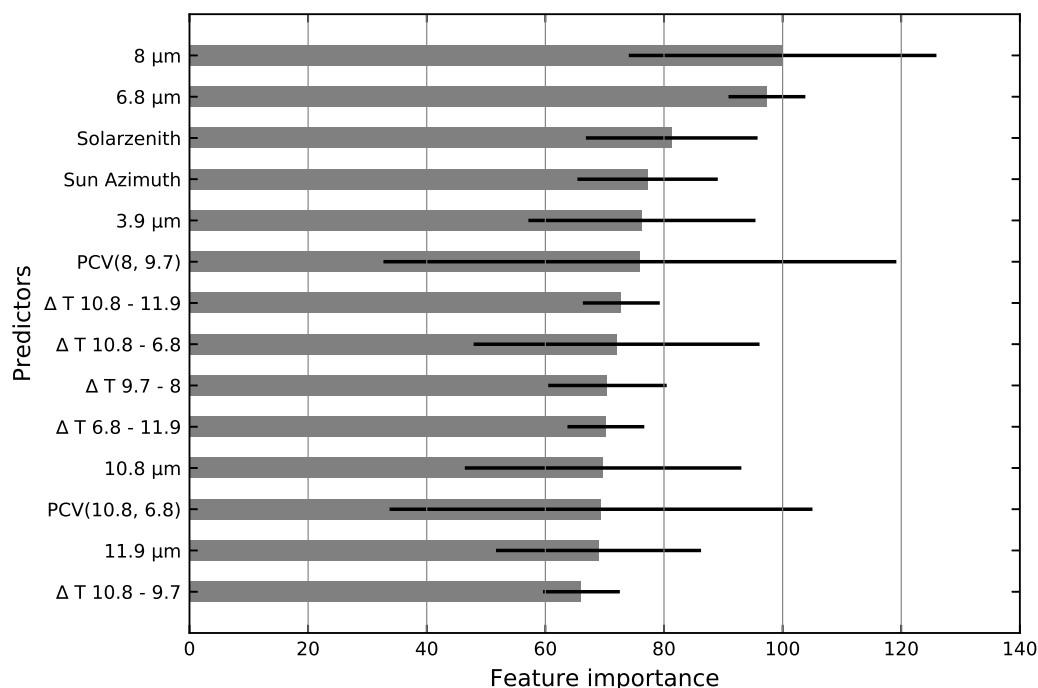
Figure 3. Mean feature importances with standard deviations based on a randomly selected subset of the training data used for the precipitation rates retrieval. The most important feature is set to 100 (*x*-axis) and the other features were depicted accordingly. The standard deviation from each predictor was used to calculate the error bars. Depending on the variability of the errors, few of them exceed 100. PCV = pseudo cross variogram. ∆T = band differences.

Because the frequencies of low and high precipitation rates are not evenly distributed, we undertook two different undersampling methods. Undersampling means the removal of pixels from the data according to specific rules. The first undersampling was used in the feature selection and RF parameter tuning on a monthly basis. Here, we defined precipitation rate ranges for each month (1 mm/h) and checked the average probability of occurrence (APO) for the whole precipitation rates. If the number of the probability of occurrence was higher than the APO, we undersampled our data by taking the number of the average precipitation rate class as threshold and selected as many pixels as the APO from the other available classes according to this threshold. If the classes of precipitation rate ranges contained less pixels than the APO, we took all the pixels from these classes (cf Figure 4).

The second undersampling was used in the feature selection, tuning, and training. Here, we used the sample weight in the fit command of the RF, which assigns a balanced weight to each precipitation class.

In a last step, the results from the precipitation area delineation and precipitation rates retrieval were combined in the final product. This means that the precipitation rates were only assigned in those areas that were defined by the precipitation area delineation (see "retrieval" in Figure 1).

3.2. Validation of the Random Forest Models

We validated our predicted precipitation rates while using the 20% of precipitating pixels from each scene that were not used for feature selection, model tuning, and model training. The precipitation rate was assigned to these pixels for every scene using the respective RF models and was then compared to the corresponding precipitation rate of the gauge calibrated MW precipitation validation data set.
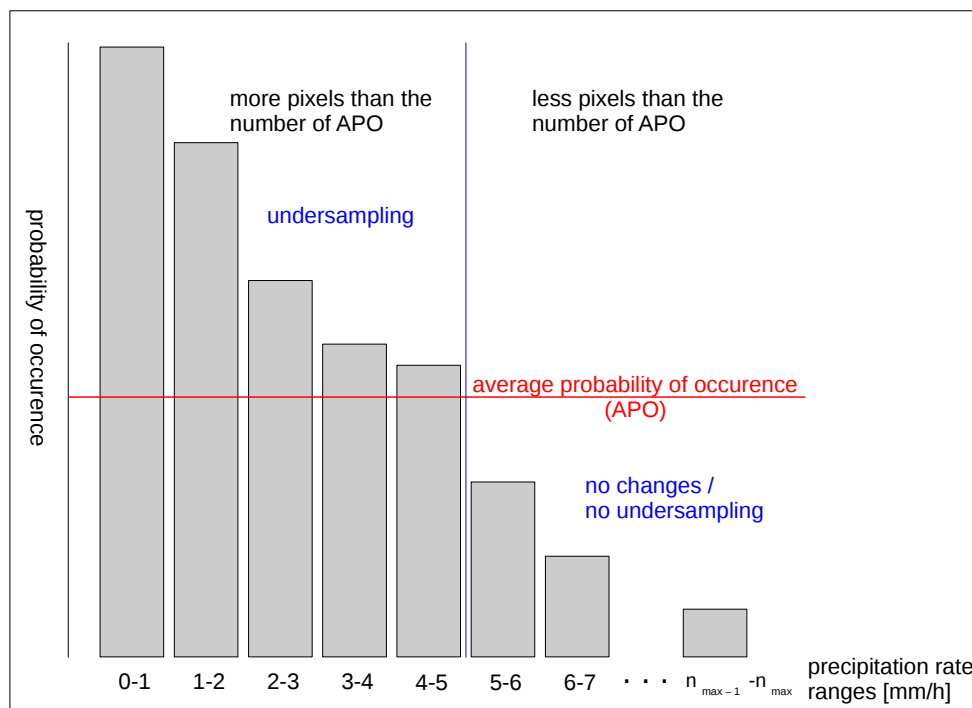
Figure 4. Overview of the first undersampling method.

In addition, we compared the precipitation rate of the IR only precipitation validation data set to the gauge calibrated MW precipitation validation data set in order to assess the improvement of performance by using the RF model.

We calculated the correlation coefficient (R), the mean absolute error (MAE), and root-mean-square error (RMSE) to evaluate the performance of the RF models (Table 1) for (i) the validation of predicted precipitation rates against the gauge calibrated MW precipitation rates and (ii) the validation of IR only precipitation rates against the gauge calibrated MW precipitation rates. Because the spatial resolution of the IMERG product is 0.1°, the above mentioned validation steps were undertaken in the corresponding spatial resolution of 0.1° (11 km).

Table 1. Overview of the validation measures used for the validation of the precipitation rates retrieval. y_i represents the prediction and x_i is the observation. $\overline{x}$ and $\overline{y}$ are the mean values of x and y, respectively.

Validation Measure	Equation	Range	Optimal Value
Correlation coefficient	$R = \dfrac{\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\overline{x})^2(y_i-\overline{y})^2}}$	$[-1,1]$	-1 or 1
Mean Absolute Error	$MAE = \left(\frac{1}{N}\right)\sum_{i=1}^{N}|y_i - x_i|$	$[0,\infty]$	0
Root Mean Square Error	$RMSE = \sqrt{\left(\frac{1}{N}\right)\sum_{i=1}^{N}(y_i - x_i)^2}$	$[0,\infty]$	0

We computed the precipitation totals of PRETIP and divided them through the number of available scenes (average validation measures) in order calculate the spatial distribution of the average validation measures for the study period May–September 2017. In addition, we divided the average validation measures through the averaged precipitation rate. The spatial distribution of these maps can be found in the results section in Figure 10.

