Article

Optimization of X-Band Radar Rainfall Retrieval in the Southern Andes of Ecuador Using a Random Forest Model

Johanna Orellana-Alvear [1,2,*], Rolando Célleri [2,3], Rütger Rollenbeck [1] and Jörg Bendix [1]

[1] Laboratory for Climatology and Remote Sensing (LCRS), Faculty of Geography, University of Marburg, D-35032 Marburg, Germany

[2] Departamento de Recursos Hídricos y Ciencias Ambientales, Universidad de Cuenca, Cuenca EC010207, Ecuador

[3] Facultad de Ingeniería, Universidad de Cuenca, Cuenca EC010207, Ecuador

[*] Correspondence: johanna.orellana@ucuenca.edu.ec; Tel.: +593-7-405-1000

check for updates

Abstract: Despite many efforts of the radar community, quantitative precipitation estimation (QPE) from weather radar data remains a challenging topic. The high resolution of X-band radar imagery in space and time comes with an intricate correction process of reflectivity. The steep and high mountain topography of the Andes enhances its complexity. This study aims to optimize the rainfall derivation of the highest X-band radar in the world (4450 m a.s.l.) by using a random forest (RF) model and single Plan Position Indicator (PPI) scans. The performance of the RF model was evaluated in comparison with the traditional step-wise approach by using both, the Marshall-Palmer and a site-specific Z–R relationship. Since rain gauge networks are frequently unevenly distributed and hardly available at real time in mountain regions, bias adjustment was neglected. Results showed an improvement in the step-wise approach by using the site-specific (instead of the Marshall-Palmer) Z–R relationship. However, both models highly underestimate the rainfall rate (correlation coefficient < 0.69; slope up to 12). Contrary, the RF model greatly outperformed the step-wise approach in all testing locations and on different rainfall events (correlation coefficient up to 0.83; slope = 1.04). The results are promising and unveil a different approach to overcome the high attenuation issues inherent to X-band radars.

Keywords: radar; rainfall retrieval; machine-learning; mountain region; Andes; X-band

1. Introduction

Accurate rainfall estimations are needed for meteorological and hydrological applications. Rain gauge and disdrometer observations are frequently considered as ground truth data. Nevertheless, their reliability and spatial representation are limited. In contrast, weather radar technology provides far better coverage in space and also in time. Even though radar provides spatially distributed reflectivity data, they still need to be converted to rainfall rates. Due to the spatio-temporal variability in precipitation (i.e., drop size distribution) as well as signal attenuation through rain cells, it is very difficult to find an adequate relationship to transform reflectivity measurements to rainfall rate [1]. Consequently, radar rainfall retrieval has become a field of major interest and development in the research community.

Most of the studies on rain radar systems generally refer nowadays to dual-polarization devices [2,3]. They are capable of deriving relevant features used for the classification of hydrometeors and thus the quality of final rainfall products is benefited. However, this radar technology is expensive and even unaffordable in some regions. Lately, straghforward X-band radar technology derived from ship radars has emerged as a cost-effective alternative. This technology is well-suited for monitoring rainfall close

to the ground in limited areas with high spatial and temporal resolution. Low costs and maintenance, compared to the most operational, more sophisticated C- and S-band radars, have increased the use of this equipment worldwide. According to a recent report of the World Meteorological Organisation (WMO), around 20% of the operational weather radars are now X-band [4]. Many of them are part of relevant radar networks used by the Collaborative Adaptive Sensing of the Atmosphere (CASA) Integrated Project 1 [5], or the hydrology-oriented Iowa Flood Studies (iFloodS) field campaign [6], just to name a few. The main limitation of X-band radars come from their attenuation issues. Nonetheless, several studies [2,7,8], given their financial capabilities, have used the key advantage of the derived features of dual-polarized radars to perform attenuation correction efficiently. On the other hand, the single polarized version of X-band, despite its restrictions and uncertainties [9] compared to dual-polarization, has also been successfully deployed (e.g., [10–12]), even in remote areas such as high mountains. Recently, Bendix et al. [13] reported the first X-band radar network at high altitudes in the southern Andes of Ecuador (Radarnet-Sur). Several studies [14–16] have shown encouraging results using the former LAWR (Local Area Weather Radar) in the network. However, due to differences in radar technology, and particularly in spatial exposure (i.e., higher altitudes up to 4450 m a.s.l. in mountain regions), quantitative precipitation estimation (QPE) of the other radar systems (Rainscanner, Selex GmbH) in the network remains a challenging task among the steep slopes of the Andes.

Prior studies have implemented mainly two different approaches to perform QPE using reflectivity (Z) from single polarized X-band radars: (a) a step-wise correction approach and (b) data-driven machine learning models. The former addresses one-by-one the well-known radar issues inherent to reflectivity observations such as clutter field, beam blockage [17], attenuation [18], radome attenuation [19] to name the most important ones, and then applies either a generalized or a locally derived Z–R relationship such as in van de Beek et al. [20,21]. However, this process is usually site specific and intricate, as corrections in reflectivity values also add noise and uncertainty to radar images [22]. Frequently, an additional step of bias adjustment using rain gauges is performed, as in Morin and Gabella [23]. Nonetheless, bias adjustment mostly depends on the rain gauge network distribution, operation and real-time availability, which are usually scarce and irregular in complex terrain. In contrast, the use of machine learning techniques for quantitative precipitation estimation has increased in recent years due to its simplicity and promising results [24–27]. This approach aims to reduce the complexity of the reflectivity correction process while properly simulate the relationship between reflectivity and rainfall. Several studies have used Artificial Neural Networks (ANN) as a radar rainfall retrieval method using information based on the vertical profile of rain [28,29] and at distances with reduced attenuation impact, which is still often a major issue when using X-band radar technology. Although ANNs are widely used, they have some intrinsic disadvantages such as slow convergence speed and less generalizing performance, which in turn makes ANN susceptible to overfitting. Moreover, ANN parameterization is hard to tune and thus expert knowledge is required for calibration. Alternatively, ensemble methods such as Random Forest have been proven to be robust and easy to implement. This decision-tree derived technique has been efficiently used in remote sensing applications [30] and radar QPE using reflectivity of several heights [31,32]. However, its application is still undocumented on radar rainfall retrieval studies using single Plan Position Indicator (PPI) scans.

Thus, this study aims to develop and evaluate a random forest model to be used as a rainfall retrieval method in a mountainous area with a scarce and unevenly distributed gauge network. The model will use reflectivity data from a single polarized X-band radar (CAXX radar) in the Southern Andes of Ecuador. This is the highest X-band radar in the world (4450 m a.s.l.) and thus, affected by the inherent attenuation issues mentioned above. To overcome these particularities, radar reflectivity will be used not only in its original form (i.e., raw reflectivity data) but also several derived features (e.g., cumulative reflectivity along the beam, average reflectivity along the beam, and so on) will be included as inputs in the model. Finally, the performance of the random forest model will be compared with the traditional step-wise approach through statistics of goodness of fit.

