



Downscaling Daily Satellite-Based Precipitation Estimates Using MODIS Cloud Optical and Microphysical Properties in Machine-Learning Models

Sergio Callaú Medrano, Frédéric Satgé, Jorge Molina-Carpio, Ramiro Pillco Zolá, Marie-Paule Bonnet

► To cite this version:

Sergio Callaú Medrano, Frédéric Satgé, Jorge Molina-Carpio, Ramiro Pillco Zolá, Marie-Paule Bonnet. Downscaling Daily Satellite-Based Precipitation Estimates Using MODIS Cloud Optical and Microphysical Properties in Machine-Learning Models. *Atmosphere*, 2023, 14 (9), pp.1349. 10.3390/atmos14091349 . hal-04375770

HAL Id: hal-04375770

<https://hal.science/hal-04375770>




Submitted on 5 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Article

Downscaling Daily Satellite-Based Precipitation Estimates Using MODIS Cloud Optical and Microphysical Properties in Machine-Learning Models

Sergio Callaú Medrano ^{1,2,*} , Frédéric Satgé ^{1,2,*}, Jorge Molina-Carpio ² , Ramiro Pillco Zolá ² and Marie-Paule Bonnet ¹ 

¹ Espace-Dev Univ Montpellier, IRD, Univ Antilles, Univ Guyane, Univ Réunion, 34093 Montpellier, France; marie-paule.bonnet@ird.fr

² Instituto de Hidráulica e Hidrología, Universidad Mayor de San Andrés, La Paz 15000, Bolivia; jamolinacarpio@gmail.com (J.M.-C.); rami_lund99@hotmail.com (R.P.Z.)

* Correspondence: callau.s.m@gmail.com (S.C.M.); frederic.satge@ird.fr (F.S.)

Abstract: This study proposes a method for downscaling the spatial resolution of daily satellite-based precipitation estimates (SPEs) from 10 km to 1 km. The method deliberates a set of variables that have close relationships with daily precipitation events in a Random Forest (RF) regression model. The considered variables include cloud optical thickness (COT), cloud effective radius (CER) and cloud water path (CWP), derived from MODIS, along with maximum and minimum temperature (Tx, Tn), derived from CHIRTS. Additionally, topographic features derived from ALOS-DEM are also investigated to improve the downscaling procedure. The approach consists of two main steps: firstly, the RF model training at the native 10 km spatial resolution of the studied SPEs (i.e., IMERG) using rain gauge observations as targets; secondly, the application of the trained RF model at a 1 km spatial resolution to downscale IMERG from 10 km to 1 km over a one-year period. To assess the reliability of the method, the RF model outcomes were compared with the rain gauge records not considered in the RF model training. Before the downscaling process, the CC, MAE and RMSE metrics were 0.32, 1.16 mm and 6.60 mm, respectively, and improved to 0.48, 0.99 mm and 4.68 mm after the downscaling process. This corresponds to improvements of 50%, 15% and 29%, respectively. Therefore, the method not only improves the spatial resolution of IMERG, but also its accuracy.

Keywords: IMERG; downscaling; random forest model; MODIS cloud optical and microphysical properties; CHIRTS



Citation: Medrano, S.C.; Satgé, F.; Molina-Carpio, J.; Zolá, R.P.; Bonnet, M.-P. Downscaling Daily Satellite-Based Precipitation Estimates Using MODIS Cloud Optical and Microphysical Properties in Machine-Learning Models. *Atmosphere* **2023**, *14*, 1349. <https://doi.org/10.3390/atmos14091349>

Academic Editors: Xiaoming Shi, Berry Wen and Lisa Milani

Received: 25 June 2023

Revised: 18 August 2023

Accepted: 19 August 2023

Published: 27 August 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In the context of climate change, precipitation patterns around the world have suffered spatial and temporal changes due to the rising temperatures [1–3]. In order to assess the impacts of changes in precipitation on the dynamic processes of water resources, ecosystems and society, it is desirable to obtain accurate precipitation estimations [4–6]. Traditionally, these estimations have been obtained from rain gauge records. However, rain gauge measurements provide a limited characterization of precipitation, since each station monitors only a small area of the studied region. Therefore, in remote areas, where the rain gauge networks are sparse and unevenly distributed [7,8], their application in the analysis of the spatio-temporal heterogeneity of precipitation entails a certain degree of uncertainty [9,10].

Alternatively, satellite-based precipitation estimates (SPEs) have the capacity to estimate precipitation continuously around the globe on regular grids. In this context, SPEs have become an alternative source of information for various water-related applications, particularly for regions encompassing poorly monitored catchments [11–13]. For example, SPEs have been used for flash flood monitoring [14–16], streamflow simulations [13,17,18],

drought monitoring [19,20] and precipitation monitoring [21–23]. The aforementioned investigations demonstrated that, due to their coarse spatial resolution (i.e., several kilometers), SPEs do not provide an effective source of data for local-scale applications, but do for large-scale applications. Indeed, SPEs' low spatial resolution prevents an optimal consideration of climate and/or topographic characteristics influencing precipitation at the local scale.

To overcome this limitation, different studies have reported spatial downscaling procedures. Two techniques have been developed to yield better spatial resolution in SPEs: (i) dynamical and (ii) statistical downscaling.

The dynamic downscaling approach is based on regional climate models, which consist of high-resolution models nested in low-resolution global climate models. This technique relies on deterministic weather models that provide high-resolution precipitation estimates and other climate variables by simulating physical processes of the land–atmosphere system. This approach is known to be limited by the substantial computational resources it requires, as well as its sensitivity to the boundary conditions obtained from global climate models [24–26].

The statistical downscaling approach considers a regression analysis between environmental variables and precipitation estimates. The environmental variables involve high-resolution auxiliary datasets as the independent variables of the regression analysis. The precipitation data (i.e., the dependent variable in the regression analysis) are characterized by their low resolution. The resolution of the precipitation dataset is downscaled to a finer resolution using the auxiliary data products through the regression procedure. This data-driven downscaling method requires a large amount of data in order to establish statistical relationships and, unlike the dynamical methods, has lower computational demands [12,26–28].

In the statistical downscaling approach, nonlinear regression methods such as artificial neural networks (ANNs), random forests (RF) and support vector machine (SVM) have become feasible alternatives for downscaling low-resolution datasets due to their high potential for complex and nonlinear input–output mapping [12,24].

Easily accessible through satellite datasets, topography features, land surface temperature (LST), and normalized difference vegetation index (NDVI) are generally used as auxiliary data in statistical downscaling techniques [29,30]. The use of these auxiliary data in statistical downscaling techniques often outperforms simpler interpolation methods (e.g., bilinear interpolation) [31].

Nevertheless, at fine temporal scales (i.e., daily), the use of NDVI can become a source of uncertainty, since the interaction between precipitation and this environmental variable is weak. Similarly, LST estimates are generally unavailable for the days with precipitation events, as cloud cover prevents LST measurement being performed from satellites. Therefore, these approaches are mainly used to downscale precipitation at the monthly time step [30–32].

One alternative to overcome the limitations of using auxiliary data that have weak interactions with precipitation at finer temporal scales (i.e., daily) is the use of cloud optical and microphysical property datasets. Indeed, clouds are a prerequisite for the occurrence of precipitation [8], and precipitation is a key factor that controls the amount of clouds on the globe [33]. Campos Braga et al. [34], and Kobayashi and Masuda [33,35], studied the relationship between cloud optical and microphysical properties, such as cloud optical thickness (COT) and cloud effective radius (CER), and precipitation. They found that the CER was a crucial parameter for establishing the initiation and amount of precipitation at the top of convective clouds. Moreover, the changes in COT resulted from changes in the size distribution of cloud droplets which are associated with precipitation.

In light of the context described above, in this study, the potential of cloud optical and microphysical properties, as observed by MODIS, to enhance the spatial resolution of SPEs across the Ramis River basin, the most important tributary of Lake Titicaca, located in the South American Andean Plateau, is investigated. To overcome the computational

demand limitations of dynamic downscaling methods, an RF regression model is selected to implement a statistical downscaling approach. MODIS cloud optical and microphysical properties have been previously considered for SPE downscaling for isolated extreme daily precipitation events in central Europe [8]. Therefore, this study will explore the possibility of downscaling continuous daily precipitation for an entire year (i.e., 2012) to provide a step forward toward the construction of continuous daily precipitation observations at high spatial resolution.

2. Data

2.1. Study Area

The Ramis catchment is located in Peru to the northwest of Lake Titicaca, between latitude $14^{\circ}03' \text{ S}$ and $15^{\circ}27' \text{ S}$ and longitude $71^{\circ}07' \text{ W}$ and $69^{\circ}25' \text{ W}$, and is the largest tributary of Lake Titicaca [36] (Figure 1). The catchment area is estimated at $15,671 \text{ km}^2$ with an elevation ranging between 3810 m and 5740 m, and a mean streamflow of $73 \text{ m}^3 \cdot \text{s}^{-1}$ [37]. The catchment has a semi-arid–sub-humid climate, with a mean annual precipitation of about 700 mm [38]. The hydrological year is divided in two seasons: (i) the wet season, which is about 5 months long (November to March); and (ii) the dry season, which is about 7 months long (April to October) [6].