3.3. Combination of Precipitation Area Delineation and Precipitation Rates Retrieval

In a last step, the method for precipitation area delineation [33] is combined with the method for precipitation rates assignment into the final Precipitation REtrieval covering the TIbetan Plateau (PRETIP), which offers high spatio-temporal (30 min, 4 km) precipitation information for the TiP. Thus, PRETIP exists in a 11 km (0.1°) and a 4 km resolution. PRETIP was independently validated against daily precipitation measurements from 28 gauge stations from the Chinese Ministry of Water Resources. For this purpose, the PRETIP precipitation data (30 min, 4 km) were aggregated on a daily base from May–September 2017 using all precipitating pixels from each available scene.

4. Results

4.1. Validation of the Predicted Precipitation Rates against IMERG's Gauge Calibrated MW Precipitation Rates

First, we validated the predicted precipitation rates against the gauge calibrated MW precipitation validation data set on a scene base (20% of precipitating pixels from each scene). For the months May, June, July, August, and September 2017 are 237, 271, 273, 253, and 155 scenes available, respectively. The average correlation coefficient of 0.59 for all scenes from the study period indicates a good performance. The MAE ranges from 0.72–1 mm/h for every 30 min on average and the average RMSE ranges from 1.24–1.74 mm/h for every 30 min. There are hardly any differences in the retrieval of precipitation rates when comparing the validation measures of the rainy season (July, August) against the validation measures of the dry season (May, September), see Figure 5 for details.
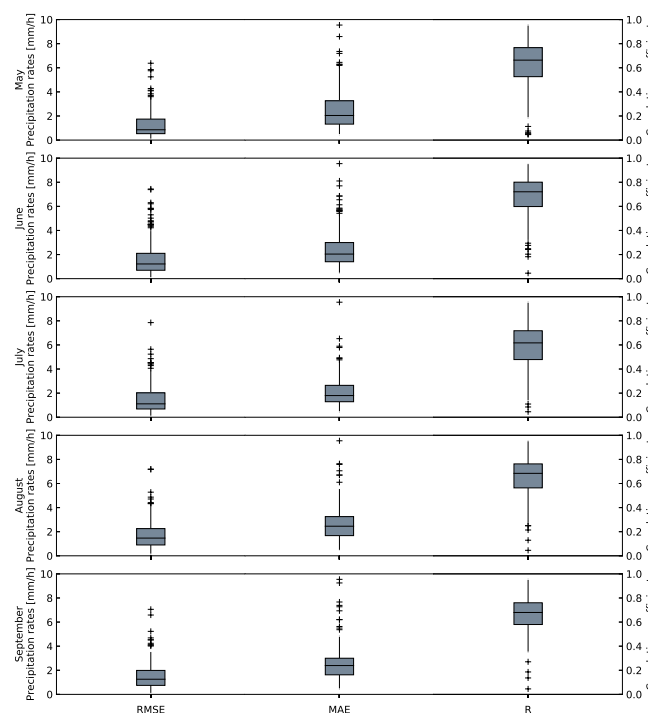

Figure 5. Performance of the validation of predicted precipitation rates against the gauge calibrated MW precipitation rates as boxplots. Validation measures RMSE, MAE (left *y*-axis) and correlation coefficient R (right *y*-axis) were calculated for each scene separately and then averaged on a monthly basis. The boxes display the percentiles (25th, 50th, and 75th). The whiskers indicate extreme data up to 1.5 times of the interquartile range. Crosses mark outliers. The width of the boxes is relative to the number of validation scenes.

We analyzed the performance of PRETIP on different elevation ranges in steps by 1000 m. The elevation was divided into following ranges: 2500–3500 m, and so on until 5500–6500 m. For each

class, between 372 (5500–6500 m) and 1148 (3500–4500 m) validation pixels were included. Therefore, MAE and RMSE can be interpreted in the unit mm/h for each 30 min. The mean value of the correlation coefficient does not differ a lot when comparing the correlation coefficients for the different elevation ranges. However, the variability of the correlation coefficient R increases with an increasing elevation. The MAE and RMSE do not show many differences in the classes of elevation ranges. For the validation measures MAE and RMSE, we find lots of outliers from the mean values in each elevation range (see Figure 6 for details).
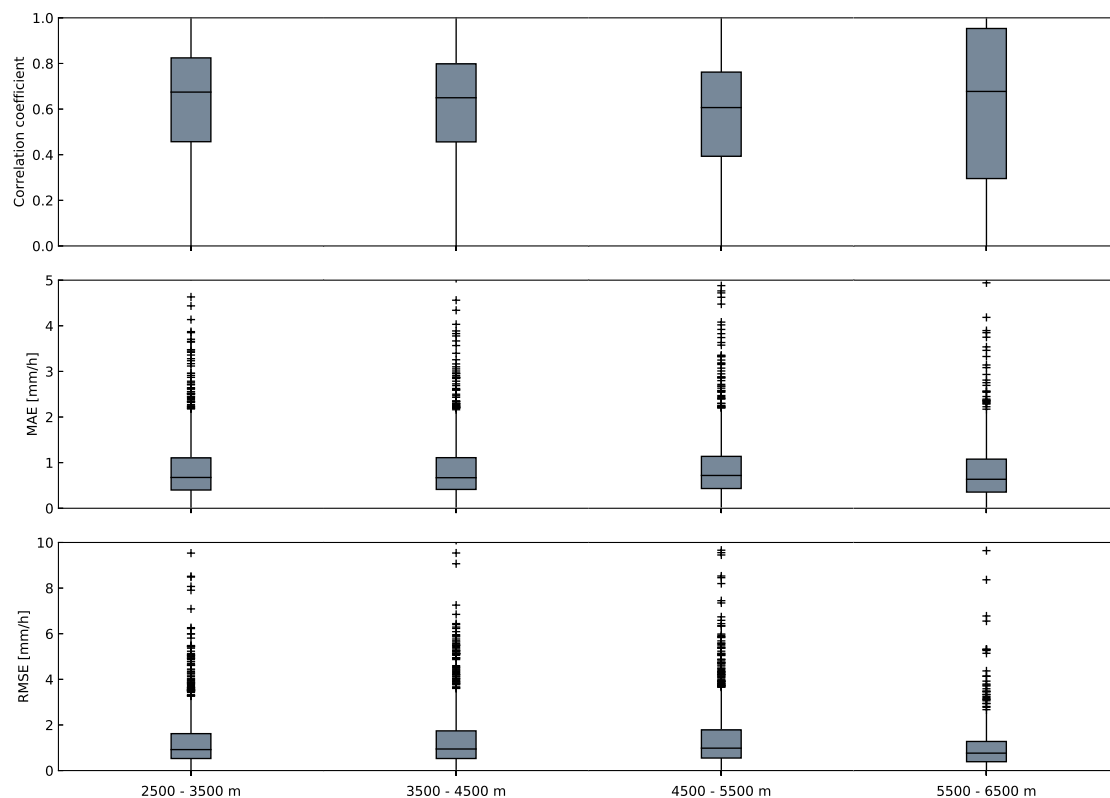

Figure 6. Validation measures correlation coefficient R, mean absolute error (MAE), and root-mean-square error (RMSE) of the predicted precipitation rates against the gauge calibrated MW precipitation rates sorted by elevation ranges in 1000 m steps. The boxes display the percentiles (25th, 50th, and 75th). The whiskers indicate extreme data up to 1.5 times of the interquartile range. Crosses mark outliers. The width of the boxes is relative to the number of validation scenes.

4.2. Results of the Comparison between IMERG's IR Only Precipitation and IMERG's Gauge Calibrated MW Precipitation

Further, we compared the gauge calibrated MW precipitation with the IR only precipitation product on a 30 min and 11 km resolution based on the 20% of the selected validation pixels. Our results reveal a correlation coefficient on average between 0.12 and 0.24, an MAE between 1.35 and 1.65, and an RMSE between 2.20 and 2.69 when comparing IR only and the gauge calibrated MW precipitation. Table 2 provides an overview of the average validation measures R, MAE, and RMSE for the predicted precipitation rates and the IR only precipitation estimates, both with reference to the gauge calibrated MW precipitation.

Table 2. Comparison of predicted precipitation rates with IR only precipitation both with reference to the gauge calibrated MW precipitation based on the validation measures.