2. Materials and Methods

2.1. Study Site

The study site constitutes the south-east quadrant of the CAXX radar coverage to an extent of 60 km radius (see Figure 1 and Section 2.2 for further details). It comprises mainly the east flank of the western Andean Cordillera in southern Ecuador at high altitudes ranging from 2300 to 4450 m a.s.l. Despite the high elevations and due to its proximity to the Equator, this tropical mountain region lacks snow covered peaks [33]. The area extends over valley and mountain landscapes where highest elevations are covered by páramo vegetation [33] while urban regions are located in the outskirts of the cordillera. The area of interest covers around 75% of the Paute basin (area of 5066 km^2), which is located in the inter-Andean depression separating the western and Real (central) cordilleras [34]. This relevant basin provides several ecosystem services to the region and the country. The third main city in the country, Cuenca, is located within the basin at 30 km distance from the CAXX radar. In addition, the hydropower generation of the Paute Integral hydroelectric system accounts for approximately 45% of energy production of the country. Precipitation over the inter-Andean valleys is of a bimodal regime, with the rainy months in the periods March–April and October–November [35].
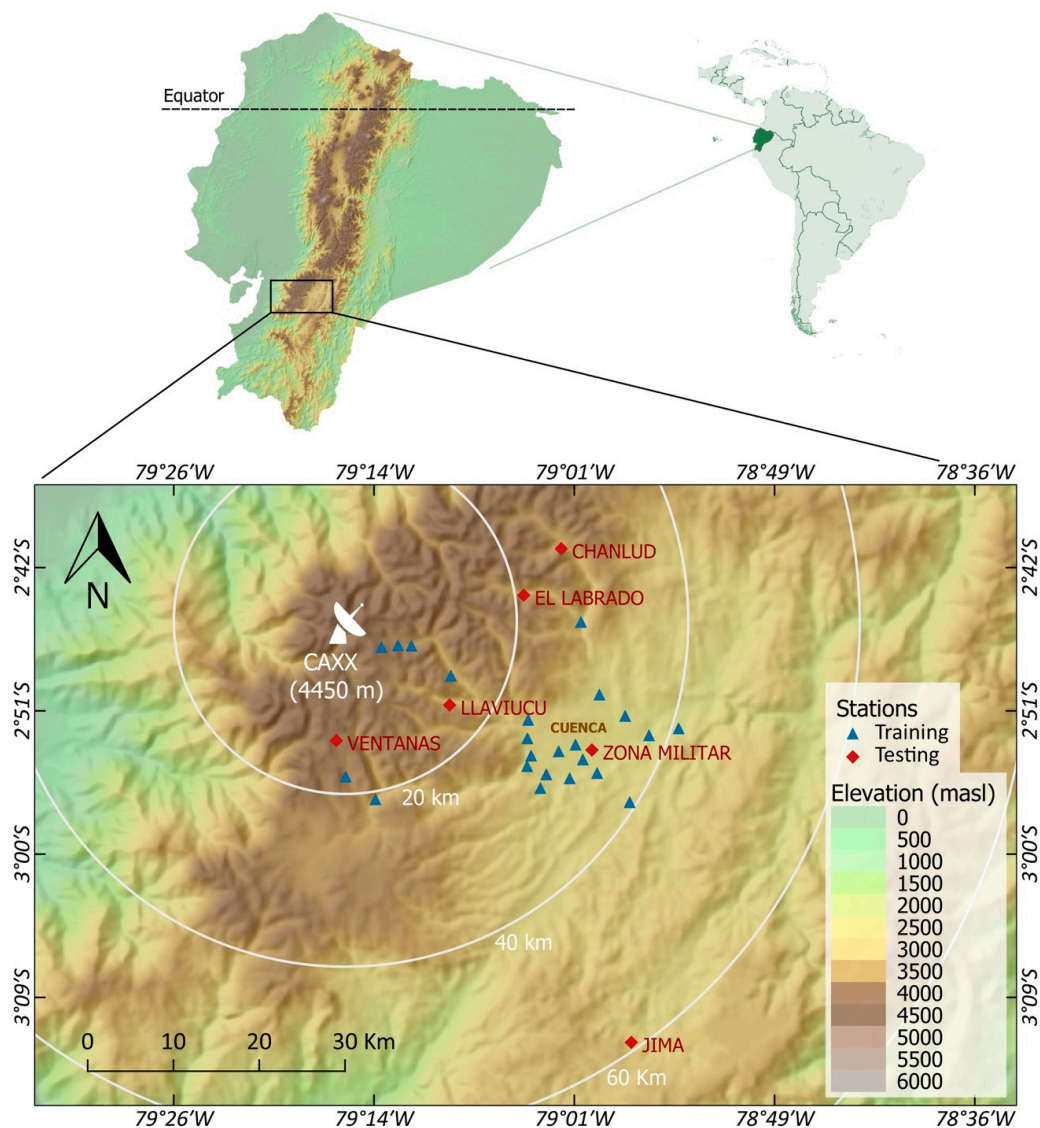
Figure 1. Map of the study site and the rain gauge network distribution.

2.2. Instruments and Data

2.2.1. Radar

A single polarized, non-Doppler, X-band radar was installed at 4450 m a.s.l. and made operational in April 2015. This is part of the Radarnet-Sur network in southern-Ecuador (CAXX; [13]). The radar was produced by SELEX GmbH and reaches a maximum range of 100 km. The SELEX RS 120 is a simplified cost-efficient instrument based on marine Radar technology and supplies 8-bit digitized samples of reflectivity along the scanlines. It is based on the traditional magnetron technology, which implies some physical limitations, of which variations of output power is the most influential factor. Magnetrons are magnetic resonators, whose efficiency depend on the strength of the internal magnet. As the magnetron needs to be heated to function properly, the magnetic field strength is slowly decaying during operation and the output power becomes lower over its lifetime. This is compensated by the internal electronics, but still affects the lower sensitivity threshold of the radar. Further technical details of the instrument are shown in Table 1. It is located at the north border of the Cajas National Park on the Paragüillas peak overlooking the city of Cuenca. The elevation angle of the radar antenna was redirected -2 degrees during its installation due to the altitude of the CAXX radar location at 4450 m a.s.l. at Paragüillas peak, one of the highest mountains in the area. It allowed the main beam to point zero degrees horizontally above the mountain range. This setting differs from other mountain radar locations where beam blockage is one of the most relevant issues. Exceptionally, at about 5 km north-west and south-west from the CAXX radar there were higher mountain peaks and therefore a strong (>80%) beam blockage was observed (see Figure 2). This X-band radar measures in two dimensions (2D). Polar images were obtained at 5-min frequency where each bin has a resolution of 2 degrees and 100 m. The 5-min mean reflectivity is internally calculated for each bin. Thus, a numerical matrix of 180 x 1000 reflectivity values was recorded at every time step. While built-in software from the SELEX company allows to apply clutter correction, a transformation to cartesian coordinates and a resampling to coarse resolutions are needed for further analyses. Thus we decided to keep up the raw polar data for this research. Reflectivity images in polar coordinates at 5-min frequency from 45 rainfall events that occurred within the period January 2016 to June 2017, were used in this study.

Table 1. Technical specifications of CAXX radar (Rainscanner 120, SelexGmbH). Peak Power is only available with a newly installed Magnetron. This peak value decreases slightly with time and hence is compensated by automatic internal recalibrations.