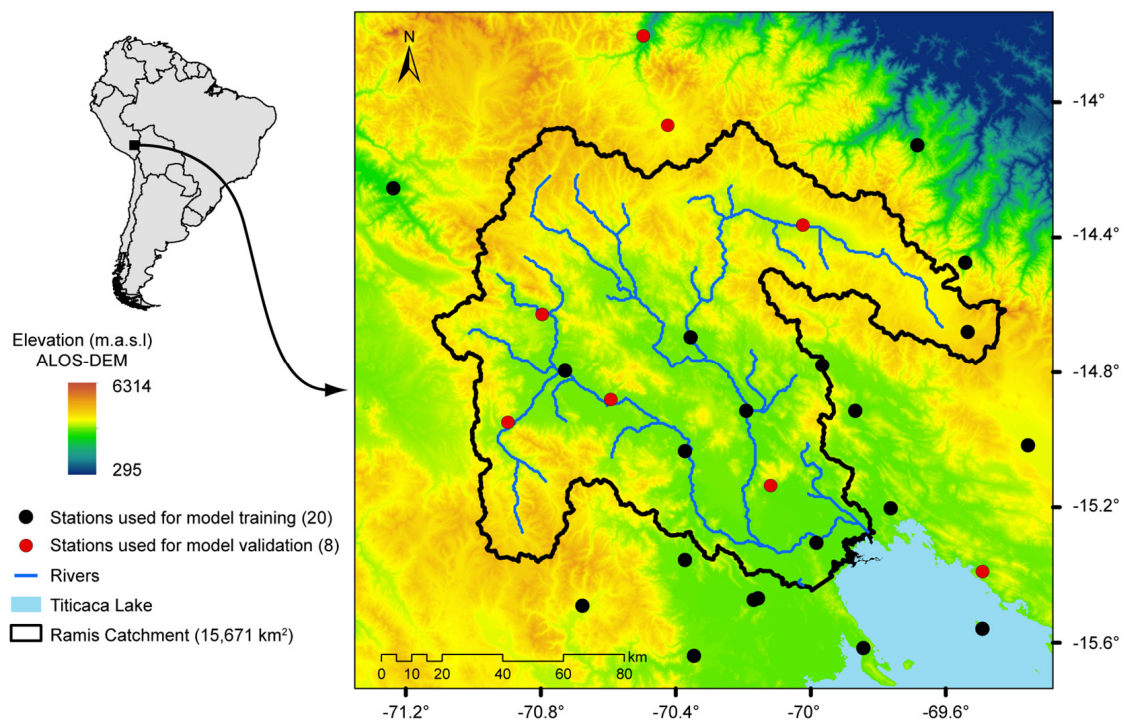


Figure 1. Geographic location and rainfall monitoring network of the Ramis catchment.

The Ramis catchment is the largest tributary to Lake Titicaca [36] which, as the largest freshwater lake in South America, represent a crucial water resource in the central Andes [39]. Along with climate warming, anthropogenic pressure on Ramis water resources has increased in recent decades due to population growth and increased evapotranspiration losses due to agricultural activity [40,41]. Furthermore, the recent extreme drought in 2022–23 in the Lake Titicaca basin also highlights the need for accurate precipitation information for sustainable water resources management. Moreover, the Ramis catchment has a rain gauge network that is denser than that of most of the Plateau, facilitating the calibration and validation of SPEs in order to extend the methodology and results to neighboring regions with similar climate contexts.

2.2. Rain Gauges

The *Servicio Nacional de Meteorología e Hidrología* (SENAMHI) in Peru is in charge of the meteorological network in the study region. Fifty-eight ground stations, collecting data on a daily basis from 1975 to 2017, were available in the study area and subjected to quality control [37]. However, many stations suffer from missing values in their temporal series from few days to complete years. In this context, we pre-selected stations with daily records for the period 2000–present ($n = 31$), which is the observation period of the considered SPE (IMERG). Then, an analysis of the precipitation records identified 2012 as the year with the highest number of operational stations ($n = 28$) with more than 80% of daily records available. Sixteen out of the twenty-eight stations were located outside the Ramis catchment, but were used to increase the reference observations, thus extending the study area (Figure 1). Based on the 28 rain gauges, the regional annual precipitation in 2012 was determined to be 822 mm, with more than 60% of the precipitation amount falling in January, February and December.

2.3. Satellite-Based Precipitation Estimates: IMERG

The Global Precipitation Measurement (GPM) mission was launched in February, 2014, as the successor of the Tropical Rainfall Measuring Mission (TRMM) [31,32]. The GPM improves the spatial resolution (from 0.25° to 0.1°), coverage (from 50° N– 50° S to 60° N– 60° S) and the sensitivity when detecting light and solid precipitation with respect to the TRMM [42]. The Integrated Multi-satellite Retrievals for GPM (IMERG) datasets are available in three different products: IMERG Early-Run, Late-Run and Final-Run. At the beginning, IMERG precipitation datasets were only derived from GPM satellite observations, and were therefore only available for the 2014–present period. However, in 2020, the algorithm used to retrieve IMERG precipitation estimates was updated to fuse observations from both the TRMM (2000–2015) and GPM (2014–present) satellites in order to obtain long-term precipitation time series from 2000 to present.

IMERG Final-Run includes a gauge-based adjustment that generally provides more reliable precipitation than its Early- and Late-Run counterparts [43]. Therefore, in the current investigation, we adopted the daily product of the IMERG Final-Run, which will hereafter referred to as IMERG. It is worth mentioning here that IMERG daily records are obtained as the sum of 48 half-hour records for the considered day (00:00 to 24:00, GMT). IMERG data were downloaded from NASA (2022) (Table 1).

Table 1. Datasets used in the study.

Variable Type	Product	Spatial Resolution	Temporal Resolution	Source
Independent	Rain gauge observations	-	Daily	https://www.senamhi.gob.pe/ , accessed on 24 June 2023
Dependent	IMERG	10 km	Daily	https://doi.org/10.1007/978-3-030-24568-9_19 , accessed on 24 June 2023
	MODIS (COT, CER and CWP)	1 km	Daily	https://doi.org/10.1109/TGRS.2002.808301 , accessed on 24 June 2023
	CHIRTS (Tx, Tn)	5 km	Daily	https://doi.org/10.15780/G2008H , accessed on 24 June 2023
	ALOS DEM (Elevation, Slope)	30 m	-	https://doi.org/10.3390/rs12071156 , accessed on 24 June 2023 Derived from ASTER GDEM

2.4. Satellite Cloud Cover Properties: MODIS

The Moderate-Resolution Imaging Spectroradiometer (MODIS) instruments are part of the Terra and Aqua platforms. Both platforms have a sun-synchronous near-polar orbit and have a global coverage of every 1 to 2 days [44]. Having twin MODIS instruments makes it possible to improve cloud imaging during the morning and afternoon [8]. Near the equator, Terra acquires cloud data at about 10:30 and 22:30 (local solar time), whereas Aqua acquires cloud data at about 15:30 and 01:30 (local solar time) during a five-minute period. Their Level 2 products provide cloud optical and microphysical properties with spatial resolutions of 1 and 5 km [45], which are observed by Terra (MOD06) and Aqua (MYD06) during both daytime (10:30 and 15:30 for Terra and Aqua, respectively) and nighttime (22:30 and 01:30 for Terra and Aqua, respectively). The cloud optical and microphysical property variables used in the downscaling methods are (i) cloud optical thickness (COT), (ii) cloud effective radius (CER) and (iii) cloud water path (CWP), at a spatial resolution of 1 km. The MOD06 and MYD06 products were pre-processed using the software HEGTOOL [46] in order to obtain COT, CER and CWP datasets with a spatial resolution of 1 km in a suitable computer file format that stores raster graphics (e.g., GeoTIFF) for the year 2012.

For each day, we calculated the mean COT, CER and CWP, as observed by MOD06 and MYD06, before using them in the downscaling procedure (Table 1). As most of the MOD06 and MYD06 nighttime records are missing, we only consider daytime observations in the daily average process to ensure consistency in the temporal series. As both convective cloudiness and the associated rainy episodes in the Andean Plateau occur mostly in the afternoon and evening [47], MOD06 and MYD06 daytime acquisitions look very adequate for estimating precipitation in this region.

2.5. Satellite-Based Temperature Estimates: CHIRTS

The Climate Hazards Group InfraRed Temperature with Station (CHIRTS), developed by the Climate Hazards Center (CHC) at the University of California Santa Barbara and the US Geological Survey (USGS) [48], is a satellite-based product with daily air temperature data. Its coverage ranges from 60° S to 70° N with a spatial resolution of $0.05^\circ \times 0.05^\circ$ (approximately 5 km \times 5 km). CHIRTS daily maximum and minimum 2 m air temperature (T_x , T_n , respectively) for 2012 were used. CHIRTS relies on temperature estimates derived from (1) thermal infra-red observations (TIR from GridSat), (2) reanalysis datasets (MERRA2 and ERA5), and (3) meteorological stations. The presence of clouds inevitably affects TIR observations. However, the TIR observations used to produce CHIRTS were previously cloud-screened to limit cloud interference [49]. Moreover, the consideration of other temperature datasets that are not directly influenced by clouds (i.e., reanalysis and meteorological stations) tend to decrease CHIRTS sensitivity to cloud cover interference even more on TIR observations. Finally, it is worth mention here that in comparison to other gridded temperature datasets (CPC, EWEMBI, ERA5, PGF and CPC), CHIRTS provided the most reliable daily temperature estimates across the study region [50]. For more information about CHIRTS daily temperature data, please refer to Funk et al. and Verdin et al. [49,51].