Validation Measure	RF Predicted Precipitation Rates (PRETIP)	IMERG IR Only Precipitation Rates
R	0.59	0.18
MAE	0.91	1.56
RMSE	1.63	2.18

Figure 7 displays the spatial distribution of predicted precipitation rates (Figure 7a), IMERG's gauge calibrated MW precipitation rates (Figure 7b) and IMERG's IR only precipitation (Figure 7c) of a sample scene on 2 August 2017 at 16:00 UTC covering the western TiP. The predicted precipitation rates tend to overestimate the precipitation amount as compared to the gauge calibrated precipitation rates. The scene is representative, because the predicted precipitation in general overestimate precipitation. IR only contains only very low precipitation rates when compared to the gauge calibrated MW precipitation. IR only highly underestimates the precipitation amount which becomes obvious in Figure 7e. This example only shows one scene; however, it exemplifies the overall trend: IR only precipitation does not capture much precipitation compared to the gauge calibrated MW precipitation.
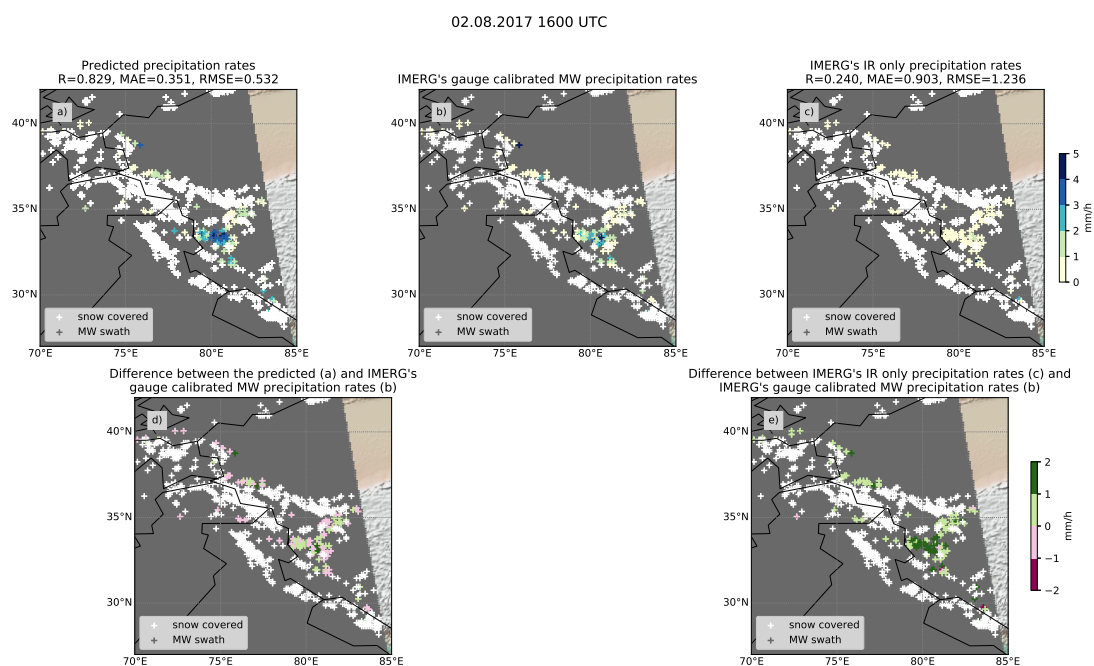

Figure 7. Comparison of (**a**) predicted precipitation rates with (**b**) gauge calibrated MW precipitation rates and (**c**) IR only precipitation rates at 2 August 2017 at 16:00 UTC. (**d**) shows the difference between (**a**) and (**b**). (**e**) displays the difference between (**c**) and (**b**). Green/pink represents the under-/overestimation of the predicted precipitation rates/IR only precipitation rates compared to the gauge calibrated MW precipitation rates. Grey represents the MW swath in which the gauge calibrated MW precipitation is available. White areas are pixels which were masked out due to snow/ice based on IMERG's quality index.

Figure 8a displays the validation results from the predicted precipitation rates for the complete TiP and the complete study period (May–September 2017). The correlation coefficient R is on average convincing with 0.59; however, it shows some outliers where the correlation between the validation pixels of the gauge calibrated MW precipitation and the prediction failed. The RMSE and MAE are on average 1.59 and 0.91, respectively. Both of them show some outliers, up to 5 mm/h. The scatter plot in

Figure 8b shows the relation between the gauge calibrated MW precipitation and the precipitation rates predicted by the RF model with use of the frequency distribution. It seems that the precipitation under- and overestimate the gauge calibrated MW precipitation. However, the majority of the validation pixels scatter around the 1:1-line. The gauge calibrated MW precipitation comprises precipitation estimates up to 50 mm/h, whereas the predicted precipitation rates do not exceed 31 mm/h. High precipitation amounts seem to be underrepresented in the RF training and can therefore not be predicted by the models.
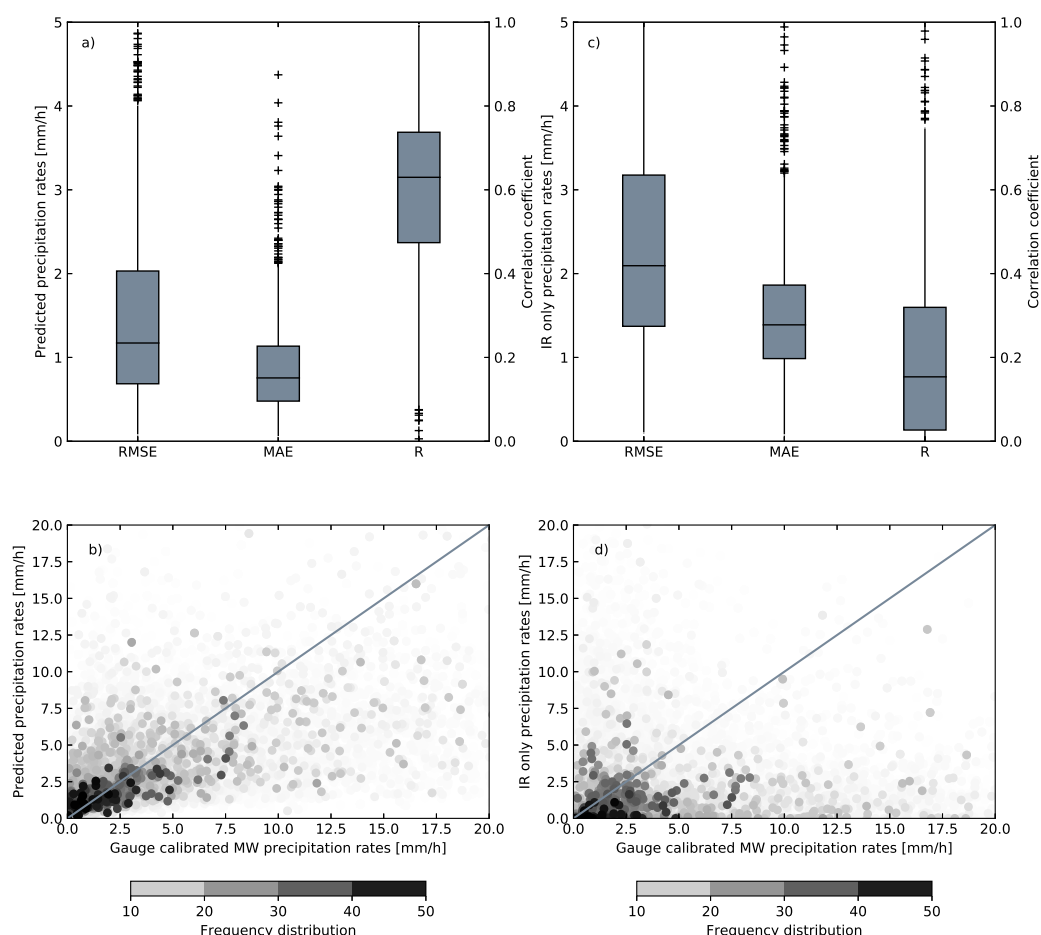

Figure 8. (**a**) Validation measures RMSE, MAE, and correlation coefficient R for the predicted precipitation rates based on the gauge calibrated MW precipitation rates from IMERG (independent validation data set: 20% of precipitating pixels from each scene). (**b**) Scatter plot between gauge calibrated MW precipitation rates and predicted precipitation rates [mm/h] with the frequency distribution of the validation pixels. (**c**) Validation measures RMSE, MAE and correlation coefficient R for the IR only precipitation rates compared to the gauge calibrated MW precipitation rates from IMERG. The boxes display the percentiles (25th, 50th, and 75th). The whiskers indicate extreme data up to 1.5 times of the interquartile range. Crosses mark outliers. The width of the boxes is relative to the number of validation scenes. (**d**) Scatterplot between gauge calibrated MW precipitation rates and IR only precipitation rates [mm/h] with the frequency distribution of the validation pixels.