Technical Components and Its Parameters	Value
Antenna	
Diameter	1.2 m
Gain	38.5 dB
Azimuth Beam Width	2°
Elevation Beam Width	2°
Rotation Rate	12 rpm
Transmitter	
Peak Power	25 KW
Pulse Duration	500–1200 ns
Pulse Length	75–180 m
Receiver	
Bandwidth (1200 ns/500 ns)	3 MHz/7 MHz
Other technical specifications	
Maximum Range	100 km
Doppler	No
Polarization	Single
Outdoor Temperature	−20 °C to +50 °C
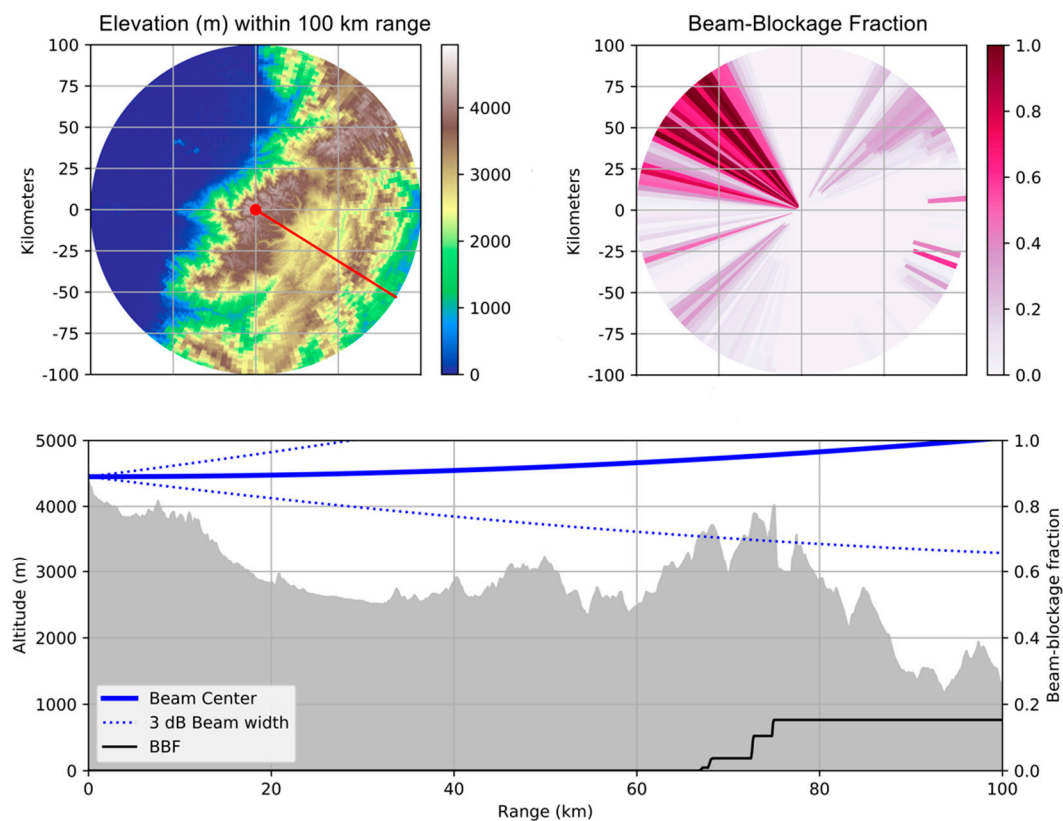
Figure 2. Beam-blockage fraction (BBF) calculated for the CAXX coverage. Transversal section for azimuth 122 (Balzay station location) is illustrated. The grey shaded area depicts the elevation of the terrain.
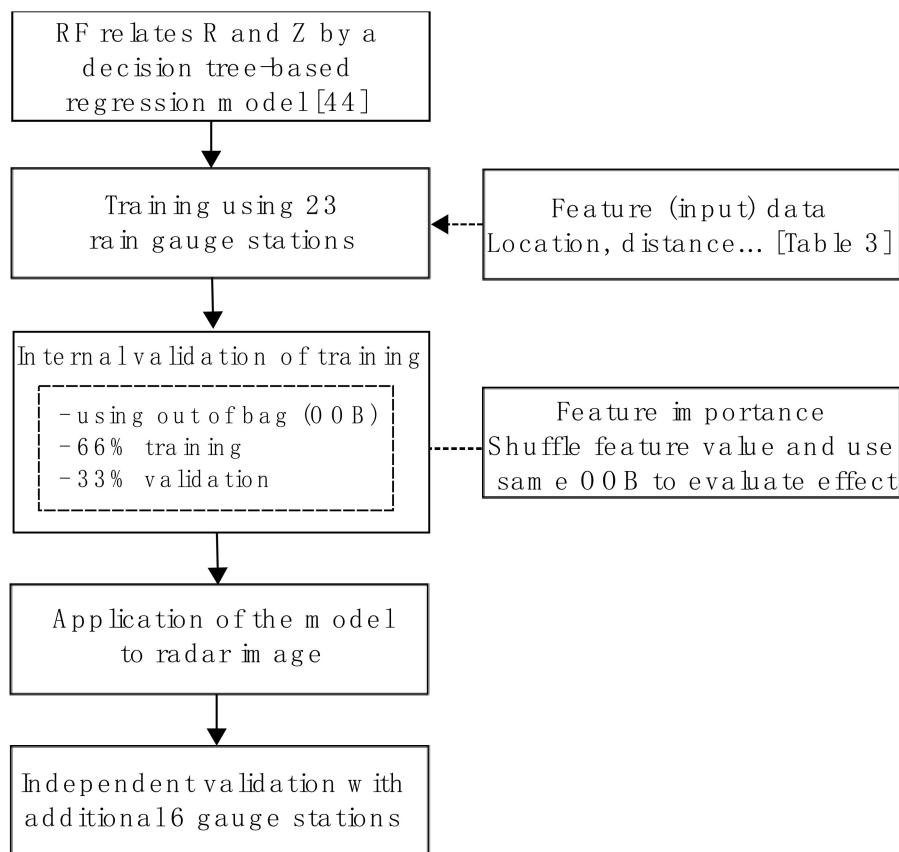
2.2.2. Disdrometers

Data from a Thies Laser Precipitation Monitor (LPM) located at 30 km south-east from the CAXX radar (azimuth = 122°) were used in this study. The disdrometer was located at Balzay station (2610 m a.s.l.). Recently, [1] validated and used its data to derive rain-type Z–R relationships in the radar coverage area. The same disdrometer dataset was used in the current study and facilitated the derivation of a k–Z attenuation relationship in the study area. A second LPM was located at 7 km from the radar (azimuth = 112°) at Virgen station (3636 m a.s.l.). For both disdrometers, reflectivity values were derived from the drop size distribution. A detailed explanation of this calculation can be found in [1]. Although these data were not used in the current study for modelling purposes, a comparison with the reflectivity records of the radar was performed. The latter ensures the proper calibration of the CAXX. Figure 3 depicts the correlation between the reflectivity measured by the radar and each LPM for two different rainfall events. Event from 02 April 2016 was registered one day after the magnetron was changed (i.e., when the radar had full output power and hence the least possible attenuation). It can be seen that at Virgen, the disdrometer and the radar reflectivity records show a good agreement. At Balzay, the decrease of the radar reflectivity is evident due to the attenuation caused by the storm nearby. Nonetheless, the radar is still able to measure up to 20 dBZ on these conditions at the Balzay station. On the other hand, the event from 10 March 2010 corresponds to a very strong event over the city. At the time, the magnetron was changed three months early. It can be observed that the radar, in general underestimates the storm at both stations when the disdrometer records reflectivity higher than 15 dBZ. Consequently, although attenuation and magnetron decay play a critical role in the radar measurements, we are certain the instrument is able to capture the rainfall variability. Further details on the disdrometers and the data can be found in [1].
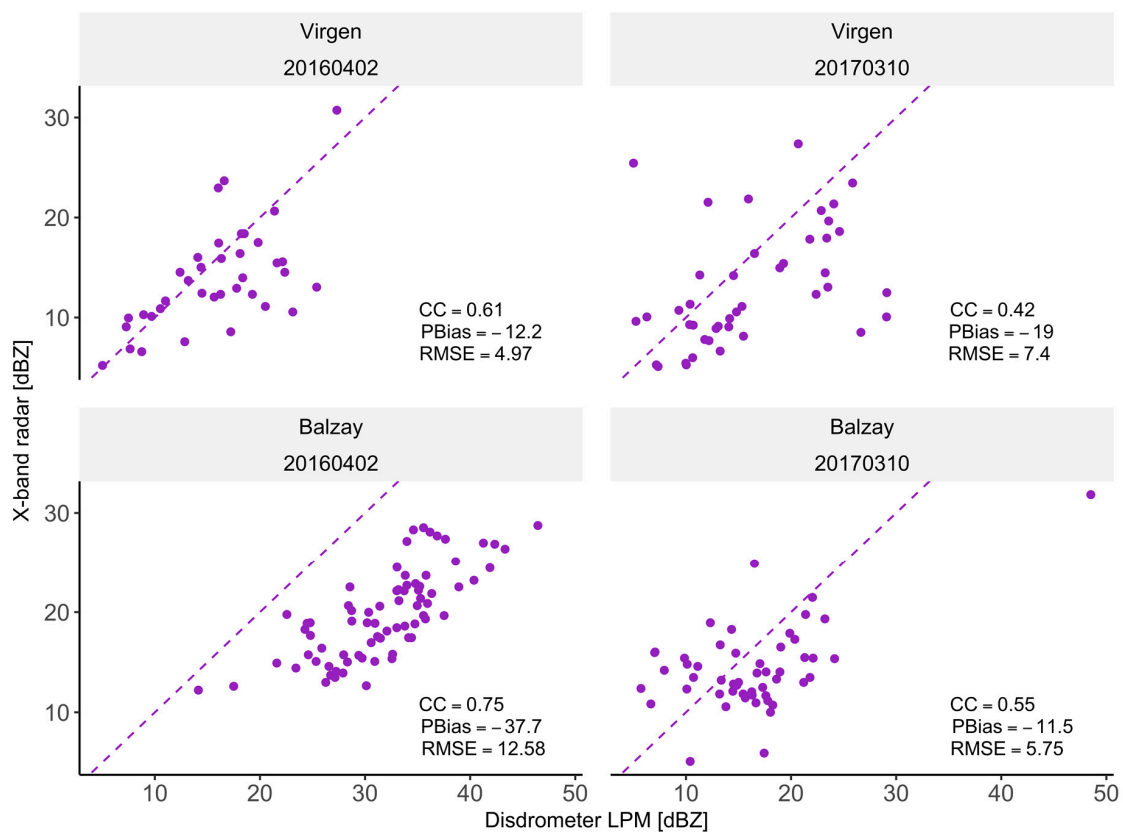
Figure 3. Reflectivity [dBZ] as recorded by both disdrometers at Virgen (7 km) and Balzay (30 km) from the radar location in comparison with the X-band CAXX measurements (>=5 dBZ) for two rainfall events. Pearson's correlation coefficient (CC), percentage bias (PBias) and root mean squared error (RMSE) are illustrated.