The use of this satellite-based temperature information could be advantageous in downscaling SPEs since the variability of the daily air temperature is linked to precipitation events. For instance, the moisture held by air increases with rising temperatures. As a result, precipitation events may become more intense [52]. This trend is in accordance with the Clausius–Clapeyron relation, which states that if relative humidity remains constant, atmospheric humidity will increase at a rate that follows the dependency of vapor saturation on temperature. This is an important consideration for understanding changes in rainfall events [53].

2.6. Digital Elevation Model: ALOS-DEM

Topography is another variable that could have an influence on the atmospheric circulation and precipitation patterns. Precipitation gradients with elevation are common, but they are usually nonlinear (e.g., [54]). The impact of topography on precipitation is greater if a source of moisture (e.g., the Titicaca Lake) is located near the analyzed region [55]. Furthermore, the relationship between precipitation and topography is frequently employed in the statistical downscaling of SPEs [31]. We used the Advanced Land Observing Satellite (ALOS) Digital Elevation Model (DEM) with a resolution of $30\text{ m} \times 30\text{ m}$ [56] to retrieve elevation and terrain slope (Table 1).

3. Methods

3.1. Machine Learning Model: Random Forest

As stated in the introduction, there are two approaches for downscaling SPEs: dynamic and statistical. Due to the substantial computational demand required by dynamic techniques [24,25], the application of a statistical approach (e.g., random forest models—RF) was deemed appropriate to downscale the daily SPEs (i.e., IMERG) in the current study.

RF is an ensemble-learning algorithm composed of several tree predictors. Each tree predictor or decision tree is constructed randomly from a collection of training data [57]. RF models have the advantage of having low rates of overfitting problems, because the prediction is carried out by averaging all decision trees [9]. Moreover, if a large amount of data is to be processed, RF has the advantage of being able to train the model fast and efficiently [26]. In addition, the RF's meta-parameters can be modified in order to optimize the model outcomes. In this study, we used the scikit-learn ensemble RF regressor and the scikit-learn Hyper-parameters tuning estimator [58] to train and optimize the RF models, respectively.

3.2. Downscaling Setup

Figure 2 illustrates the two-stage procedure followed to train and apply the RF model for downscaling IMERG at 1 km resolution.

The first stage consists of training the RF model at the native spatial resolution of IMERG (i.e., 10 km). To do so, the model inputs (or regressors) IMERG, CHIRTS (T_x , T_n), ALOS-DEM (elevation, slope) and MODIS (COT, CER, CWP) were upscaled to 10 km using a bilinear resampling method. The RF model targets entail the rainfall observations of 20 of the 28 rain gauges randomly selected (Figure 1). During RF model training, the model inputs were shaped only in the 20 cells corresponding to the locations of the 20 rain gauges. The precipitation records of the remaining 8 stations were applied to verify the downscaling outcomes.

The second stage consists of the application of the trained RF model at 1 km spatial resolution. In order to obtain the corresponding outcomes, the model inputs IMERG, CHIRTS (T_x , T_n) and ALOS-DEM (elevation, slope) were resampled to 1 km using bilinear resampling methods. No resampling process was applied to MODIS (COT, CER, and CWP), as its native resolution is 1 km. The model outcomes of the second stage consist of 366 matrices of cells (i.e., the 366 days of the year 2012) that cover the study area and have a spatial resolution of 1 km.

In order to assess the benefits of the different regressors, three RF models were contemplated, which will be called scenarios from now on. Each scenario uses different model inputs (Figure 2). Scenario-1 considers IMERG and MODIS (COT, CER and CWP); scenario-2 considers IMERG, MODIS (COT, CER and CWP) and ALOS-DEM (elevation, slope); and scenario-3 considers IMERG, MODIS (COT, CER and CWP), ALOS-DEM (elevation, slope) and CHIRTS (T_x , T_n). The performance of each scenario is verified considering (i) the 20 pixels containing the rain gauges used for RF model training (i.e., the training dataset), and (ii) the 8 pixels containing the rain gauges not used for RF model training (i.e., the validation dataset).

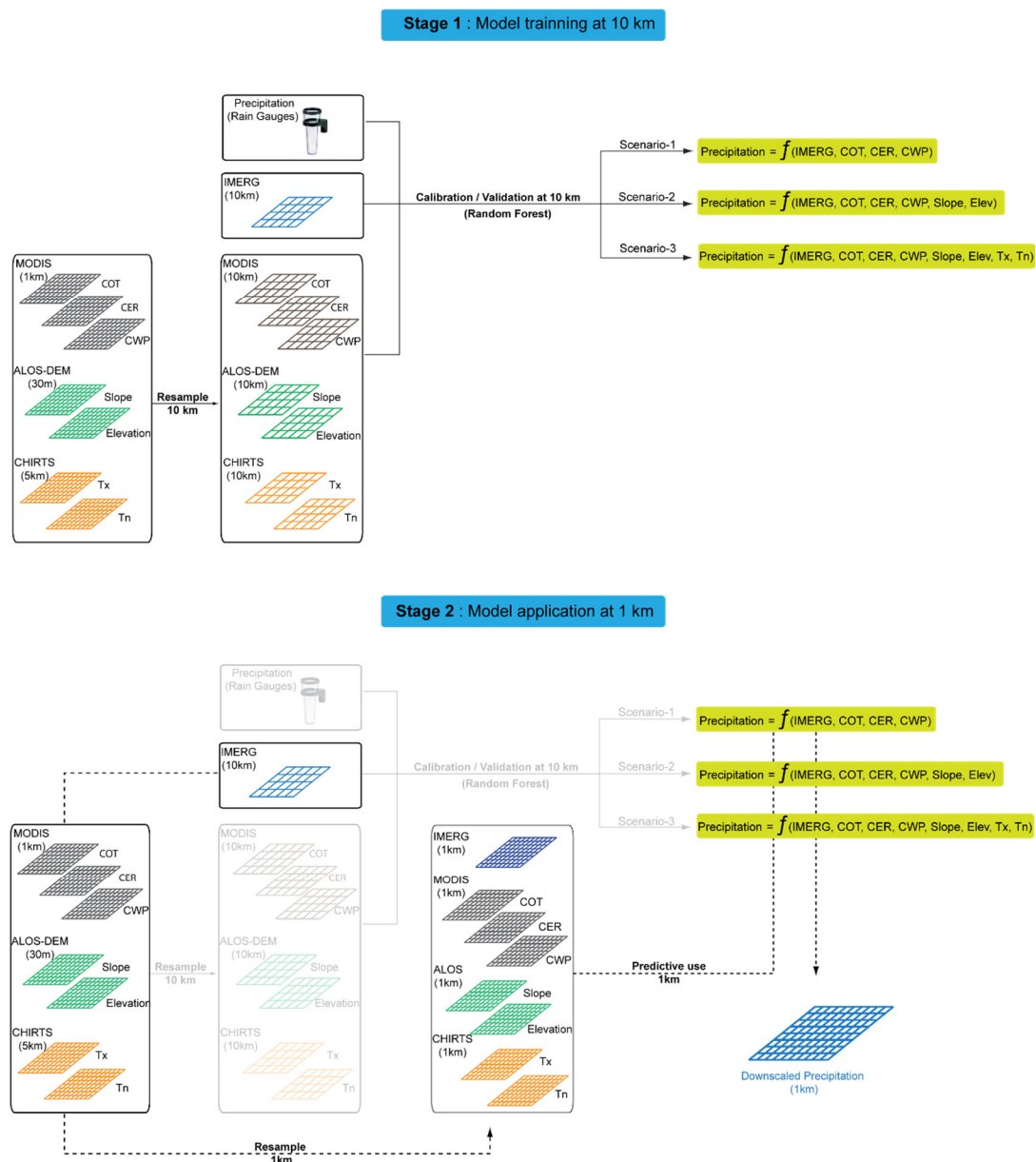


Figure 2. Flowchart of the procedures followed to downscale the spatial resolution of the SPEs.

In addition, IMERG at its native spatial resolution (i.e., 10 km) and at a spatial resolution of 1 km (after the application of a bilinear interpolation) are used as benchmarks to evaluate the different scenarios. Hereafter, these benchmarks will be referred to as IMERG-10 km and IMERG-1 km, respectively.

As it could be argued that using the precipitation time series of neighboring stations could lead to overly optimistic results, as those series are to be correlated, we tested and validated alternative RF models using different time periods rather than different stations. Their inputs were the same datasets applied in scenarios 1, 2 and 3, randomly distributed into a training set consisting of 75% of the observations and a validation set consisting of 25%. These models yielded similar metrics as the proposed methodology. Hence, we chose to keep the proposed method in the next stages of the study.