Figure 8c displays validation results from the comparison of IR only and the gauge calibrated MW precipitation using the same validation pixels as for the comparison of the predicted precipitation rates and the gauge calibrated MW precipitation for the complete study area and study period (May–September 2017). When compared to Figure 8a, the correlation coefficient R is lower, with outliers in both directions. The RMSE and MAE are both higher as compared to Figure 8a with values, on average, 1.56 mm/h for the MAE and 2.18 mm/h for the RMSE (see Table 2 for

comparison). MAE and RMSE show some outliers, up to 5 mm/h. Figure 8d shows the connection between the gauge calibrated MW precipitation and IR only precipitation with use of the frequency distribution. Whereas a perfect connection would show lots of points scattering around the grey line we find here over-/underestimation of low and high precipitation amounts for both products. Most of the validation pixels of IR only underestimate the gauge calibrated MW precipitation.

Figure 9a displays the temporal cycle (May–September 2017) of the gauge calibrated MW precipitation and the predicted precipitation rates. In general, the precipitation rates follow the trends of the gauge calibrated precipitation estimates, although they do not reach the same magnitude.
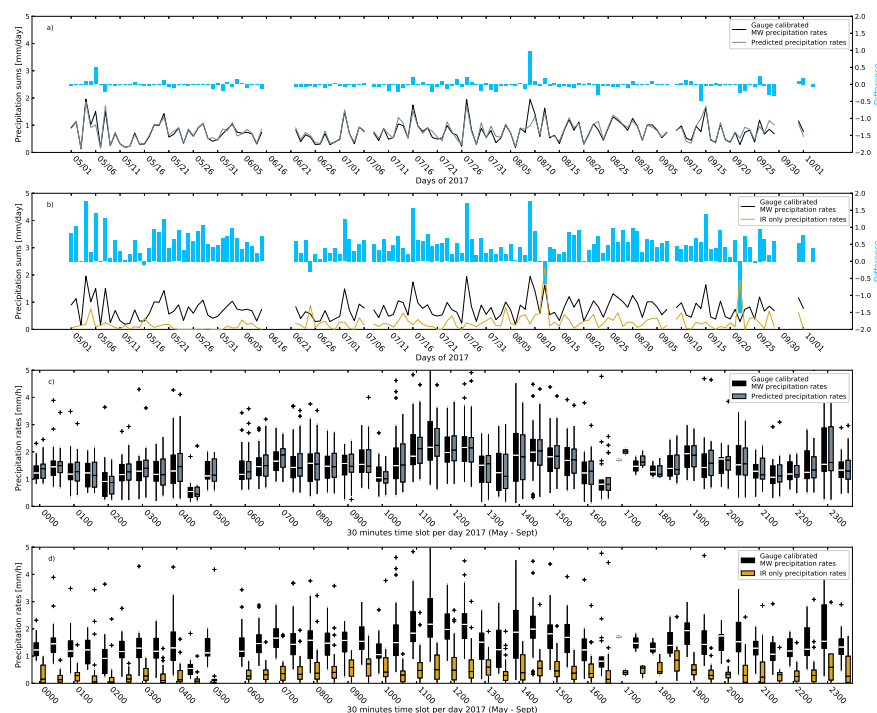

Figure 9. Daily precipitation sums for gauge calibrated MW precipitation rates with the predicted precipitation (**a**) and with IR only precipitation rates (**b**). (**c**,**d**) depict the gauge calibrated MW precipitation rates, the predicted precipitation rates the IR only precipitation rates on an averaged daily cycle as boxplots. The boxes display the percentiles (25th, 50th and 75th). The whiskers indicate extreme data up to 1.5 times of the interquartile range. Crosses mark outliers. The width of the boxes is relative to the number of validation scenes.

The temporal cycle (May–September 2017) of the gauge calibrated MW precipitation and IR only in Figure 9b demonstrates large differences. The differences are most pronounced in the months of May and June, however, at hardly any point in time is IR only precipitation able to capture the same amount of precipitation as the gauge calibrated MW precipitation. This stresses the underestimation of precipitation in IR only compared to the gauge calibrated MW precipitation.

Figure 9c compares the diurnal variation between the gauge calibrated MW precipitation and the predicted precipitation rates. The predicted precipitation rates follow the trend of the gauge calibrated MW precipitation with peaks at around 12:00 UTC, 15:00 UTC and 22:00 UTC. The occurrence of precipitation events during the day might be associated with thermal convection at the slopes and the precipitation events during night were already described by Rüthrich et al. [46].

Figure 9d displays the diurnal variation between the gauge calibrated MW precipitation and IR only precipitation. IR only precipitation does not follow the same trends as the gauge calibrated MW precipitation and it contains much lower precipitation amounts when compared to the gauge calibrated MW precipitation. The high precipitation peaks are not captured by IR only. We compare the gauge calibrated MW precipitation with (i) our predicted precipitation rates and (ii) IR only to understand the similarities and differences of the products.

Figure 10a illustrates the total precipitation sums of the gauge calibrated MW precipitation. Figure 10b,c display the differences of PRETIP (Figure 10b)/IR only (Figure 10c) to the gauge calibrated MW precipitation to give an overview of the spatial distribution of precipitation. The data are shown in the 11 km resolution. The differences of PRETIP and IR only reveal that PRETIP is wetter, especially around the south central and east central of the TiP, and that IR only does not capture much precipitation at all. Figure 10a,c display a rather dry west and a wet south at the Himalayas associated with the increasing elevation.
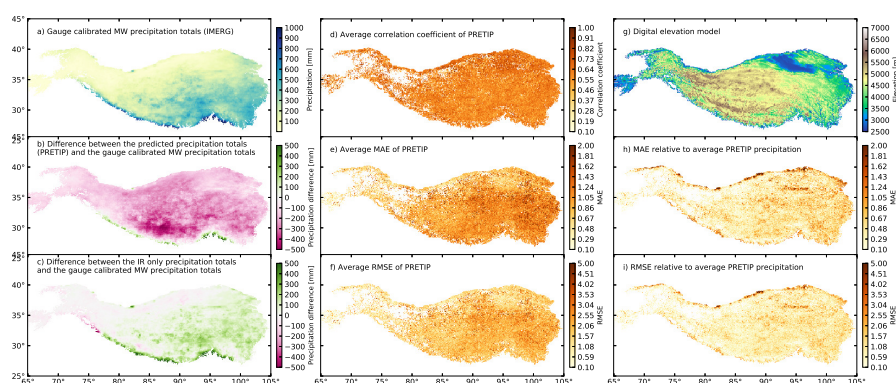

Figure 10. (**a**) depicts the precipitation totals of IMERG's gauge calibrated MW precipitation and (**b**,**c**) show the differences of IMERG's gauge calibrated MW precipitation with PRETIP and IMERG's IR only precipitation for the time period of May–September 2017. For the precipitation totals of (**a**) and the differences of (**c**) only those scenes in which PRETIP is available were used. Green/pink represents the under-/overestimation of the gauge calibrated MW precipitation rates from IMERG when compared to the predicted precipitation rates (PRETIP)/IR only precipitation (IMERG). (**d**–**f**) show the average R, MAE, and RMSE of the validation data between PRETIP and the gauge calibrated MW precipitation relative to the number of available PRETIP scenes. (**g**) shows the digital elevation model of the TiP. (**h**,**i**) illustrate the average MAE and RMSE of the validation between PRETIP and the gauge calibrated MW precipitation relative to the average PRETIP precipitation for the complete time period (May–September 2017).