2.2.3. Rain Gauges

A rain gauge network comprising 29 stations was used in this study. The resolution of the tipping buckets was 0.1 mm and data were recorded at 5-min intervals. The rain gauges are unevenly distributed within the study site (see Figure 1). Regular calibration and maintenance were performed for all instruments in the network. For the case of the random forest model (for details on the RF model the reader may refer to Section 2.4. and Figure 4), rain gauge stations were split for the training-validation and test stages. The former used 23 rain gauges and the latter 6 rainfall stations which are representative for different locations and altitudes in the study area. Testing stations were selected in order to cover the entire study area and thus test the capability of generalization of the models. Statistics of altitudes and ranges from both groups of stations as well as details of the testing stations are given in Table 2. Rainfall intensity (R) data at 5-min frequency corresponding to 30 of the 45 selected rainfall events were used as target observations for the model in the training phase. This is equivalent to around 16,000 5-min intervals of positive (R > 0.1 mm h^{-1}) rainfall. The remaining 15 rainfall events were used in the test phase.

```
┌─────────────────────────┐
│   RF relates R and Z by a │
│    decision tree-based    │
│   regression model [44]   │
└─────────────────────────┘
           │
           ▼
┌─────────────────────────┐        ┌──────────────────────────┐
│    Training using 23      │ ◄----- │    Feature (input) data    │
│   rain gauge stations     │        │  Location, distance... [Table 3] │
└─────────────────────────┘        └──────────────────────────┘
           │
           ▼
┌─────────────────────────┐
│ Internal validation of training │
│  ┌────────────────────┐ │
│  ┆ - using out of bag (OOB) ┆ │        ┌──────────────────────────┐
│  ┆ - 66% training        ┆ │┄┄┄┄┄┄┄ │    Feature importance      │
│  ┆ - 33% validation      ┆ │        │ Shuffle feature value and use │
│  └────────────────────┘ │        │  same OOB to evaluate effect │
└─────────────────────────┘        └──────────────────────────┘
           │
           ▼
┌─────────────────────────┐
│   Application of the model │
│      to radar image       │
└─────────────────────────┘
           │
           ▼
┌─────────────────────────┐
│ Independent validation with │
│  additional 6 gauge stations │
└─────────────────────────┘
```

Figure 4. Simplified workflow of the development of the random forest (RF) model for the current study.

Table 2. (**a**) Statistics of the altitudes of the rain gauges and its distances from the CAXX radar used at each stage (23 for training-validation and 6 for testing) of the modeling process. (**b**) Further details of the independent rain gauges among the study site that were used for testing the models.

	Training-Validation		Test	
	Altitude [m]	Distance [km]	Altitude [m]	Distance [km]
Min.	2418	3.70	2545	13.60
Mean	2891	24.72	3219	26.83
Max.	3950	39.10	3907	57.80
Testing Station	X [m]	Y [m]	Altitude [km]	Distance [km]
CHANLUD	718604	9703574	3851	27.51
EL LABRADO	714224	9698186	3434	21.85
JIMA	726672	9646392	2898	59.35
LLAVIUCU	705563	9685489	3154	16.04
VENTANAS	692346	9681395	3592	13.24
ZONA MILITAR	722146	9680256	2568	33.14

2.3. QPE Models

Two different approaches are implemented and evaluated in the current study. The first approach consists of reflectivity upscaling and the application of the traditional Z–R relationship. It helps to build two different models: (i) using a locally derived Z–R equation and (ii) using the Marshall-Palmer equation. The second approach consists of the tree-based machine learning technique named random forest. In the following, details of each model are presented.

Step-Wise Correction Model

This method has been widely used in the literature and is well-known for applying corrections to radar imagery sequentially (i.e., correction on reflectivity values). Individual corrections related to physical and atmospheric conditions that may influence the reflectivity values recorded by the radar are applied. The complete correction chain differs depending on the radar location and the study goals (e.g., [11,21,36]). To facilitate the reproducibility of this study, we used the methods implemented in the *wradlib* library [37,38] which is developed in python programming language. Thus, we elaborate a step-wise approach, which compiles the different steps that are applicable to our study site.

First, ground echo removal is performed by considering both, static and dynamic clutter. The static clutter was calculated using the method described in Marra and Morin [36]. Since the magnetron had to be changed regularly (i.e, six months) three different static clutter signatures were observed from the images and thus static clutter was removed depending on the time period. The dynamic clutter was calculated using the Gabella filter [39], which accounts for texture mask and compactness. Both clutter signatures were removed and polar interpolation was performed afterwards to fill in the bins marked as clutter.

The CAXX radar antenna was vertically re-directed −2 degrees during its installation. Therefore, due to the original vertical inclination of the radar instrument (see Table 1), the main beam is currently zero degrees oriented. Figure 2 denotes the beam blockage map calculated based on Bech et al. (Equation (2) and Appendix; [40]). Since the area of interest remains fully visible from the main lobe, beam blockage correction is neglected.

Afterwards, a gate-by-gate attenuation correction was applied based on the iterative approach of Krämer and Verworn [41] and evaluated in Jacobi and Heisternmann [42] with a generalized and scalable number of constraints. The parameter of maximum reflectivity (maxdBZ) was set to 95.5 dBZ in accordance with the maximum possible value of reflectivity, as seen by the X-band radar in this study. Likewise, the value of maximum Path Integrate Attenuation (PIA) was fixed to 20 dB as suggested by Jacobi and Heistermann [42]. Parameters of the k–Z relationship needed for the Kraemer method were derived by using data from the Thies disdrometer located at Balzay station.