3.3. Performance Evaluation Analysis

3.3.1. Quantitative Metrics

The scenarios' outcomes and the benchmarks were evaluated according to three quantitative metrics: (i) the correlation coefficient (CC) (Equation (1)); (ii) the mean absolute error (MAE) (Equation (2)); and (iii) the root mean square error (RMSE) (Equation (3)) [26].

$$CC = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2 \sum_{i=1}^N (y_i - \bar{y})^2}} \quad (1)$$

$$MAE = \frac{\sum_{i=1}^N |y_i - x_i|}{N} \quad (2)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (x_i - y_i)^2}{N}} \quad (3)$$

where x_i and y_i are the values of the scenarios' outcomes or benchmarks and the rain gauge observations, respectively, N is the length of the overall data, and, \bar{x} and \bar{y} are the average values of the scenarios' outcomes or benchmarks and rain gauge observations.

For each scenario, CC, MAE and RMSE were computed after the application of the RF models at 1 km considering (i) the 20 pixels including the rain gauges used for the RF model training (training dataset), and (ii) the 8 pixels including the rain gauges not used for RF model training (validation dataset).

3.3.2. Categorical Metrics

Categorical metrics were considered to assess the capability of the RF models to detect daily precipitation events. The categorical metrics consider precipitation as a discrete value with four possible situations: both the scenario (or benchmark) and the rain gauge detect a precipitation event (a), only the scenario (or benchmark) reports a precipitation event (b), only the rain gauge detects a precipitation event (c), or neither the scenario (or benchmark) nor the rain gauge report a precipitation event (d). To avoid the inconsistencies that could be produced by light precipitation events, a threshold of 1 mm.day⁻¹ was used in the precipitation estimates to separate rainy from non-rainy days.

The metrics considered in the current work are: (i) the probability of detection (POD); (ii) the false alarm ratio (FAR); and (iii) the critical success ratio (CSI). POD assesses the ability of the scenario (or benchmark) to forecast precipitation events (Equation (5)), whereas FAR evaluates how often the scenario (or benchmark) detects an event when it should not do so (Equation (6)). POD and FAR values vary from 0 to 1, with perfect scores being 1 and 0 for POD and FAR, respectively. CSI computes the ratio between the number of precipitation events correctly detected by the scenario (or benchmark) and the total number of precipitation events (Equation (6)). Values vary from 0 to 1, with a perfect score being 1.

To facilitate the inter-comparison of the scenarios and the benchmarks, the results are presented in the form of a performance diagram [59]. Performance diagrams have previously been used for daily precipitation analysis in Brazil, Bolivia and Pakistan [19,60].

$$POD = \frac{a}{a + c} \quad (4)$$

$$FAR = \frac{b}{a + b} \quad (5)$$

$$CSI = \frac{a}{a + b + c} \quad (6)$$

4. Results and Discussion

4.1. Evaluation of Results from Quantitative Indices

Figure 3 shows the median statistical metrics (i.e., CC, MAE and RMSE) obtained after the application of the RF model (Stage 2), considering both the training and validation datasets.

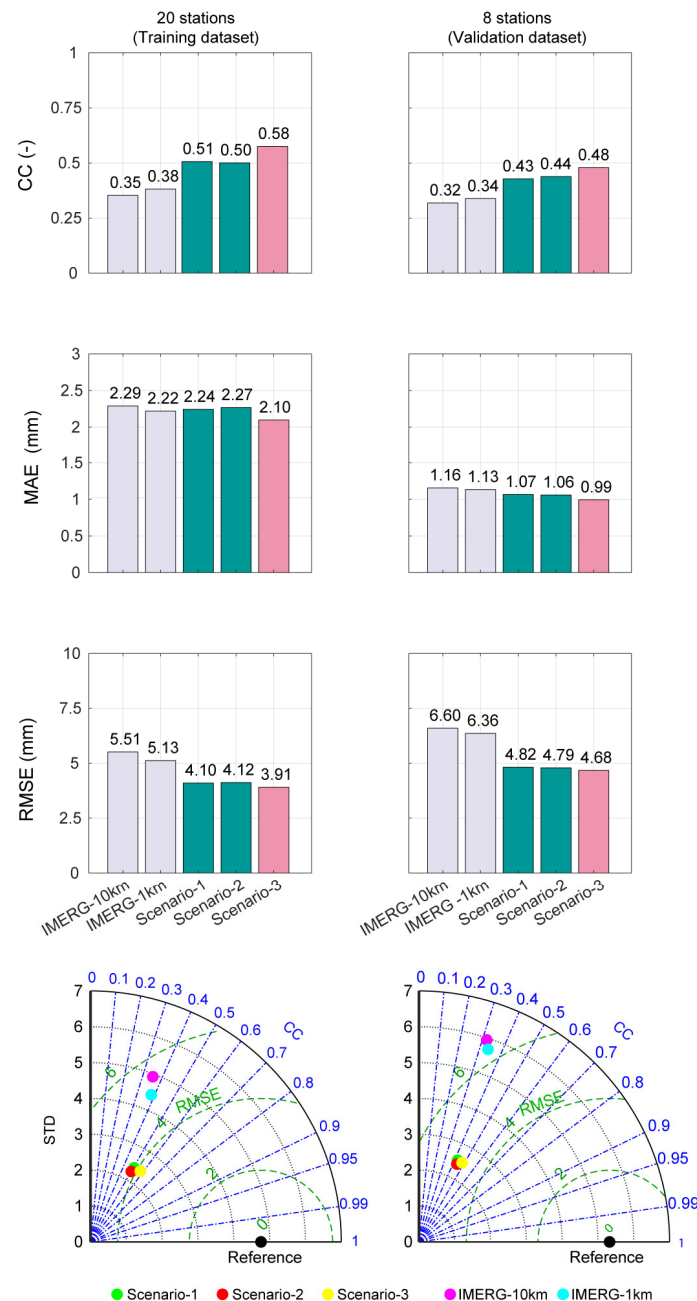


Figure 3. Statistical metrics (CC, MAE and RMSE) and Taylor diagrams constructed for the evaluation of the RF models' performance.

Applying bilinear interpolation to downscale IMERG to 1 km (i.e., IMERG-1 km) led to very slight improvements compared to the use of IMERG data at its native spatial resolution (i.e., IMERG-10 km) (Figure 3). IMERG-10 km and IMERG-1 km CC were higher for the training dataset (0.35, 0.38) than for the validation dataset (0.32, 0.34).

All of the scenarios studied had similar CC magnitudes in both analyses (i.e., considering the training and validation datasets), which indicates that there were no overfitting issues in the trained RF models. It is worth mentioning that the lower statistical scores

obtained by each scenario when considering the validation dataset are related to the lower scores overall that were also observed for the benchmarks.

The Taylor diagrams show that the scenarios reached more accurate results than the benchmarks according to the statistical indices (i.e., CC, RMSE and Standard Deviation), as all scenarios are closer to the reference dot. The results of the scenarios are ranked from scenario-3 to scenario-1 (in reverse order). Indeed, scenario-3 achieved the best results for both the training and validation datasets, with the highest CC scores (0.58 and 0.48) and the lowest MAE values (2.1 and 2.48 mm) and RMSE values (3.91 and 4.68 mm). Considering the validation dataset, these correspond to improvements of 50%, 15% and 29% when compared to IMERG-10 km and to improvements of 41%, 12% and 26% when compared to IMERG-1 km.

Scenario-1, which only considers the MODIS and IMERG variables as regressors (COT, CER, CWP, IMERG), also considerably outperformed both benchmarks. For example, when considering the validation dataset, Scenario-1 reached CC and RMSE values of 0.43 and 4.82, whereas IMERG-10 km reached values of 0.32 and 6.60, respectively. This corresponds to improvements of 34% and 27%, respectively. Therefore, even if the temporal mismatch between IMERG and MODIS observations (IMERG relies on 48 half-hour observations, whereas MODIS relies on two five-min observations) can result in a “non-rainy” cloud cover condition, despite detection of a precipitation event by IMERG (before or after the MODIS Terra and/or Aqua observations), the MODIS cloud optical and microphysical properties still constitute very valuable information for downscaling SPEs at the daily time step. All scenarios present higher CC and lower RMSE and MAE than the benchmarks, highlighting the benefits of the proposed downscaling method.

Figure 4 shows the statistical metrics (i.e., CC, MAE and RMSE) obtained after RF model application (Stage 2) at the 28 pixels that contain rainfall stations.

When considering the training dataset, the max and min CC for IMERG-10 km were 0.52 and 0.22, and increased to 0.71 and 0.46 in scenario-3 (i.e., the scenario with the best results). This corresponds to improvements of 36% and 109%, respectively. Similarly, the scenarios' RMSE and MAE values were lower than those of the benchmarks (Figure 4). For example, the max and min RMSE (MAE) values decreased from 7.64 mm and 4.09 mm (4.39 mm and 1.71 mm) for IMERG-10 km to 6.65 mm and 2.78 mm (3.86 mm and 1.47 mm) for scenario-3. This corresponds to improvements of 13% and 32% for max and min RMSE, respectively, and to improvements of 12% and 14% for max and min MAE, respectively.