Figure 10d–f show the validation measures correlation coefficient R, MAE, and RMSE relative to the number of the available scenes for the validation of the predicted precipitation rates against the gauge calibrated MW precipitation. This helps to identify differences in the various regions of the TiP. Parts of the western TiP are frequently masked out due snow and ice. IMERG adds IR information in

areas where snow and ice interferes with the backscattering of the microwave signatures. This leads to a decrease in the quality index (below 0.9). Hence, we exclude these data from our modeling process [25]. Therefore, the western part of the TiP experiences data gaps. There are not many differences in the spatial distribution of the correlation coefficient. Concerning the MAE and RMSE, we find higher values in the central eastern part of the TiP, which is probably associated with the higher amount of precipitation.

Figure 10g displays the digital elevation model of the TiP. Figure 10h,i display the average validation measures MAE and RMSE relative to the average PRETIP precipitation. The validation measures increase with the increase of precipitation. The variability of very low precipitation in the west, together with the limitation of MW use on ice and snow surfaces and monsoon influenced precipitation makes the modeling of precipitation difficult. Lower validation measures are retrieved in the central TiP and high errors can be found at the lower elevation ranges, especially at the northern border of the study area.

4.3. Validation of PRETIP against 28 Chinese Rain Gauge Observations

We compared the independent gauge observations which are available for each day with the corresponding pixel in PRETIP to see if precipitation from the gauges is captured by our new precipitation product. We found a correlation coefficient of R = 0.49 averaged over all 28 Chinese rain gauge observations. The worst correlation is R = 0.23 and the best correlation is R = 0.73. We found an average MAE of 7.1 and an average RMSE of 12.3, which shows the high variability of the precipitation captured by the gauge observations. The lowest/highest MAE is 3.94/9.69 and the lowest/highest RMSE is 8.09/16.32 (see Figure 11 for details).

Figure 11 displays the spatial distribution of the validation measures correlation coefficient R, MAE, and RMSE for validation of PRETIP (4 km resolution) with the 28 Chinese rain gauge observations. The colours indicate the level of the validation measures based on the colour scheme for each measure separately. The correlation coefficients are strongest in the southern and in the eastern TiP. The western stations rather receive low correlations. The MAE is highest in the central and south of this spatial extent. The MAE is comparably low in the western and eastern TiP. The RMSE is remarkably high in the central south, where correlations and MAE are also high. Low RMSE values occur in the western and the eastern TiP.
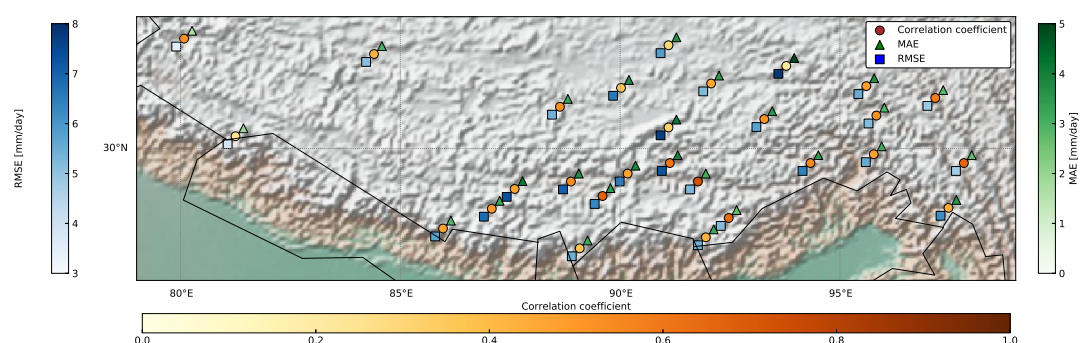

Figure 11. Spatial distribution of validation measures correlation coefficient (yellow/brown circles), MAE (green triangles), and RMSE (blue squares).

Figure 12a shows the temporal variation of precipitation from both, PRETIP and the gauge observations averaged over the study period. In general, PRETIP follows the trends of the gauge observations and seems to match the timing of precipitation that occurs in the gauge measurements; however, the magnitude strongly differs. PRETIP clearly receives more precipitation than the gauge observations. The scatterplot of PRETIP and the gauge observations (Figure 12b) clearly demonstrate the overestimation of PRETIP as compared to the gauge observations. Especially the high precipitation values differ from the gauge observations. Figure 12c displays the variation of the validation measures RMSE, MAE and correlation coefficient R als boxplots.
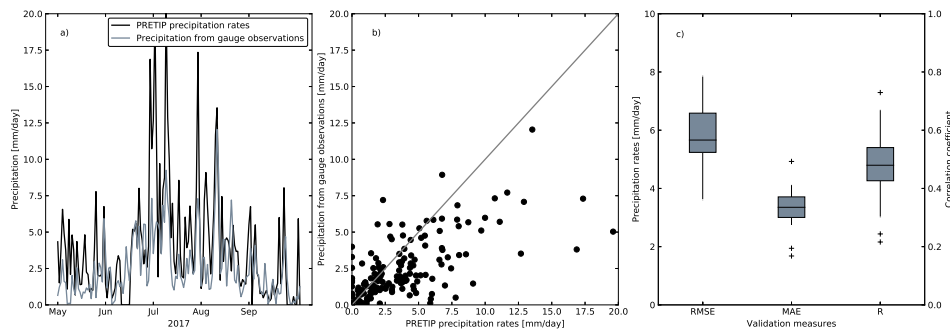

Figure 12. Temporal variation (**a**) and scatterplot (**b**) of daily precipitation from the Chinese rain gauge stations and the corresponding PRETIP precipitation pixels averaged for the complete study period. (**c**) shows the boxplots of the validation measures correlation coefficient R (right *y*-axis), MAE, and RMSE (both left *y*-axis) for the comparison of PRETIP in the 4 km resolution with the 28 Chinese rain gauge observations on the daily scale.

5. Discussion

The validation of the retrieval is overall promising; however, it has some limitations, which are mainly due to the used reference data set for RF model development. GPM IMERG is the best satellite based precipitation product available for the TiP [20,21]. However, MW does not work accurately over high elevation and snow/ice covered areas due to scattering effects [21]. Therefore, the training of the RF models is restricted to areas which are not masked out because of snow and ice [21]. Thus, our approach is neither used nor applicable over snow and ice covered areas. In addition, GPM IMERG is known to underestimate precipitation over the Tianshan Mountains and the TiP, since it underestimates precipitation at high altitudes [21].

PRETIP tends to overestimate the precipitation area (cf. [33]) and the precipitation amount as compared to the gauge calibrated MW precipitation from IMERG. On the one hand, the precipitation amount is overestimated due to the overestimation of the precipitation area, and on the other hand, it is overestimated possibly due to the undersampling methods which undersample the majority class (very low precipitation amount) and, therefore, impact the learning algorithm of RF. In general, the precipitation patterns are captured quite well. Additionally, retrieving precipitation in a region which is partly (semi-)arid and partly influenced by monsoon is challenging due to the large differences in the precipitation amount.

We found differences within the validation measures that are visible in all months (compare Figure 5). This spread originates from the scene based trained models, which work for some scenes well, and for others not. There are hardly any differences in the sample sizes of each month; however, September contains the least number of scenes (155 scenes), whereas the other months contain between 237 and 273 scenes per month. In general, we find similar validation measures and variations for all months.

Our precipitation product contains multiple spectral bands from two GEO satellites, whereas the IR only product only makes use of a single GEO band. A major drawback is that multispectral

information is not available globally or back in time to 1998, which is a limitation for products, such as IMERG IR only and PERSIANN. New GEO satellites with higher spectral resolution will improve the retrieval of precipitation from space. This offers great potential for the future. However, for climatological analyses of precipitation, the IR only methods are still needed, because only those are available from the older GEO satellites. We showed that our precipitation retrieval performs better than IMERG's IR only precipitation as compared to the gauge calibrated MW precipitation estimates. The correlation coefficient for the half hourly and 11 km resolution is 0.24 on average, and MAE and RMSE are 1.65 mm/h and 2.6 mm/h, respectively. This shows that the use of multispectral IR bands outperform IMERG's single band IR only precipitation.

Further, we compared our product with Chinese gauge measurements and found a correlation coefficient of R = 0.49 while using the 4 km resolution of PRETIP. The MAE is 3.3 mm/day and the RMSE is 5.3 mm/day on average regarding the 4 km resolution. The complex terrain of the TiP and associated orographic effects limit the representativity between single station data points and pixels that represent precipitation information for an area with a size of 4 km * 4 km [21].