Furthermore, two Z–R relationships were evaluated: (a) the Marshall-Palmer equation $Z = 200R^{1.6}$ [43] and (b) a site-specific Z–R relationship $Z = 103R^{2.06}$ derived by Orellana-Alvear et al. [1] for the Balzay station. Usually, the last stage of the step-wise methodology for radar correction includes a bias adjustment using rain gauge data. However, this step will be excluded because the aim of this study is to compare the step-wise method and the random forest model without the use of rain gauges for later adjustment. This is because rain gauge data normally is only available with large delays in Ecuador and rain gauge networks are furthermore unevenly distributed in mountain regions as the Andes making processes as bias or kriging adjustment hardly applicable on hourly and sub-hourly time steps.

2.4. Random Forest Model

Figure 4 illustrates a simplified workflow of the random forest model implemented for quantitative precipitation estimation in the current study. Random Forest (RF) [44] is a decision-tree based machine learning technique that consists of random and equally distributed decision trees. This is one of the so-called ensemble methods. The final result is a combination; the mean in the simplest case, of all outputs of the trees in the forest. This technique has some advantages in contrast to other artificial intelligence methods. The RF algorithm for regression first obtains *n* datasets of random samples with replacement (bootstrapping), which are selected to construct *n* trees. Thus, a different subset is used for the construction of each decision tree of the model. A third part of data (out-of-bag OOB) is often used for the internal validation process with the purpose of obtaining unbiased estimations of the regression error, as well as feature importance estimations of the variables used in the tree construction. A regression tree is built for each set of samples. Then, for each node, a random subset of predictor variables is used to create the binary rule. The selection of the used variable to define the binary rule at

each step is based on the sum of square residuals. The tree is expanded without pruning. Finally, each observation of the OOB subset is evaluated in all regression trees and the average of the predictions is taken as the final result.

The use of this OOB data subset generates an internal validation of the algorithm based on the independent data from the training process and in consequence the generalization error is unbiased. On the other hand, the random structure of the subsets of the predictor variables used in each tree branch ensures a low correlation between trees, which increases the model robustness. Nonetheless, the drawback of RF lies in the limited interpretability since prediction comes from a forest of trees instead of a specific one. Finally, since the predicted value is the mean value of the whole all-tree predictions, it cannot exceed those from the training set. In consequence, RF tends to overestimate the low values and underestimate the high extremes.

For this study, an RF model was trained by using data of 30 selected rainfall events from 23 training rain gauge stations. From these, around ~16,000 rain observations were obtained at 5-min frequency and used as target variable. Input features described in Table 3 were extracted from the 5-min radar images corresponding to those rainfall events. Usually, to ensure spatial validation, leave-one-station-out cross-validation would be applied. However, due to the different number of (rain) observations at each location, optimized hyper-parameters (e.g., number of trees, number of features to construct each tree and depth of the tree) for the RF model were found through a grid search approach by using a 10-fold cross-validation. In addition, 15 different rainfall events at 6 testing stations were used for both, temporal and spatial independent testing.

It should be stressed that a test dataset is unnecessary when using an RF model due to its inner OOB error self-validation. However, the use of the independent rainfall events allows for a separate evaluation of the temporal and spatial performance of the model. This is particularly helpful to identify if the model is able to properly predict rainfall rates at longer distances from the radar and under different rainfall conditions. Hourly rainfall accumulation was performed for the evaluation from the original ~6000 rain observations at 5-min frequency corresponding to the 15 events. Testing rainfall events include three storms that are representative from different rainfall formation processes: (a) 09 March 2016 14:00 – 10 March 2016 07:30 (flash flood event in the city of Cuenca); (b) 19 March 2017 13:00 – 19 March 2017 17:00 (convective formation) and; (c) 10 June 2016 02:00–10 June 2016 09:00 (big cell event). All dates are stated at local time.

Input Features

The use of textural features as inputs of data-driven models has proven to be effective on non-coherent radar applications [45]. Thus, input features for the RF model were derived to include textural characteristics as well as features that stand for the rainfall evolution along the beam. Table 3 shows the interpretation of each feature used in the RF model. All of them are related to the location of the station. In addition, a feature importance analysis [44] was performed to better understand the features' influence on the model. It was accomplished by evaluating the increase of the error model on the OOB sample (i.e., 33% of the training dataset) after each feature of the model has been shuffled (i.e., breaking the relation of the feature and the output of the model). For this the scikit-learn library of python was used [46]. It should be pointed out that since the number of features is reduced and the purpose of this feature importance analysis was mainly to have some insights on the usefulness of derived features, we kept the trained RF model without further modifications (i.e., feature selection).

Table 3. Features used in the random forest model for radar rainfall retrieval.

Feature Name	Description
[Alt]	Altitude
[Dist]	Distance from radar
[AvgN]	Spatial average reflectivity (3x3 window) at the time step
[StdN]	Standard deviation of spatial reflectivity (3x3 window)
[Cum_dBZ]	Cumulative reflectivity along the beam until one bin before the station
[Cum_dBZ_npx]	Number of bins that exceeded a reflectivity threshold of 6 dBZ (rain signature according to Gabella and Notarpietro [39]).
[Cum_dBZ_lastbin]	Distance of the last rain bin (i.e., bin with rain signature) along the beam before reaching the station location.
[Std_temp]	Standard deviation of temporal evolution (lag-10 – lag-1 point image)
[Cum_dis]	Average reflectivity along the beam until the station [Cum_dBZ]/[Dist]
[AvgN_alt]	Quotient of average spatial reflectivity and altitude [AvgN]/[Alt]
[AvgN_cum]	Sum of cumulative dBZ along the beam [Cum_dBZ] and the spatial average reflectivity [AvgN] at the station. Strong differences with respect to the independent features [Cum_dBZ] or [AvgN] may help to increase the information gain regarding an attenuation situation.
[Cum_avg]	Quotient of cumulative reflectivity and number of bins where rainfall was occurring [Cum_dbZ]/[Cum_dBZ_npx]. Average of reflectivity by only considering rainy bins (i.e., average of "intensity" of rainfall).
[Cum_dis_lastbin]	Average of reflectivity until the last rain bin before reaching the station [Cum_dBZ]/[Cum_dBZ_lastbin].

2.5. Evaluation

Evaluation was performed by a comparison of the estimated (predicted) and observed rainfall rates for the test dataset. It was carried out considering both, temporal and spatial independence. Thus, the test dataset corresponded to 15 independent rainfall events which kept out of the training stage and the spatial locations of the stations were furthermore not used during training at all. For this, rain gauge observations and 5-min rainfall estimations were accumulated and compared to an hourly basis. The goodness of fit between observations and predictions was measured by means of the mean squared error (MSE), Pearson's coefficient correlation (CC) and the slope of the linear regression. Here, a linear regression model y = ax + b was derived, where y is the rain gauge rainfall and x is the estimated radar rainfall. In addition, similarly to Anagnostou et al. [47] the following statistical metrics were used: (a) relative mean error (rME); (b) relative root mean squared error (rRMSE) and; (c) efficiency score (Eff) described by Nash and Sutcliffe [48]. Both, ME and RMSE were normalized by the mean of the reference values. Finally, the measurement error was calculated through Equation (1). According to Thurai et al. [49], this metric is related to the radar parameters used in a specific algorithm.

$$\frac{var(R)}{\left(\overline{R}\right)^2} \tag{1}$$

Here, $var(.)$ denotes the variance of the quantity inside the brackets and the superscript bar denotes its mean.

3. Analysis and Discussion

3.1. Step-Wise Correction Model

The step-wise procedure was applied on the radar imagery to derive rainfall rates as explained in Section 2.3. In the following the intermediate results are discussed.

First of all, it was found that there is no influence of beam blockage in the study area as seen in Figure 2. Moreover, derivation of static clutter was performed for all different time periods where the magnetron was replaced in the radar. The very high reflectivity produced by mountain returns

in the near north-west and south-west region of the radar consistently appeared in all clutter maps. Contrary, minor differences in the static clutter signature (i.e., very low reflectivity values) of the near surroundings of the radar (~5 km) were found. Although physical calibration parameters remain unaltered in the SELEX ES Gematronik built-in software, such differences can be attributed to the power of the magnetron signal which may vary due to storage time and weather conditions.