To assess the overall performance of the scenarios, it is convenient to focus on the results obtained with the eight stations not considered in the training procedure (i.e., the validation dataset). The scenario-3 outperforms IMERG-10 km for all the statistical indices. The improvement rates in each maximum/minimum statistical metric were 67%/52% for the CC metric, 24%/16% for the MAE metric, and 31%/33% for the RMSE metric.

It is worth mentioning that scenario-1, which considers MODIS and IMERG variables as regressors (COT, CER, CWP, IMERG) also considerably outperformed the benchmarks. For example, when considering the validation dataset, scenario-1 presented max and min CC values of 0.60 and 0.33, corresponding to improvements of 50% and 32% when compared to IMERG-10 km and to improvements of 39% and 27% when compared to IMERG-1 km. The inclusion of topography features (slope and elevation) in scenario-2 did not improve the precipitation estimates compared to scenario-1. Indeed, the max and min values of the indices for both scenarios were very close. As topography might control the cloud dynamics and accumulation [61–63], topography and cloud properties may be redundant, which could explain why scenario-1 and scenario-2 performed similarly.

On the other hand, the consideration of air temperature in scenario-3 led to a significant improvement in precipitation estimates. In accordance with the Clausius–Clapeyron relation, air temperature changed over time according to the climate conditions. Those variations influence the atmospheric humidity and help in the formation of clouds, and thus precipitation events. Therefore, air temperature may be complementary to MODIS

cloud optical and microphysical properties [52,53], as it supports the discretization of rainy from non-rainy clouds.

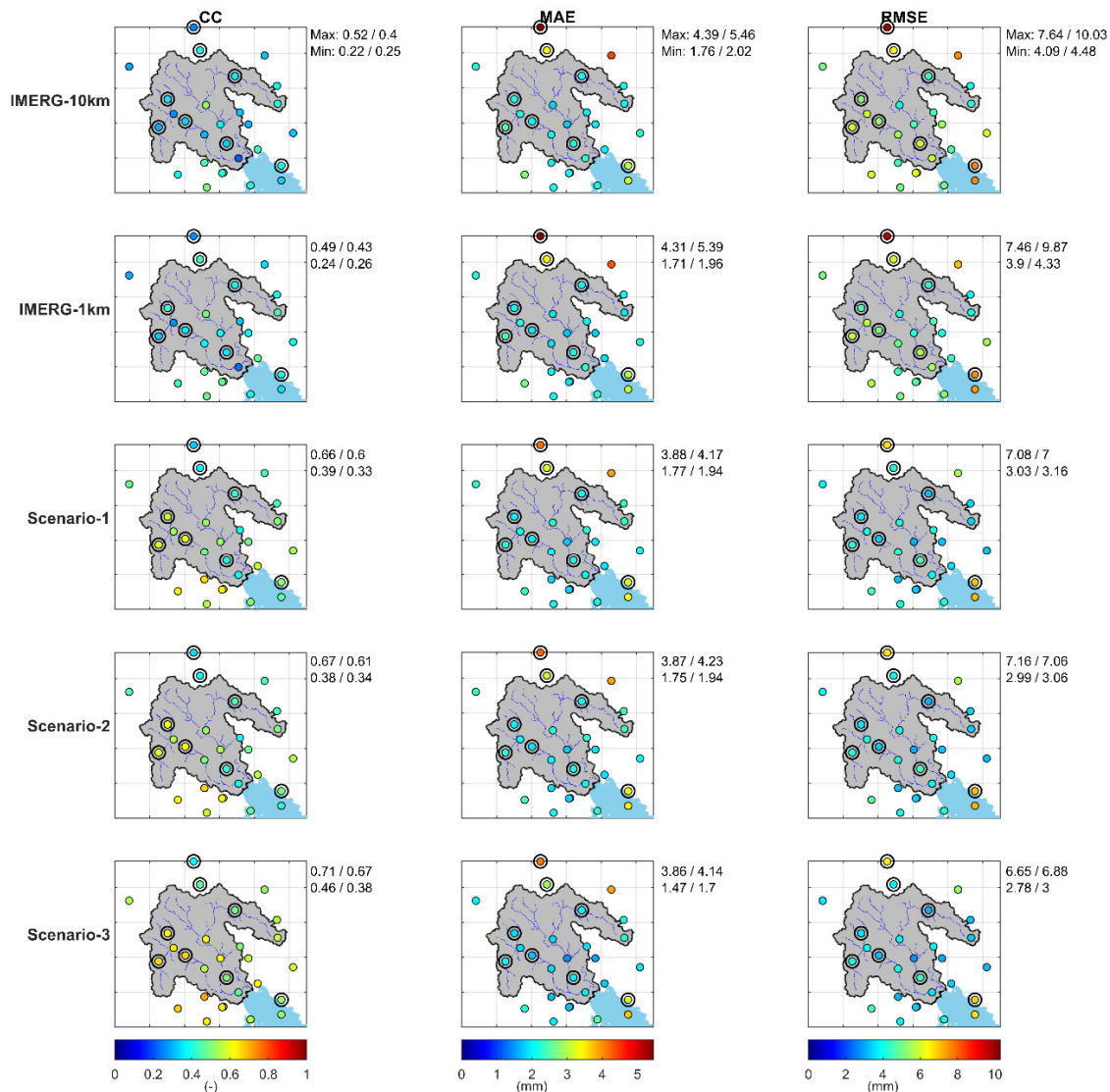


Figure 4. CC, MAE and RMSE calculated by comparing all scenarios and benchmarks against the rain gauge observations. The rain gauges with rings correspond to the 8 stations not used for the model training (i.e., the validation dataset). For max and min values, the first and the second value are those obtained with the training and the validation datasets, respectively.

4.2. Evaluation from Qualitative Indices

Figure 5 illustrates the precipitation event forecasting capability exhibited by the scenarios and the benchmarks. The POD was higher for the scenarios than for the benchmarks, demonstrating that the scenarios have better forecasting performance. Nevertheless, the FAR values indicate that the scenarios detect some precipitation events when they should not. Despite these shortcomings, the overall performance of the scenarios is comparable to that of the benchmarks, since the critical success indexes (CSI) computed for both groups (i.e., the benchmarks and the scenarios) have similar values (around 0.5).

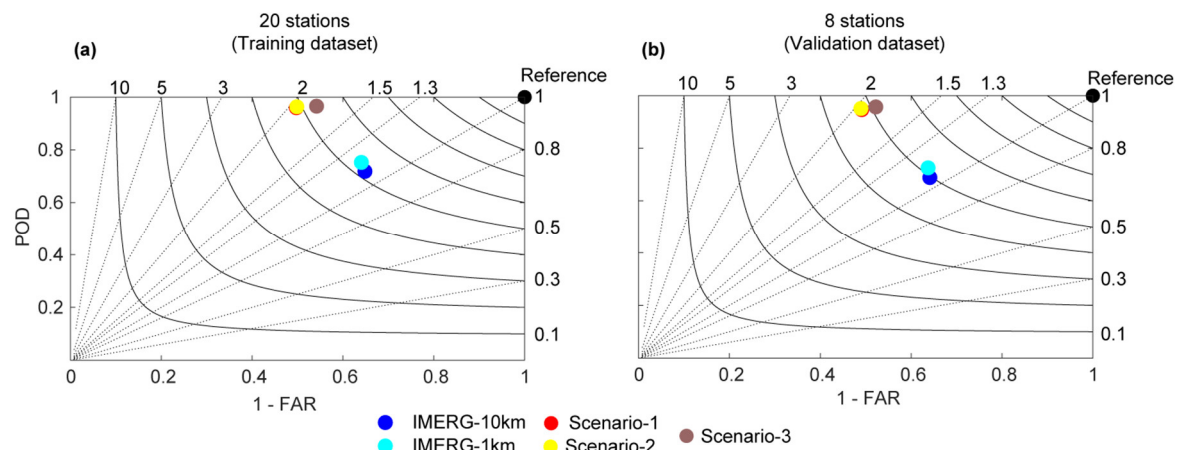


Figure 5. Performance diagram used to assess the SPEs' daily prediction performance using the training dataset (a) and the validation dataset (b). Curved lines and straight lines represent the CSI and BIAS values, respectively. The reference black dot represents a perfect statistical score.

A recent study in China proposed two-step machine-learning modeling to merge six SPEs with rain gauge information [9]. First, a classification model was trained to identify rainy and non-rainy days, and secondly, a regression model was used to estimate the rainfall amounts of the classified rainy days. The resulting merge datasets outperformed the six studied SPEs in terms of categorical metrics (including POD, FAR and CSI) and substantially alleviated the temporal and spatial biases in the rainfall estimates. This two-step modeling approach could be used as a guideline for further studies aiming to improve the categorical (and quantitative) metrics of the proposed scenarios.

4.3. Assessment of the Precipitation Spatial Pattern

The climate variability over the study area depends on the easterly flow that transports humid air from the Amazon region (summer) and the westerly flow that comes from the Pacific region (winter). The westerly flow inhibits moisture transport toward Lake Titicaca, and instead provides the study region with dry air [64].