Concerning the spatial distribution of the validation measures, we found that the low MAE and RMSE in the western TiP appear because of the low precipitation amount. In the central and eastern TiP, the precipitation amount is higher. The central and eastern part of the TiP are influenced by the Indian summer monsoon and the East Asian summer monsoon. Due to this atmospheric influence, these regions of the TiP receive more precipitation when compared to the dry west. In the eastern and central part of the TiP, the agreement between PRETIP and the Chinese rain gauge observations is stronger (high correlation coefficient, Rmax = 0.73). The western part of the TiP is influenced by the westerlies, where less precipitation occurs [47–50]. In the south of the TiP the MAE and RMSE increase. The temporal pattern between PRETIP and the Chinese gauge observations is captured quite well, but with some errors.

Bai & Liu [7] calculated the median value for the comparison between the satellite based precipitation products against daily gauge observations over the TiP and found rather low correlation coefficients for TMPA, CMORPH, PERSIANN-CDR and CHIRPS with an R between 0.17–0.18. Only MSWEP reached an R of 0.32. The average correlation coefficient for the comparison with the available stations for PRETIP is 0.49. However, the mean RMSE for the validation of PRETIP with gauge observations is 12.3 mm/day, whereas the RMSE for the other studies are between 3.6 mm/day and 5.4 mm/day. Gao & Liu [9] evaluated satellite based precipitation data sets over the TiP using gauge observations and divided them into different climate zones. The correlation coefficient and the RMSE both strongly depend on the climate zone, because the range of the validation measures is quite strong. TMPART performs worst with a correlation between 0.22–0.6 and an RMSE up to 14.95 mm/day. PERSIANN received an R of 0.27–0.48 and an RMSE of 2.74–10.36. TMPA and CMORPH perform better. For TMPA, they find an R between 0.32–0.67 and an RMSE between 2.63–8.93 mm/day. CMORPH performs best with an R of 0.4–0.67 and an RMSE of 2.12–8.6 mm/day. In general, the validation measures are comparable with the validation measures calculated between the daily gauge observations and PRETIP. Lu & Panmao [51] validate CMORPH and TMPA 3B42 over the TiP and find an R between 0.4–0.51, which is comparable to the mean correlation coefficient from the present study (R = 0.59). Here, the daily MAE is comparable low and ranges between 1–1.25 mm/day.

Figure 10a depicts the precipitation totals for the whole time period for gauge calibrated MW precipitation and shows the differences to PRETIP (Figure 10b) and the IR only precipitation (Figure 10c). PRETIP overestimates precipitation, especially in the south eastern part of the TiP, whereas IR only underestimates precipitation in the same region. All of the products show very low precipitation in the western part of the TiP. The spatial distribution of the validation measures do not differ much on average (Figure 10d–f), but a slight increase in MAE and RMSE in the eastern TiP are visible. This is due to the higher precipitation amount. In addition, Figure 10h,i show the average distribution of the validation measures with regard to PRETIP precipitation. The validation

measures increase with the increase in precipitation and show therefore a similar pattern of the precipitation retrieval.

Other studies also retrieved precipitation while using GEO satellite data over various regions. For example, Kühnlein et al. [31] received an MAE between 0.72–0.8 mm/h and an RMSE between 1.26 mm/h and 1.46 mm/h on an hourly scale in Germany. In another study where Kühnlein et al. [32] estimated precipitation from MSG SEVIRI for various times of the day they calculated an MAE between 0.52–0.99 mm/h and an RMSE of 0.89–1.94 mm/h for the hourly base with regard to the different times of the day (day, twilight, night) and different seasons (April–September, October–March) in Germany. Meyer et al. [28] mapped precipitation over southern Africa and gained an RMSE of 0.72 mm/h for the hourly basis. Please refer to Table 3 for comparison.

Some other studies retrieved precipitation over the TiP. Chao et al. [52] merged satellite based precipitation data from CMORPH with gauge observations using weighted regression methods over the eastern part of the TiP and computed a correlation of 0.86 and an RMSE of 7.3 mm/day. The Modified-IMSRA (M-IMSRA) algorithm of the Indian Meteorological Department (IMD) is restricted to the Indian subcontinent and does therefore not cover the TiP [53]. Shen et al. [54] produced a 1 km gauge-radar-satellite merged precipitation data set based on gauge observations, CMORPH and interpolation merging strategies. However, due to missing gauge and radar measurement, this method fails to capture precipitation over the TiP precisely. The authors find wide ranging correlation coefficients between 0.1–0.9 for the six hourly resolution and an RMSE up t 14 mm/6 h.

Several studies found that many satellite based precipitation failed to correctly capture precipitation over the TiP. Hussain et al. [55] questioned the use of CMORPH, TMPA, and PERSIANN precipitation data over the western, arid, and glacial part of the TiP. The overestimation of precipitation in all of the data sets is probably associated with radiation absorption over snow covered regions. They receive for their comparison of PERSIANN, TMPA, and CMRORPH on various elevation ranges (plain, mountain, glacial) correlation coefficients between 0.29 and 0.81, whereas the glacial parts perform the worst when compared to the plains and mountains. The RMSE draws a similar pattern: plains and mountains are comparable with large ranges across the data sets which range from 44–97 mm/month, and glacial receives on average for all data sets an RMSE of 60 mm/month. Yamamoto et al. [56] found that PERSIANN differed a lot from gauge observations especially in Winter and also contained a time lag of several hours. PERSIANN also seems to overestimate medium and high precipitation [57]. By contrast, CMORPH underestimates medium and high precipitation and showed significant biases in summer and winter [21,57].

Song et al. [58] stressed the large discrepancies among various data sets (satellite based, reanalysis) in the Himalayas and Karakoram. In addition, the coarse spatial and temporal resolution from satellite based precipitation products highlight the need of a high tempo spatial precipitation product for the TiP. Our product received convincing results with its high temporal resolution (30 min).

Table 3. Comparison of the validation measures of PRETIP with other studies.

Validation Measure	PRETIP	Kühnlein et al., 2014a	Kühnlein et al., 2014b	Meyer et al., 2017	Min et al., 2019
R/Rsq	0.59/0.35	-/0.36–0.52	-/0.48–0.5	-/0.33	-
MAE [mm/h]	0.91	0.52–0.99	0.72–0.8	-	0.51
RMSE [mm/h]	1.63	0.89–1.94	1.26–1.46	0.72	2.0

6. Conclusions

The lack of precipitation measurements from the ground highly limits the information density on the TiP. Here, we present a new approach for a retrieval method based on machine learning (Random Forest), which overcomes this data gap by retrieving precipitation rates from brightness temperatures from the two GEO satellites Insat-3D and Elektro-L2. We found that undersampling during the RF regression enhanced the learning process of the models and, therefore, leads to promising validation

results. Kolbe et al. [33] showed promising results of the precipitation area delineation; however, with a tendency to overestimate the areas. In combination with the results from the precipitation rates retrieval, we conclude that the complete product tends to overestimate precipitation due to the overestimation of the precipitation area. However, our new product PRETIP outperformed IMERG's IR only precipitation estimates, which, in general, tends to underestimate precipitation. In addition, we found a comparable convincing agreement with the Chinese rain gauge observations (Raverage = 0.49), especially at gauge stations with high precipitation amount (Rmax = 0.73). To further enhance the prediction of precipitation on the TiP, additional predictor variables derived from numerical weather predictions might be an asset. Here, we combine Elektro-L2 and Insat-3D for the period May–September 2017. Because Insat-3D is frequently available since 2014 it would be interesting to check whether the algorithm works as precisely when only Insat-3D data are used and, if so, PRETIP could be extended to longer time series.