Figure 5 illustrates the cumulative effect of the reflectivity correction related to the clutter removal and attenuation for the rainfall event from 10 February 2017 15:45 local time. It can be observed from the left hand side: (i) static clutter is concentrated in the close range of the radar (~5–7 km), (ii) static and dynamic clutter identification and removal showed that filtering static clutter still allowed some corrupted bins to remain in the image. This in turn causes the following polar interpolation to generate a weak reflectivity halo close to the radar. This is in agreement with [9] that highlighted that the effectiveness of any ground clutter algorithm is truncated due to the error source produced when ground clutter echoes are mistakenly identified/unidentified. Nonetheless, the high reflectivity in the near north-west and south-west regions of the radar were removed and thus the filtering properly worked under topographic enhanced conditions (i.e., very high reflectivity due to mountain returns). Regarding the dynamic clutter, as expected, small objects were dismissed (see Figure 5, region B). A similar effect was obtained in [39] as a result of the trade-off by taking out undesired reflectivity regions and the exclusion of very small rain cells and (c) attenuation correction by means of PIA. It can be observed that an enhancement of the reflectivity along the beam is accomplished in the radar scene. This is more notorious when the core intensity of the rain cells at the east is clearly augmented. A further detailed evaluation of PIA remains out the scope of this study.
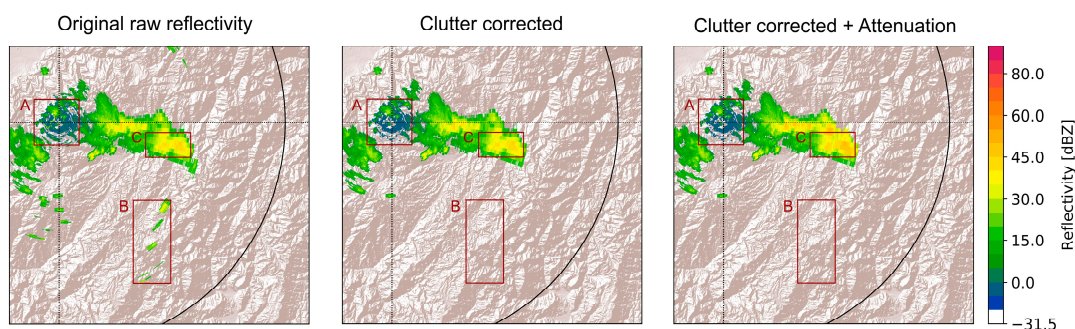

Figure 5. Static-Dynamic clutter and attenuation correction applied to the radar scene from 2017.02.10 15:45 local time. From left to right: (i) original uncorrected reflectivity, (ii) upscaled reflectivity after static (region A) and dynamic clutter (region B) removal, and (iii) upscaled reflectivity after clutter removal plus attenuation correction (region C).

3.2. Random Forest Model

An optimized random forest model was developed and used to estimate rain rate at 5-min interval. The best model configuration was found with 400 trees, 4 random features and 30 levels depth to build each tree. In the following, the results related to the (i) feature importance analysis and (ii) evaluation of the performance of the model are described. For the latter, rainfall estimations at 5-min steps were accumulated to an hourly basis and compared with rain gauge observations of the test dataset.

Figure 6 shows that there are significant differences in the feature importance for the RF model. For instance, the quotient between the spatial average of reflectivity *AvgN* (3 x 3 window) and the *altitude* of the station is the most important feature, while in contrast, the absolute altitude value has a lower influence on the model. A similar relation was found for *cum_dis* and *distance* features. This confirms that the model performance is highly influenced by data representation in machine learning techniques and consequently support our decision of including derived features to properly train the model. Furthermore, similarly to another study in the region [50], we found out through this

analysis that the terrain altitude does not play a significant role (i.e., low feature importance) in QPE. Finally, as expected, the higher the importance of the feature the higher its standard deviation.
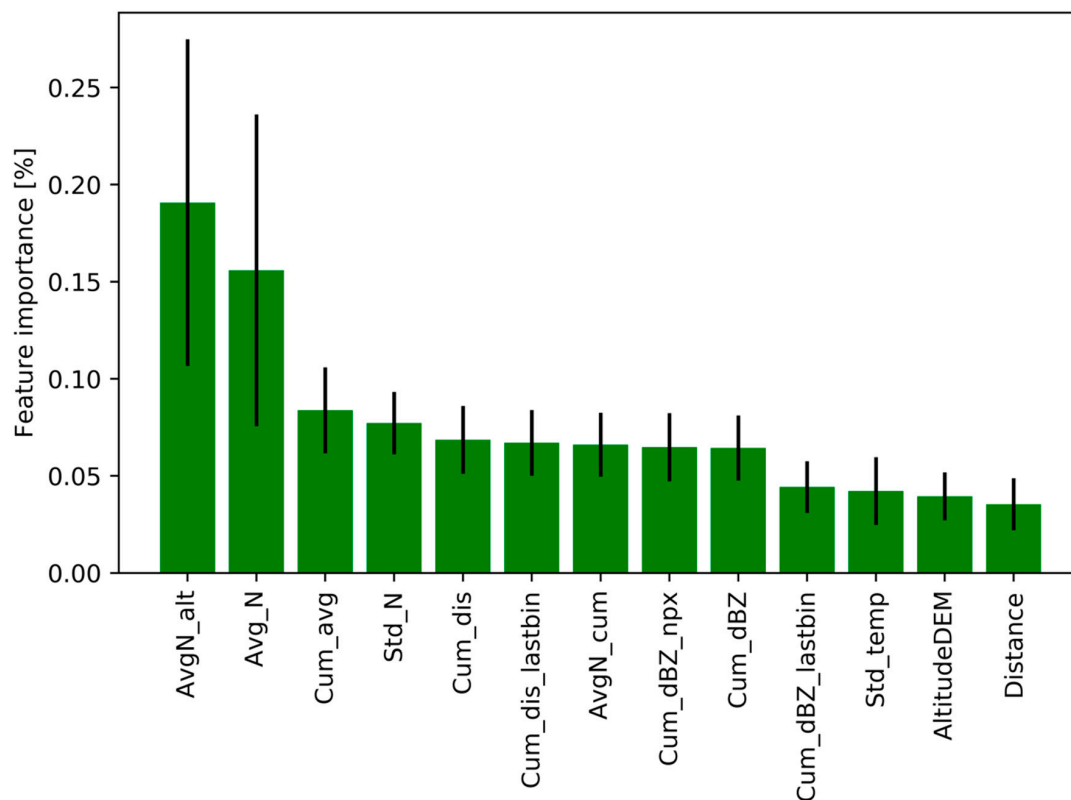
Figure 6. Feature importance and its standard deviation on the RF model. Out-of-bag (OOB) error was used to rank the features in the training stage.