The model presented in the current study does not consider the previously mentioned atmospheric circulation, instead focusing on the analysis of precipitation patterns contemplating only the amount of rainfall that fell in 2012. In that sense, the scenarios and the benchmarks allow us to quantify the annual precipitation over the study area (Figure 6). The spatial pattern of the annual precipitation is more conspicuous in the benchmarks than in the scenarios. Indeed, all scenarios underestimate the amounts of precipitation observed by the benchmarks across the northern Amazonian region, whereas they overestimate the precipitation amounts across the rest of the study area (Figure 6).

According to the isohyet maps of the Titicaca–Desaguadero–Poopó–Salar de Coipasa system [65], a positive precipitation anomaly can be observed over Lake Titicaca (i.e., there is more precipitation over the lake surface than in the surrounding land). Indeed, Lake Titicaca allows more solar radiation absorption, resulting in higher temperatures on the lake surface than for the surrounding land. The air masses pick up lake moisture while their temperature rises, increasing rainfall above the lake [66]. This strong land–water contrast (in terms of radiation absorption, temperature, and humidity) is a source of error for SPEs, which underestimate precipitation over the Lake Titicaca [6]. This climatic behavior is correctly captured by scenario-3 (Figure 6e), which estimates an annual precipitation of approximately 1000 mm across the lake. The consideration of T_x and T_n in scenario-3 allows a better representation of that behavior and improves the precipitation estimates across Lake Titicaca (i.e., positive rainfall anomaly).

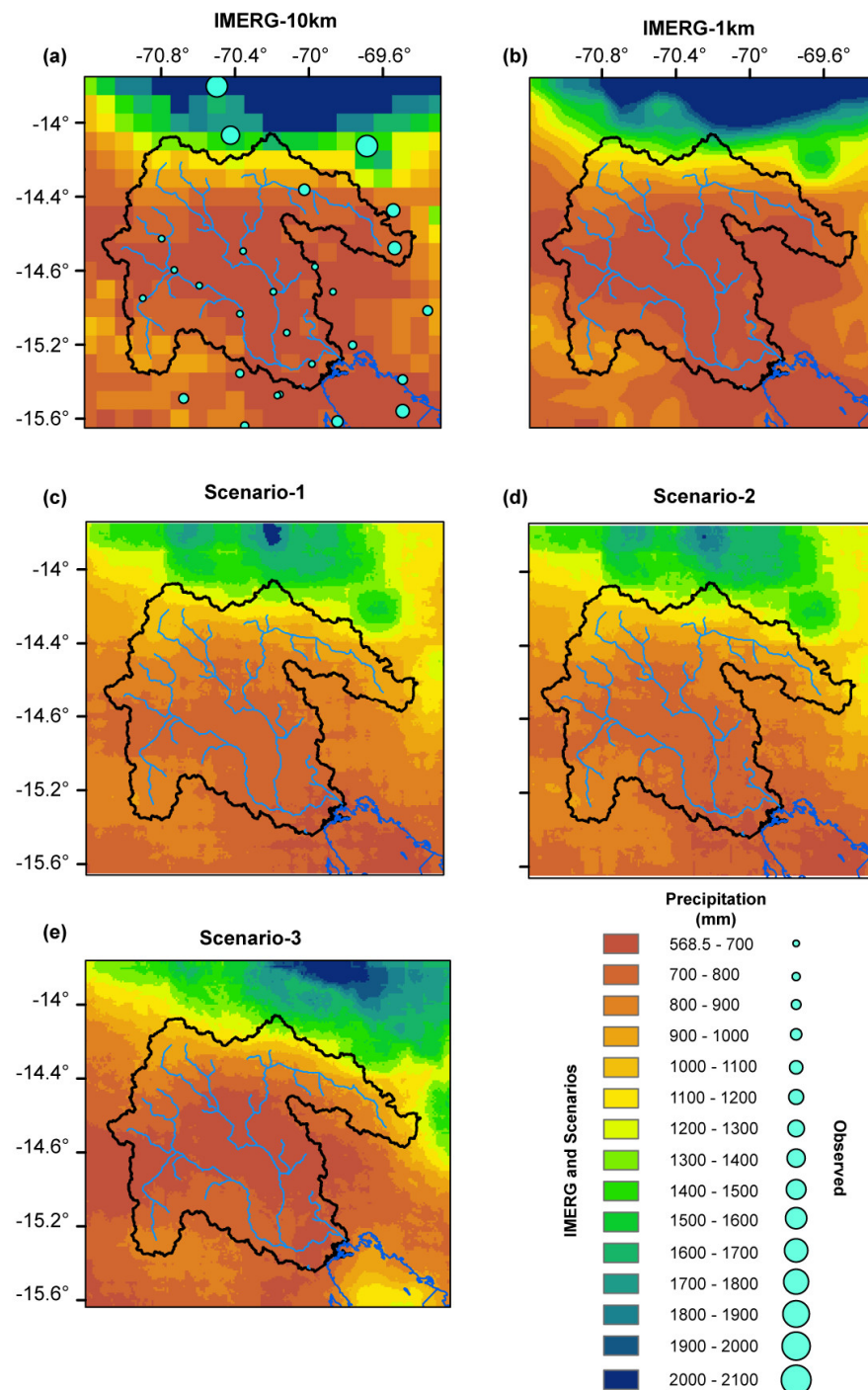


Figure 6. Annual rainfall (2012) pattern retrieved from each of the scenarios and benchmarks: (a) IMERG-10 km and observed data; (b) IMERG-1 km; (c) scenario-1; (d) scenario-2; (e) scenario-3.

5. Conclusions

This study assessed the benefits of using cloud optical and microphysical properties derived from MODIS along with temperature estimates derived from CHIRTS as input variables in a random forest (RF) model to downscale daily satellite-based precipitation estimates (SPEs) (i.e., IMERG) from 10 km to 1 km. Unlike others regressors generally used in statistical downscaling approaches (e.g., NDVI datasets), cloud optical and microphysical properties and air temperature have the substantial advantages of providing variables that are closely related with precipitation events at finer temporal scales (e.g., daily).

The results showed that the consideration of MODIS cloud optical thickness (COT), cloud effective radius (CER) and cloud water path (CWP) as auxiliary data provides a suitable alternative for daily SPE downscaling. Indeed, the downscaled precipitation estimates obtained with this dataset improved the reliability of IMERG by 34% and 27% for CC and RMSE, respectively. The additional use of CHIRTS maximum and minimum air temperature estimates (Tx and Tn, respectively) led to additional improvements. The combination of the MODIS and CHIRTS datasets improved the CC and RSME by 50% and 29%, respectively, compared to IMERG at its native spatial resolution (i.e., 10 km). This result suggests that the downscaled precipitation estimates are prone to improvement with the inclusion of a new set of regressors. Along these lines, the consideration of additional variables that are linked to precipitation at the daily time step (e.g., aerosol concentration, total water content, and wind flow patterns) should be considered as additional model inputs for improving the model outcomes. A future challenge will be the use of two RF models, one to detect rainfall events (i.e., a classification model) and another to estimate the amount of precipitation in the detected rainfall events (i.e., a regression model) in order to improve SPEs' daily precipitation forecasting.

Author Contributions: Conceptualization, S.C.M. and F.S.; methodology, S.C.M. and F.S.; formal analysis, S.C.M.; investigation, S.C.M.; data curation S.C.M.; Writing—original draft preparation, S.C.M. and F.S.; Writing—review and editing, S.C.M., F.S., J.M.-C., R.P.Z. and M.-P.B.; supervision F.S., J.M.-C. and R.P.Z.; project administration, F.S.; funding acquisition, F.S. and M.-P.B. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the Centre National d'Etudes Spatiales (CNES) in the framework of the IBYS project (Irrigation by satellite). The first author is grateful to the IRD (Institut de Recherche pour le Développement) for its financial support.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Fallah, A.; Sungmin, O.; Orth, R. Climate-dependent propagation of precipitation uncertainty into the water cycle. *Hydrol. Earth Syst. Sci.* **2020**, *24*, 3725–3735. [\[CrossRef\]](#)
2. Wild, M. The global energy balance as represented in CMIP6 climate models. *Clim. Dyn.* **2020**, *55*, 553–577. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Yang, D.; Yang, Y.; Xia, J. Hydrological cycle and water resources in a changing world: A review. *Geogr. Sustain.* **2021**, *2*, 115–122. [\[CrossRef\]](#)
4. Allan, R.P.; Barlow, M.; Byrne, M.P.; Cherchi, A.; Douville, H.; Fowler, H.J.; Gan, T.Y.; Pendergrass, A.G.; Rosenfeld, D.; Swann, A.L.S.; et al. Advances in understanding large-scale responses of the water cycle to climate change. *Ann. N. Y. Acad. Sci.* **2020**, *1472*, 49–75. [\[CrossRef\]](#)
5. Chen, F.; Gao, Y.; Wang, Y.; Li, X. A downscaling-merging method for high-resolution daily precipitation estimation. *J. Hydrol.* **2020**, *581*, 124414. [\[CrossRef\]](#)
6. Satgé, F.; Bonnet, M.-P.; Gosset, M.; Molina, J.; Lima, W.H.Y.; Pillco Zolá, R.; Timouk, F.; Garnier, J. Assessment of satellite rainfall products over the Andean plateau. *Atmos. Res.* **2016**, *167*, 1–14. [\[CrossRef\]](#)
7. Baez-Villanueva, O.M.; Zambrano-Bigiarini, M.; Ribbe, L.; Nauditt, A.; Giraldo-Osorio, J.D.; Thinh, N.X. Temporal and spatial evaluation of satellite rainfall estimates over different regions in Latin-America. *Atmos. Res.* **2018**, *213*, 34–50. [\[CrossRef\]](#)
8. Sharifi, E.; Saghafian, B.; Steinacker, R. Downscaling Satellite Precipitation Estimates with Multiple Linear Regression, Artificial Neural Networks, and Spline Interpolation Techniques. *J. Geophys. Res. Atmos.* **2019**, *124*, 789–805. [\[CrossRef\]](#)
9. Lei, H.; Zhao, H.; Ao, T. A two-step merging strategy for incorporating multi-source precipitation products and gauge observations using machine learning classification and regression over China. *Hydrol. Earth Syst. Sci.* **2022**, *26*, 2969–2995. [\[CrossRef\]](#)
10. Liang, S.; Wang, J. (Eds.) *Advanced Remote Sensing: Terrestrial Information Extraction and Applications*, 2nd ed.; Academic Press: Amsterdam, The Netherlands, 2020; ISBN 978-0-12-815826-5.
11. Goshime, D.W.; Absi, R.; Ledésert, B. Evaluation and Bias Correction of CHIRP Rainfall Estimate for Rainfall-Runoff Simulation over Lake Ziway Watershed, Ethiopia. *Hydrology* **2019**, *6*, 68. [\[CrossRef\]](#)
12. Salimi, A.H.; Masoompour Samakosh, J.; Sharifi, E.; Hassanvand, M.R.; Noori, A.; von Rautenkranz, H. Optimized Artificial Neural Networks-Based Methods for Statistical Downscaling of Gridded Precipitation Data. *Water* **2019**, *11*, 1653. [\[CrossRef\]](#)