Abbreviations

The following abbreviations are used in this manuscript:

APO	Average Probability of Occurrence
APRHODITE	Asian Precipitation—Highly-Resolved Observational Data Integration Towards Evaluation
CHIRPS	Climate Hazards Group InfraRed Precipitation with Station data
CMA	China Meteorological Administration
CMORPH	Climate Prediction Centre MORPHing technique
DEM	Digital Elevation Model
GEO	Geostationary
ECMWF	European Centre for Medium-range Weather Forecasts
GPCC	Global precipitation climatology center
GPM	Global Precipitation Measurement Mission
GSMaP-MVK	Global Satellite Mapping of Precipitation—Moving Vector with Kalman Filter
HE(M)	Hydro estimate (Method) from Insat-3D
IMERG	Integrated Multi-satellitE Retrievals for GPM
IMSRA	INSAT Multi-Spectral Rainfall Algorithm
IR	InfraRed
IR only	InfraRed only precipitation product
JAXA	Japan Aerospace Exploration Agency
LEO	Low Earth Orbit
MAE	Mean Absolute Error
MOSDAC	Meteorological and Oceanographic Satellite Data Archival Centre
MSG	Meteosat Second Generation
MSWEP	Multi-Source Weighted-Ensemble Precipitation
MW	Microwave
NASA	National Aeronautics and Space Administration
PERSIANN	Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks
PERSIANN-CDR	PERSIANN climate data record
PCV	Pseudo Cross Variogram
PrecipCal	Multi-satellite precipitation estimate with gauge calibration
PRETIP	Precipitation Retrieval covering the Tibetan Plateau
R	Correlation Coefficient
RF	Random Forest
RMSE	Root Mean Square Error
SEVIRI	Spinning Enhanced Visible and Infrared Imager
TiP	Tibetan Plateau
TMPA	TRMM Multi-satellite Precipitation Analysis
TMPART	TMPA Real Time version
TRMM	Tropical Rainfall Measuring Mission
VIS	Visible

References

1. Immerzeel, W.W.; van Beek, L.P.; Bierkens, M.F. Climate Change Will Affect the Asian Water Towers. *Science* **2010**, *328*, 1382–1385. [CrossRef]

2. Kidd, C.; Huffman, G.J. Global precipitation measurement. *Meteorol. Appl.* **2011**, *18*, 334–353. [CrossRef]

3. Hsu, K.L.; Gao, X.; Sorooshian, S.; Gupta, H.V. Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks. *J. Appl. Meteorol.* **1997**, *36*, 1176–1190. [CrossRef]

4. Huffman, G.J.; Bolvin, D.T.; Nelkin, E.J.; Wolff, D.B.; Adler, R.F.; Gu, G.; Hong, Y.; Bowman, K.P.; Stocker, E.F. The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. *J. Hydrometeorol.* **2007**, *8*, 38–55. [CrossRef]

5. Joyce, R.J.; Janowiak, J.E.; Arkin, P.A.; Xie, P. CMORPH: A Method that Produces Global Precipitation Estimates from Passive Microwave and Infrared Data at High Spatial and Temporal Resolution. *J. Hydrometeorol.* **2004**, *5*, 487–503. [CrossRef]

6. Huffman, G.J.; Bolvin, D.T.; Nelkin, E.J. Integrated Multi-satellitE Retrievals for GPM (IMERG) Technical Documentation. *NASA/GSFC Code* **2015**, *612*, 47.

7. Bai, P.; Liu, X. Evaluation of five satellite-based precipitation products in two gauge-scarce basins on the Tibetan Plateau. *Remote Sens.* **2018**, *10*, 1316. [CrossRef]

8. Lu, D.; Yong, B. Evaluation and hydrological utility of the latest GPM IMERG V5 and GSMaP V7 precipitation products over the Tibetan Plateau. *Remote Sens.* **2018**, *10*, 2022. [CrossRef]

9. Gao, Y.C.; Liu, M.F. Evaluation of high-resolution satellite precipitation products using rain gauge observations over the Tibetan Plateau. *Hydrol. Earth Syst. Sci.* **2013**, *17*, 837–849. [CrossRef]

10. Ma, L.; Zhao, L.; Tian, L.M.; Yuan, L.M.; Xiao, Y.; Zhang, L.L.; Zou, D.F.; Qiao, Y.P. Evaluation of the integrated multi-satellite retrievals for global precipitation measurement over the Tibetan Plateau. *J. Mt. Sci.* **2019**, *16*, 1500–1514. [CrossRef]

11. Tong, K.; Su, F.; Yang, D.; Zhang, L.; Hao, Z. Tibetan Plateau precipitation as depicted by gauge observations, reanalyses and satellite retrievals. *Int. J. Climatol.* **2014**, *34*, 265–285. [CrossRef]

12. Singh, A.K.; Singh, V.; Singh, K.K.; Tripathi, J.N.; Kumar, A.; Soni, A.K.; Sateesh, M.; Khadke, C. A Case Study: Heavy Rainfall Event Comparison Between Daily Satellite Rainfall Estimation Products with IMD Gridded Rainfall Over Peninsular India During 2015 Winter Monsoon. *J. Indian Soc. Remote Sens.* **2018**, *46*, 927–935. [CrossRef]

13. Reddy, M.V.; Mitra, A.K.; Momin, I.M.; Mitra, A.K.; Pai, D.S. Evaluation and inter-comparison of high-resolution multi-satellite rainfall products over India for the southwest monsoon period. *Int. J. Remote Sens.* **2019**, *40*, 4577–4603. [CrossRef]

14. Mitra, A.K.; Kaushik, N.; Kumar Singh, A.; Parihar, S.; Bhan, S.C. Evaluation of INSAT-3D satellite derived precipitation estimates for heavy rainfall events and its validation with gridded GPM (IMERG) rainfall dataset over the Indian region. *Remote Sens. Appl. Soc. Environ.* **2018**, *9*, 91–99. [CrossRef]

15. Chen, F.; Li, X. Evaluation of IMERG and TRMM 3B43 Monthly Precipitation Products over Mainland China. *Remote Sens.* **2016**, *8*, 472. [CrossRef]

16. Lu, X.; Wei, M.; Tang, G.; Zhang, Y. Evaluation and correction of the TRMM 3B43V7 and GPM 3IMERGM satellite precipitation products by use of ground-based data over Xinjiang, China. *Environ. Earth Sci.* **2018**, *77*, 209. [CrossRef]

17. Ma, Y.; Tang, G.; Long, D.; Yong, B.; Zhong, L.; Wan, W.; Hong, Y. Similarity and error intercomparison of the GPM and its predecessor-TRMM multisatellite precipitation analysis using the best available hourly gauge network over the Tibetan Plateau. *Remote Sens.* **2016**, *8*, 569. [CrossRef]

18. Wang, W.; Lu, H.; Zhao, T.; Jiang, L.; Shi, J. Evaluation and Comparison of Daily Rainfall From Latest GPM and TRMM Products Over the Mekong River Basin. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 2540–2549. [CrossRef]

19. Xu, R.; Tian, F.; Yang, L.; Hu, H.; Lu, H.; Hou, A. Ground validation of GPM IMERG and TRMM 3B42V7 rainfall products over Southern Tibetan plateau based on a high-density rain gauge network. *J. Geophys. Res.* **2017**, *122*, 910–924. [CrossRef]

20. Yang, M.; Li, Z.; Anjum, M.N.; Gao, Y. Performance Evaluation of Version 5 (V05) of Integrated Multi-Satellite Retrievals for Global Precipitation Measurement (IMERG) over the Tianshan Mountains of China. *Water* **2019**, *11*, 1139. [CrossRef]

21. Zhang, C.; Chen, X.; Shao, H.; Chen, S.; Liu, T.; Chen, C.; Ding, Q.; Du, H. Evaluation and Intercomparison of High-Resolution Satellite Precipitation Estimates—GPM, TRMM, and CMORPH in the Tianshan Mountain Area. *Remote Sens.* **2018**, *10*, 1543. [CrossRef]

22. Schneider, U.; Becker, A.; Ziese, M.; Rudolf, B. Global Precipitation Analysis Products of the GPCC. *Internet Publ.* **2011**, 1–13. [CrossRef]

23. Kidd, C.; Kniveton, D.R.; Todd, M.C.; Bellerby, T.J. Satellite Rainfall Estimation Using Combined Passive Microwave and Infrared Algorithms. *J. Hydrometeorol.* **2003**, *4*, 1088–1104. [CrossRef]

24. Kidd, C.; Levizzani, V. Status of satellite precipitation retrievals. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 1109–1116. [CrossRef]