The correlation of observed and estimated hourly rainfall applying the QPE random forest model to the entire test dataset is shown in Figure 7. It can be observed that lower rain rates are overestimated by the model while higher rain rates are underestimated. This is a well-known effect of random forest regression models due to the averaging of the resulting predictions of all trees in the forest. Nonetheless, an acceptable performance of the model (CC = 0.76; Eff = 0.55; rME = −0.1; MSE = 1.4) on the test dataset is reached. A similar performance of correlation coefficient of 0.79 was found in [20], but in contrast the authors implemented an exhaustive step-wise approach on the X-band radar imagery under about 15 km coverage. Furthermore, [20] found that rainfall event totals less than 3 mm h^{-1} were overestimated, while higher total events were mostly underestimated by using an X-band radar in the Netherlands. A similar situation holds for the performance of the RF model in the current study. It can be seen in Figure 7 that rain observations higher than 5 mm h^{-1} are underestimated in all cases. There are possible explanations for this situation: (i) lack of training samples under similar conditions. As reported by [1], contrary to what can be expected in tropical regions, the relative frequency of these rain rates is low in the study area. (ii) the accumulation and therefore the propagation of the error through the increasing of the temporal resolution (i.e., from 5 min to 60 min) can negatively affect the rainfall retrieval estimation. This is more notorious in higher rain rates. iii) when the core of a raincell occurs just above the rain gauge for almost the entire hour, features as *avgN* and *avgN_alt* (the highest ranked features) should be of utmost importance for a proper prediction. Probably all trees that did not pick these features to construct the tree produced very low predictions, which in turn generates a strong underestimation.
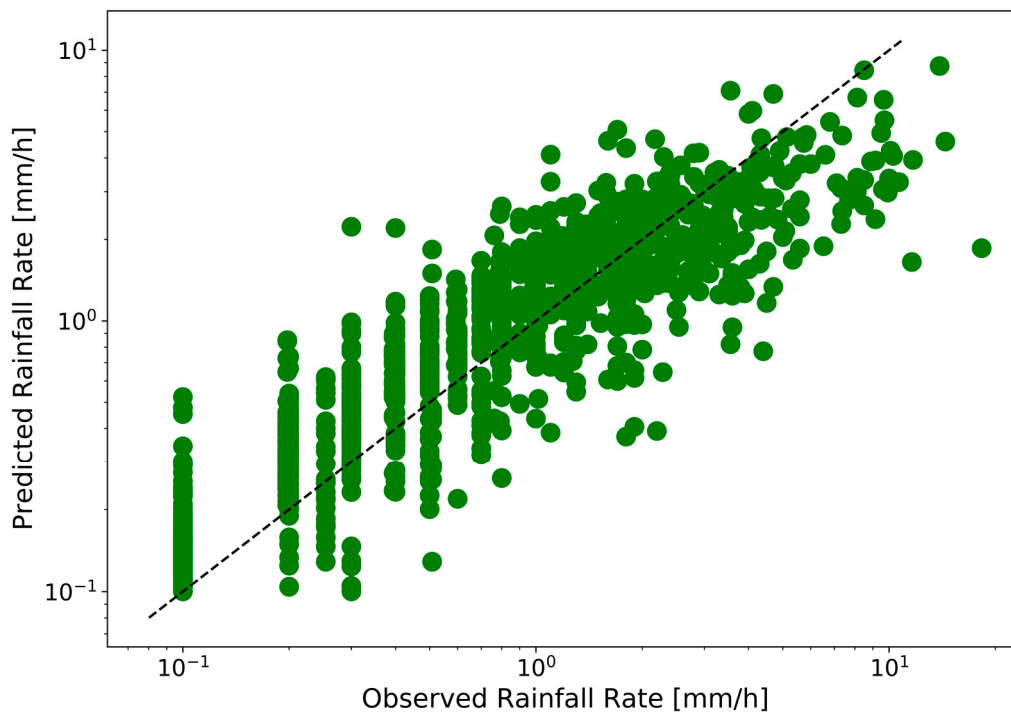
Figure 7. Correlation of observed and estimated hourly rainfall using QPE random forest model on the entire test dataset (15 independent events) is illustrated. The 1-1 line is show in black dotted.

3.3. Comparison of the Models - Temporal and Spatial Evaluation

The RF model consistently outperformed the other models at all spatial locations as seen in Figure 8 and Table 4. Despite of the distance from the radar, the estimations of the RF model remain mostly acceptable (CC > 0.65). This finding is of utmost importance since the range is well known to negatively impact on the rainfall estimation at X-band radars. Thus, contrary to other studies such as [21], one could overcome this limitation and use the relevant radar imagery at longer distances (e.g., 50 km). The lower statistical measures of RF occur at Jima station (CC = 0.66; Eff = 0.38; rME = −0.27, MSE = 4.30). It is located at about 60 km from the radar, which increases the possibility of uncertainties along the beam and thus the derived features that came from reflectivity data. Thus, all together, the RF model performed remarkably with a goodness of fit up to CC = 0.83 and Eff = 0.67 at different stations and what is more importantly without any further bias adjustment using rain gauge data. The latter is a critical difference to other studies that had accomplished similar goodness of fit 0.6 < CC < 0.85 only mostly using high density rain gauge networks for bias correction as in [23] and [36] or vertical reflectivity profiles in the case of machine learning approaches as in [26,27]. The performance of the implemented RF technique in [20] (correlation coefficient of 0.82) was likewise comparable to our RF model (CC = 0.76) using a short-term dataset. However, it should be noticed that a long-term evaluation (e.g., at least one year) was not possible in our study since the CAXX radar was temporarily shut down on some occasions. Supplementary maintenance was required due to the harsh climatic conditions in the Andean mountain range. Therefore, a complementary evaluation in the future is recommended.
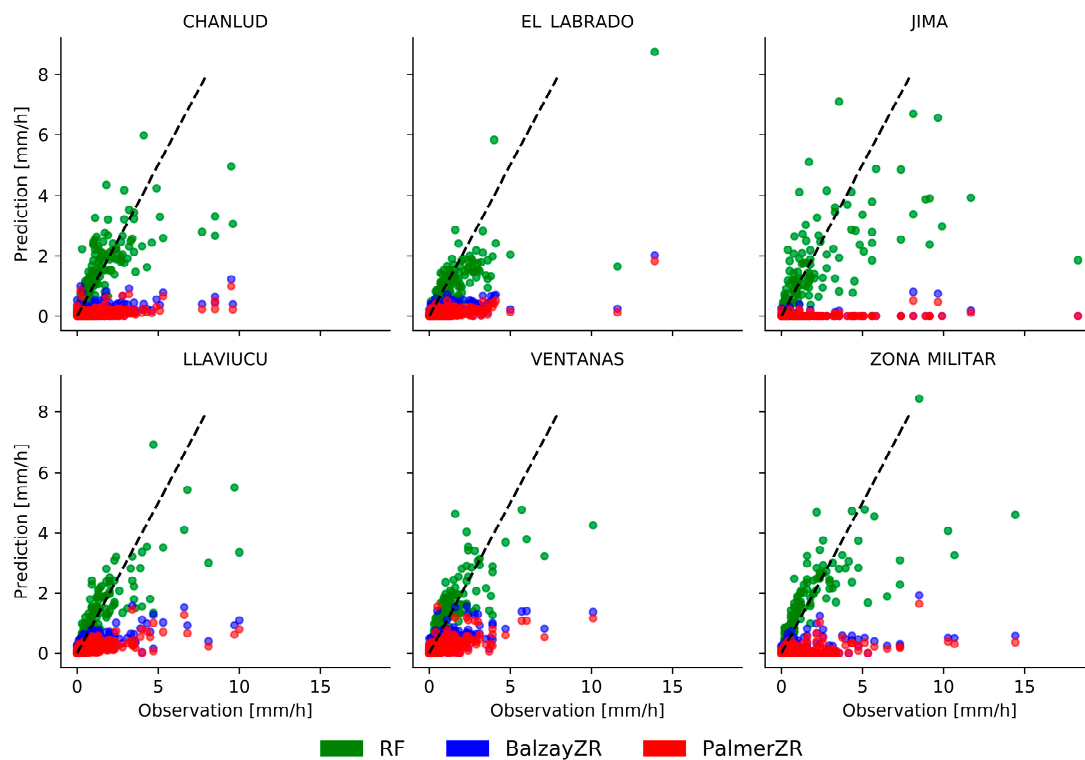
Figure 8. Scatter plot of observed rain rate vs. predicted rain rate for hourly accumulated rainfall at all six testing stations using the test data set (15 independent events). The 1-1 line is shown in black dotted. Three models are illustrated: random forest model (green), step-wise model using site-specific Balzay Z–R relation (blue) and step-wise model using Marshall-Palmer Z–R relation (red).