13. Satgé, F.; Hussain, Y.; Molina-Carpio, J.; Pillco, R.; Laugner, C.; Akhter, G.; Bonnet, M. Reliability of SM2RAIN precipitation datasets in comparison to gauge observations and hydrological modelling over arid regions. *Int. J. Climatol.* **2021**, *41*, E517–E536. [\[CrossRef\]](#)
14. Chiang, Y.-M.; Hsu, K.-L.; Chang, F.-J.; Hong, Y.; Sorooshian, S. Merging multiple precipitation sources for flash flood forecasting. *J. Hydrol.* **2007**, *340*, 183–196. [\[CrossRef\]](#)
15. de Coning, E. Optimizing Satellite-Based Precipitation Estimation for Nowcasting of Rainfall and Flash Flood Events over the South African Domain. *Remote Sens.* **2013**, *5*, 5702–5724. [\[CrossRef\]](#)
16. Saber, M.; Yilmaz, K. Evaluation and Bias Correction of Satellite-Based Rainfall Estimates for Modelling Flash Floods over the Mediterranean region: Application to Karpuz River Basin, Turkey. *Water* **2018**, *10*, 657. [\[CrossRef\]](#)
17. Jiang, D.; Wang, K. The Role of Satellite-Based Remote Sensing in Improving Simulated Streamflow: A Review. *Water* **2019**, *11*, 1615. [\[CrossRef\]](#)
18. Wang, Q.; Xia, J.; She, D.; Zhang, X.; Liu, J.; Zhang, Y. Assessment of four latest long-term satellite-based precipitation products in capturing the extreme precipitation and streamflow across a humid region of southern China. *Atmos. Res.* **2021**, *257*, 105554. [\[CrossRef\]](#)
19. Satgé, F.; Xavier, A.; Pillco Zolá, R.; Hussain, Y.; Timouk, F.; Garnier, J.; Bonnet, M.-P. Comparative Assessments of the Latest GPM Mission's Spatially Enhanced Satellite Rainfall Products over the Main Bolivian Watersheds. *Remote Sens.* **2017**, *9*, 369. [\[CrossRef\]](#)
20. Zubieta, R.; Molina-Carpio, J.; Laqui, W.; Sulca, J.; Ilbay, M. Comparative Analysis of Climate Change Impacts on Meteorological, Hydrological, and Agricultural Droughts in the Lake Titicaca Basin. *Water* **2021**, *13*, 175. [\[CrossRef\]](#)
21. Behrangi, A.; Hsu, K.; Imam, B.; Sorooshian, S.; Huffman, G.J.; Kuligowski, R.J. PERSIANN-MSA: A Precipitation Estimation Method from Satellite-Based Multispectral Analysis. *J. Hydrometeorol.* **2009**, *10*, 1414–1429. [\[CrossRef\]](#)
22. Wang, J.; Petersen, W.A.; Wolff, D.B. Validation of Satellite-Based Precipitation Products from TRMM to GPM. *Remote Sens.* **2021**, *13*, 1745. [\[CrossRef\]](#)
23. Wang, Z.; Zhong, R.; Lai, C.; Chen, J. Evaluation of the GPM IMERG satellite-based precipitation products and the hydrological utility. *Atmos. Res.* **2017**, *196*, 151–163. [\[CrossRef\]](#)
24. Sun, L.; Lan, Y. Statistical downscaling of daily temperature and precipitation over China using deep learning neural models: Localization and comparison with other methods. *Int. J. Climatol.* **2021**, *41*, 1128–1147. [\[CrossRef\]](#)
25. Wang, F.; Tian, D.; Lowe, L.; Kalin, L.; Lehrter, J. Deep Learning for Daily Precipitation and Temperature Downscaling. *Water Resour. Res.* **2021**, *57*, e2020WR029308. [\[CrossRef\]](#)
26. Yan, X.; Chen, H.; Tian, B.; Sheng, S.; Wang, J.; Kim, J.-S. A Downscaling–Merging Scheme for Improving Daily Spatial Precipitation Estimates Based on Random Forest and Cokriging. *Remote Sens.* **2021**, *13*, 2040. [\[CrossRef\]](#)
27. Pham, H.X.; Shamseldin, A.Y.; Melville, B.W. Projection of future extreme precipitation: A robust assessment of downscaled daily precipitation. *Nat. Hazards* **2021**, *107*, 311–329. [\[CrossRef\]](#)
28. Xiang, L.; Xiang, J.; Guan, J.; Zhang, F.; Zhao, Y.; Zhang, L. A Novel Reference-Based and Gradient-Guided Deep Learning Model for Daily Precipitation Downscaling. *Atmosphere* **2022**, *13*, 511. [\[CrossRef\]](#)
29. Ulloa, J.; Ballari, D.; Campoazano, L.; Samaniego, E. Two-Step Downscaling of Trmm 3b43 V7 Precipitation in Contrasting Climatic Regions with Sparse Monitoring: The Case of Ecuador in Tropical South America. *Remote Sens.* **2017**, *9*, 758. [\[CrossRef\]](#)
30. Xu, S.; Wu, C.; Wang, L.; Gonsamo, A.; Shen, Y.; Niu, Z. A new satellite-based monthly precipitation downscaling algorithm with non-stationary relationship between precipitation and land surface characteristics. *Remote Sens. Environ.* **2015**, *162*, 119–140. [\[CrossRef\]](#)
31. Chen, C.; Hu, B.; Li, Y. Easy-to-use spatial random-forest-based downscaling-calibration method for producing precipitation data with high resolution and high accuracy. *Hydrol. Earth Syst. Sci.* **2021**, *25*, 5667–5682. [\[CrossRef\]](#)
32. Karbalaye Ghorbanpour, A.; Hessels, T.; Moghim, S.; Afshar, A. Comparison and assessment of spatial downscaling methods for enhancing the accuracy of satellite-based precipitation over Lake Urmia Basin. *J. Hydrol.* **2021**, *596*, 126055. [\[CrossRef\]](#)
33. Kobayashi, T.; Masuda, K. Effects of precipitation on the relationships between cloud optical thickness and drop size derived from space-borne measurements. *Geophys. Res. Lett.* **2008**, *35*, L24809. [\[CrossRef\]](#)
34. Campos Braga, R.; Rosenfeld, D.; Krüger, O.O.; Ervens, B.; Holanda, B.A.; Wendisch, M.; Krisna, T.; Pöschl, U.; Andreae, M.O.; Voigt, C.; et al. Linear relationship between effective radius and precipitation water content near the top of convective clouds. *Atmos. Chem. Phys.* **2021**, *21*, 14079–14088. [\[CrossRef\]](#)
35. Kobayashi, T.; Masuda, K. Changes in Cloud Optical Thickness and Cloud Drop Size Associated with Precipitation Measured with TRMM Satellite. *J. Meteorol. Soc. Jpn. Ser. II* **2009**, *87*, 593–600. [\[CrossRef\]](#)
36. Gammons, C.H.; Slotton, D.G.; Gerbrandt, B.; Weight, W.; Young, C.A.; McNearny, R.L.; Cámac, E.; Calderón, R.; Tapia, H. Mercury concentrations of fish, river water, and sediment in the Río Ramis-Lake Titicaca watershed, Peru. *Sci. Total Environ.* **2006**, *368*, 637–648. [\[CrossRef\]](#)
37. Satgé, F.; Ruelland, D.; Bonnet, M.-P.; Molina, J.; Pillco, R. Consistency of satellite-based precipitation products in space and over time compared with gauge observations and snow- hydrological modelling in the Lake Titicaca region. *Hydrol. Earth Syst. Sci.* **2019**, *23*, 595–619. [\[CrossRef\]](#)
38. Lujano, E.; Sosa, J.D.; Lujano, R.; Lujano, A. Performance evaluation of hydrological models GR4J, HBV and SOCONT for the forecast of average daily flows in the Ramis river basin, Peru. *Rev. Ing. UC* **2020**, *27*, 189–199.