25. Huffman, G.J.; Bolvin, D.T.; Nelkin, E.J. Integrated Multi-satellitE Retrievals for GPM (IMERG) Technical Documentation. *NASA/GSFC Code* **2018**, *612*, 47. [CrossRef]

26. Tan, J.; Petersen, W.A.; Tokay, A. A Novel Approach to Identify Sources of Errors in IMERG for GPM Ground Validation. *J. Hydrometeorol.* **2016**, *17*, 2477–2491. [CrossRef]

27. Yin, Z.Y.; Zhang, X.; Liu, X.; Colella, M.; Chen, X. An Assessment of the Biases of Satellite Rainfall Estimates over the Tibetan Plateau and Correction Methods Based on Topographic Analysis. *J. Hydrometeorol.* **2008**, *9*, 301–326. [CrossRef]

28. Meyer, H.; Drönner, J.; Nauss, T. Satellite-based high-resolution mapping of rainfall over southern Africa. *Atmos. Meas. Tech.* **2017**, *10*, 2009–2019. [CrossRef]

29. Min, M.; Bai, C.; Guo, J.; Sun, F.; Liu, C.; Wang, F.; Xu, H.; Tang, S.; Li, B.; Di, D.; et al. Estimating Summertime Precipitation from Himawari-8 and Global Forecast System Based on Machine Learning. *IEEE Trans. Geosci. Remote Sens.* **2019**, *57*, 2557–2570. [CrossRef]

30. Upadhyaya, S.; Ramsankaran, R. Modified-INSAT Multi-Spectral Rainfall Algorithm (M-IMSRA) at climate region scale: Development and validation. *Remote Sens. Environ.* **2016**, *187*, 186–201. [CrossRef]

31. Kühnlein, M.; Appelhans, T.; Thies, B.; Nauss, T. Improving the accuracy of rainfall rates from optical satellite sensors with machine learning—A random forests-based approach applied to MSG SEVIRI. *Remote Sens. Environ.* **2014**, *141*, 129–143. [CrossRef]

32. Kühnlein, M.; Appelhans, T.; Thies, B.; Nauss, T. Precipitation estimates from MSG SEVIRI daytime, nighttime, and twilight data with random forests. *J. Appl. Meteorol. Climatol.* **2014**, *53*, 2457–2480. [CrossRef]

33. Kolbe, C.; Thies, B.; Egli, S.; Lehnert, L.; Schulz, H.; Bendix, J. Precipitation Retrieval over the Tibetan Plateau from the Geostationary Orbit—Part 1: Precipitation Area Delineation with Elektro-L2 and Insat-3D. *Remote Sens.* **2019**, *11*, 2302. [CrossRef]

34. Skofronick-Jackson, G.; Huffman, G.; Stocker, E.; Petersen, W. Successes with the Global Precipitation Measurement (GPM) mission. In Proceedings of the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, China, 10–15 July 2016.

35. Beusch, L.; Foresti, L.; Gabella, M.; Hamann, U. Satellite-based rainfall retrieval: From generalized linear models to artificial neural networks. *Remote Sens.* **2018**, *10*, 939. [CrossRef]

36. Indian Space Research Organisation. *Insat-3D. India's Advanced Weather Satellite*; Technical Report; Indian Space Research Organisation: Bangalore, India, 2013.

37. National Satellite Meteorological Centre. *Insat-3D Data Products Catalog*; Technical Report; National Satellite Meteorological Centre: New Delhi, India, 2014.

38. News, S.P. INSAT-3D Now Operational. 2014. Available online: https://www.wmo-sat.info/vlab/2014/02/insat-3d-now-operational/ (accessed on 29 June 2020).

39. MOSDAC, Space Applications Centre, Indian Space Research Organisation, Govt. of India. Meteorological & Oceanographic Satellite Data Archival Centre. Available online: https://www.mosdac.gov.in/ (accessed on 29 June 2020).

40. Zak, A. Zenit Delivers Elektro-L2. 2016. Available online: http://www.russianspaceweb.com/elektro-l2.html (accessed on 29 June 2020).

41. Wijaya, A.; Marpu, P.; Gloaguen, R. Geostatistical Texture Classification of Tropical Rainforest in Indonesia. In *Quality Aspects in Spatial Data Mining*; Number 1; CRC Press: Freiberg, Germany, 2008; pp. 199–210. [CrossRef]

42. ORNL DAAC. *Digital Elevation Model: Global 30 Arc-Second Elevation Data Set (GTOPO30)*; 2017. Available online: https://webmap.ornl.gov/ogc/dataset.jsp?dg_id=10003_1 (accessed on 29 June 2020).

43. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

44. Breiman, L. Bagging predictors. *Mach. Learn.* **1996**, *24*, 123–140.10.1007/BF00058655. [CrossRef]

45. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.

46. Rüthrich, F.; Reudenbach, C.; Thies, B.; Bendix, J. Lake Related Cloud-Dynamics on the Tibetan Plateau: Spatial Patterns and Interannual Variability. *J. Clim.* **2015**, *28*, 9080–9104. [CrossRef]

47. Bookhagen, B.; Burbank, D.W. Toward a complete Himalayan hydrological budget: Spatiotemporal distribution of snowmelt and rainfall and their impact on river discharge. *J. Geophys. Res. Earth Surf.* **2010**, *115*, 1–25. [CrossRef]

48. Bothe, O.; Fraedrich, K.; Zhu, X. Precipitation climate of Central Asia and the large-scale atmospheric circulation. *Theor. Appl. Climatol.* **2012**, *108*, 345–354. [CrossRef]

49. Curio, J.; Maussion, F.; Scherer, D. A 12-year high-resolution climatology of atmospheric water transport over the Tibetan Plateau. *Earth Syst. Dyn.* **2015**, *6*, 109–124. [CrossRef]

50. Mölg, T.; Maussion, F.; Scherer, D. Mid-latitude westerlies as a driver of glacier variability in monsoonal High Asia. *Nat. Clim. Chang.* **2014**, *4*, 68–73. [CrossRef]

51. Wu, L.; Zhai, P. Validation of daily precipitation from two high-resolution satellite precipitation datasets over the Tibetan Plateau and the regions to its east. *Acta Meteorol. Sin.* **2012**, *26*, 735–745. [CrossRef]

52. Chao, L.; Zhang, K.; Li, Z.; Zhu, Y.; Wang, J.; Yu, Z. Geographically weighted regression based methods for merging satellite and gauge precipitation. *J. Hydrol.* **2018**, *558*, 275–289. [CrossRef]

53. Mishra, A.K.; Gairola, R.M.; Varma, A.K.; Agarwal, V.K. Improved rainfall estimation over the Indian region using satellite infrared technique. *Adv. Space Res.* **2011**, *48*, 49–55. [CrossRef]

54. Shen, Y.; Hong, Z.; Pan, Y.; Yu, J.; Maguire, L. China's 1 km Merged Gauge, Radar and Satellite Experimental Precipitation Dataset. *Remote Sens.* **2018**, *10*, 264. [CrossRef]

55. Hussain, Y.; Satgé, F.; Hussain, M.B.; Martinez-Carvajal, H.; Bonnet, M.P.; Cárdenas-Soto, M.; Roig, H.L.; Akhter, G. Performance of CMORPH, TMPA, and PERSIANN rainfall datasets over plain, mountainous, and glacial regions of Pakistan. *Theor. Appl. Climatol.* **2018**, *131*, 1119–1132. [CrossRef]

56. Yamamoto, M.K.; Ueno, K.; Nakamura, K. Comparison of Satellite Precipitation Products with Rain Gauge Data for the Khumb Region, Nepal Himalayas. *J. Meteorol. Soc. Jpn.* **2011**, *89*, 597–610. [CrossRef]

57. Alazzy, A.A.; Lü, H.; Chen, R.; Ali, A.B.; Zhu, Y.; Su, J. Evaluation of Satellite Precipitation Products and Their Potential Influence on Hydrological Modeling over the Ganzi River Basin of the Tibetan Plateau. *Adv. Meteorol.* **2017**, *2017*, 1–23. [CrossRef]

58. Song, C.; Huang, B.; Ke, L.; Ye, Q. Precipitation variability in High Mountain Asia from multiple datasets and implication for water balance analysis in large lake basins. *Glob. Planet. Chang.* **2016**, *145*, 20–29. [CrossRef]