39. Pillco Zolá, R.; Bengtsson, L.; Berndtsson, R.; Martí-Cardona, B.; Satgé, F.; Timouk, F.; Bonnet, M.-P.; Mollericon, L.; Gamarra, C.; Pasapera, J. Modelling Lake Titicaca's daily and monthly evaporation. *Hydrol. Earth Syst. Sci.* **2019**, *23*, 657–668. [\[CrossRef\]](#)
40. Satgé, F.; Hussain, Y.; Xavier, A.; Pillco Zolá, R.; Salles, L.; Timouk, F.; Seyler, F.; Garnier, J.; Frappart, F.; Bonnet, M.-P. Unraveling the impacts of droughts and agricultural intensification on the Altiplano water resources. *Agric. For. Meteorol.* **2019**, *279*, 107710. [\[CrossRef\]](#)
41. Satgé, F.; Espinoza, R.; Zolá, R.; Roig, H.; Timouk, F.; Molina, J.; Garnier, J.; Calmant, S.; Seyler, F.; Bonnet, M.-P. Role of Climate Variability and Human Activity on Poopó Lake Droughts between 1990 and 2015 Assessed Using Remote Sensing Data. *Remote Sens.* **2017**, *9*, 218. [\[CrossRef\]](#)
42. Huffman, G.J.; Bolvin, D.T.; Braithwaite, D.; Hsu, K.; Joyce, R.J.; Kidd, C.; Nelkin, E.J.; Sorooshian, S.; Tan, J.; Xie, P. *IMERG_ATBD_VALgorithm Theoretical Basis Document (ATBD) Version 06—NASA Global Precipitation Measurement (GPM) Integrated Multi-Satellite Retrievals for GPM (IMERG)*; NASA: Washington, DC, USA, 2020.
43. Nwachukwu, P.N.; Satge, F.; Yacoubi, S.E.; Pinel, S.; Bonnet, M.-P. From TRMM to GPM: How Reliable Are Satellite-Based Precipitation Data across Nigeria? *Remote Sens.* **2020**, *12*, 3964. [\[CrossRef\]](#)
44. Crosson, W.L.; Al-Hamdan, M.Z.; Hemmings, S.N.J.; Wade, G.M. A daily merged MODIS Aqua–Terra land surface temperature data set for the conterminous United States. *Remote Sens. Environ.* **2012**, *119*, 315–324. [\[CrossRef\]](#)
45. Platnick, S.; King, M.D.; Meyer, K.G.; Wind, G.; Amarasinghe, N.; Marchant, B.; Arnold, G.T.; Zhang, Z.; Hubanks, P.A.; Ridgway, B.; et al. *MODIS Cloud Optical Properties: User Guide for the Collection 6 Level-2 MOD06/MYD06 Product and Associated Level-3 Datasets*; NASA: Washington, DC, USA, 2015.
46. HDF Group. *HDF-EOS to GeoTIFF Conversion Tool (HEG) Stand-Alone User's Guide*; Riverdale; Raytheon Company: Waltham, MA, USA, 2019; p. 118.
47. Garreaud, R. Multiscale Analysis of the Summertime Precipitation over the Central Andes. *Mon. Weather Rev.* **1999**, *127*, 901–921. [\[CrossRef\]](#)
48. Funk, C.; Peterson, P.; Landsfeld, M.; Pedreros, D.; Verdin, J.; Shukla, S.; Husak, G.; Rowland, J.; Harrison, L.; Hoell, A.; et al. The climate hazards infrared precipitation with stations—A new environmental record for monitoring extremes. *Sci. Data* **2015**, *2*, 150066. [\[CrossRef\]](#) [\[PubMed\]](#)
49. Funk, C.; Peterson, P.; Peterson, S.; Shukla, S.; Davenport, F.; Michaelsen, J.; Knapp, K.R.; Landsfeld, M.; Husak, G.; Harrison, L.; et al. A High-Resolution 1983–2016 Tmax Climate Data Record Based on Infrared Temperatures and Stations by the Climate Hazard Center. *J. Clim.* **2019**, *32*, 5639–5658. [\[CrossRef\]](#)
50. Satgé, F.; Pillco Zolá, R.; Molina-Carpio, J.; Pacheco Mollinedo, P.; Bonnet, M.-P. Reliability of gridded temperature datasets to monitor surface air temperature variability over Bolivia. *J. Climatol.* **2023**. [\[CrossRef\]](#)
51. Verdin, A.; Funk, C.; Peterson, P.; Landsfeld, M.; Tuholske, C.; Grace, K. Development and validation of the CHIRTS-daily quasi-global high-resolution daily temperature data set. *Sci. Data* **2020**, *7*, 303. [\[CrossRef\]](#)
52. Martinkova, M.; Kysely, J. Overview of Observed Clausius-Clapeyron Scaling of Extreme Precipitation in Midlatitudes. *Atmosphere* **2020**, *11*, 786. [\[CrossRef\]](#)
53. Blenkinsop, S.; Chan, S.C.; Kendon, E.J.; Roberts, N.M.; Fowler, H.J. Temperature influences on intense UK hourly precipitation and dependency on large-scale circulation. *Environ. Res. Lett.* **2015**, *10*, 054021. [\[CrossRef\]](#)
54. Arias, P.A.; Garreaud, R.; Poveda, G.; Espinoza, J.C.; Molina-Carpio, J.; Masiokas, M.; Viale, M.; Scaff, L.; Van Oevelen, P.J. Hydroclimate of the Andes Part II: Hydroclimate Variability and Sub-Continental Patterns. *Front. Earth Sci.* **2021**, *8*, 505467. [\[CrossRef\]](#)
55. Prudhomme, C.; Reed, D.W. Relationships between extreme daily precipitation and topography in a mountainous region: A case study in Scotland. *Int. J. Climatol.* **1998**, *18*, 1439–1453. [\[CrossRef\]](#)
56. Tadono, T.; Ishida, H.; Oda, F.; Naito, S.; Minakawa, K.; Iwamoto, H. Precise Global DEM Generation by ALOS PRISM. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2014**, *II-4*, 71–76. [\[CrossRef\]](#)
57. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [\[CrossRef\]](#)
58. Buitinck, L.; Louppe, G.; Blondel, M.; Pedregosa, F.; Mueller, A.; Grisel, O.; Niculae, V.; Prettenhofer, P.; Gramfort, A.; Grobler, J.; et al. API design for machine learning software: Experiences from the scikit-learn project. *arXiv* **2013**, arXiv:arXiv.1309.0238. [\[CrossRef\]](#)
59. Roebber, P.J. Visualizing Multiple Measures of Forecast Quality. *Weather Forecast.* **2009**, *24*, 601–608. [\[CrossRef\]](#)
60. Oliveira, R.; Maggioni, V.; Vila, D.; Morales, C. Characteristics and Diurnal Cycle of GPM Rainfall Estimates over the Central Amazon Region. *Remote Sens.* **2016**, *8*, 544. [\[CrossRef\]](#)
61. Bendix, J.; Rollenbeck, R.; Göttlicher, D.; Cermak, J. Cloud occurrence and cloud properties in Ecuador. *Clim. Res.* **2006**, *30*, 133–147. [\[CrossRef\]](#)
62. Kumar, S.; Vidal, Y.-S.; Moya-Álvarez, A.S.; Martínez-Castro, D. Effect of the surface wind flow and topography on precipitating cloud systems over the Andes and associated Amazon basin: GPM observations. *Atmos. Res.* **2019**, *225*, 193–208. [\[CrossRef\]](#)
63. Valjarević, A.; Morar, C.; Živković, J.; Niemets, L.; Kićović, D.; Golijanin, J.; Gocić, M.; Bursać, N.M.; Stričević, L.; Žiberna, I.; et al. Long Term Monitoring and Connection between Topography and Cloud Cover Distribution in Serbia. *Atmosphere* **2021**, *12*, 964. [\[CrossRef\]](#)
64. Canedo, C.; Pillco Zolá, R.; Berndtsson, R. Role of Hydrological Studies for the Development of the TDPS System. *Water* **2016**, *8*, 144. [\[CrossRef\]](#)

65. PNUMA. *Perspectivas del Medio Ambiente en el Sistema Hídrico—Titicaca-Desaguadero-Poopó-Salar de Coipasa (TDPS)—GEO Titicaca*; PNUMA: Nairobi, Kenya, 2011.
66. Roche, M.-A.; Fernández-Jáuregui, C.A.; Aliaga Rivera, A. (Eds.) *Balance Hídrico Superficial de Bolivia*; Ministerio de Medio Ambiente y Agua: Montevideo, Uruguay, 1992; ISBN 978-92-9089-029-4.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.