Table 4. Statistical measures of the performance of the different rainfall retrieval models at all test locations.

Metric	Model	Chanlud	El Labrado	Jima	Llaviucu	Ventanas	Z. Militar
	RF	0.97	0.90	4.30	0.78	0.55	1.85
MSE	BalzayZR	2.97	2.52	10.21	2.43	1.55	5.46
	PalmerZR	3.16	2.72	10.30	2.72	1.79	5.71
	RF	0.77	0.81	0.66	0.83	0.81	0.78
CC	BalzayZR	0.60	0.67	0.27	0.68	0.64	0.47
	PalmerZR	0.55	0.69	0.28	0.67	0.62	0.46
Efficiency	RF	0.58	0.57	0.38	0.67	0.65	0.57
Score	BalzayZR	−0.27	−0.19	−0.47	−0.03	0.02	−0.27
	PalmerZR	−0.35	−0.29	−0.48	−0.15	−0.13	−0.33
	RF	−0.01	−0.24	−0.27	−0.06	0.09	−0.05
rME	BalzayZR	−0.90	−0.86	−0.99	−0.73	−0.66	−0.89
	PalmerZR	−0.94	−0.92	−0.99	−0.84	−0.78	−0.93
	RF	0.91	0.93	1.11	0.85	0.78	0.96
rRMSE	BalzayZR	1.58	1.55	1.72	1.50	1.31	1.64
	PalmerZR	1.63	1.61	1.72	1.59	1.40	1.68
	RF	1.14	1.38	1.26	1.22	1.08	1.31
Slope	BalzayZR	4.77	4.72	7.31	3.67	2.47	3.77
	PalmerZR	6.03	6.62	12.00	4.78	2.99	5.10

Some studies such as [11] that did not use a rain gauge network for adjustment of the rainfall derivation but accomplished a similar or better model performance were mostly used in very short ranges (15–30 km). Therefore, despite their good performance, they would be of very limited

applicability in mountain regions as the Andes. For instance, this study differs from [11] in the evaluation at longer distances from radar, 60 km instead of 20 km and more significantly in the Z–R relationship performance. Contrary to [11] that used a unique event-based derived Z–R for the rainfall derivation of the Palermo radar, [1] has already demonstrated the high variability of Z–R relationships at the southern Andes of Ecuador and its impact on rainfall derivation.

　　Another difference in the evaluation in comparison to previous studies is the spatial resolution. Since CAXX radar imagery has not been transformed to cartesian coordinates, the effect of spatial averaging has been minimized and as a result, the inadequate assumption of good performance of a model due to a coarser spatial resolution. In other words, we are certain our results are unbiased regarding a point-pixel evaluation. This in turn makes difficult to compare our results with studies that used machine learning techniques with lower spatial resolution by means of S-, C-band radar data as in [26–28] (e.g., 1 km instead of 100 m) whose spatial resolution could minimize differences in rain gauge area coverage and radar volumetric content. While the latter is desirable, it also obscures the rainfall spatial variability which is remarkable in high mountain regions like the Andes. Furthermore, additional vertical profile of reflectivity may also contribute to the good performance (correlation coefficient ~0.8–0.9) of neural networks applications as in [26,27]. Nonetheless, also a lower performance (correlation coefficient of 0.37) was found by [28] in a test dataset from a new location 10 km apart from the radar. Thus, as already mentioned above, neural networks have their own drawbacks compared to a more simplified approach like random forest.

　　On the other hand, as expected, the step-wise approach by using both Z–R relationships showed a better adjustment for higher rain rates at the nearest stations Ventanas and Llaviucu (around 15 km from the radar) in comparison to other locations. X-band radars are prone to strong attenuation issues along the beam, thus rain rates at closer stations are more likely to properly be estimated by the step-wise model due to the lower attenuation influence at short distances as in [2].

　　In relation to the measurement errors, it can be seen in Table 5, that RF model performs the best, showing the lowest error in comparison to the BalzayZR and PalmerZR models. In general, lowest and highest rain rates are more sensitive to the use of radar parameters (i.e., reflectivity) than the mid-range rain rates ($1 < R < 10$). This can be explained by the fact that the proper detection of lower rain rates depends more on the sensitivity of the radar signal, while the quality of higher rain rate assessments is usually more affected by attenuation effects.

Table 5. Measurement errors for all rainfall retrieval models by considering different rain rate intervals.

Rain Rate [mm h^{-1}]	RF	BalzayZR	PalmerZR
<1	0.69	2.30	3.36
1 <= R < 2	0.20	1.16	1.52
2 <= R < 5	0.19	0.90	1.28
5 <= R < 10	0.16	1.15	1.43
>=10	0.24	0.72	1.01

　　Furthermore, the hourly basis evolution of the rainfall observations and estimations for the RF model at the three testing rainfall events, that were described in Section 2.4 as representative for different rainfall formation process, shows a slight overestimation of drizzle and underestimation of the peaks, which holds for all rainfall events as seen in Figure 9. Both models that used Z–R relationships, on the other hand, strongly and permanently underestimate the rainfall. Although, the RF model presents this jumping effect, the general tendency of the estimations remains similar to the observations despite the individual spatial and intensity characteristics of each rainfall event. This is key for the proper evaluation of the generalization of the model under different rainfall events and proves the promising performance of the RF model by using single PPI scans of the X-band radars in Ecuador.
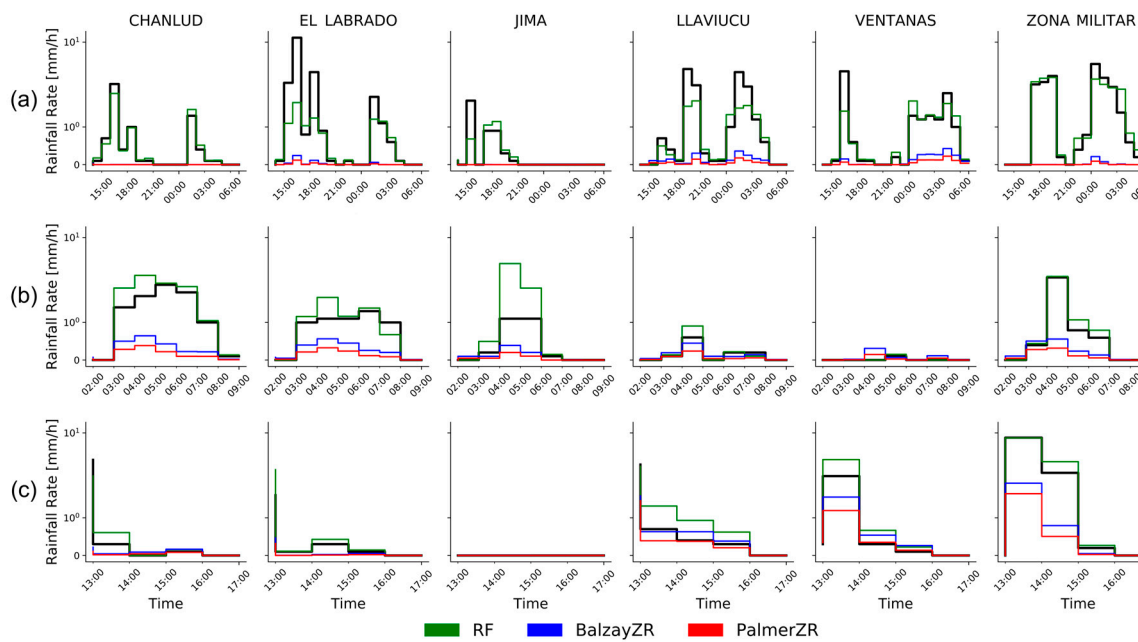
Figure 9. Hyetograms of the three representative rainfall events from the independent test set: (**a**) 09 March 2016 14:00–10 March 2016 07:30, (**b**) 10 June 2016 02:00– 10 June 2016 09:00 and (**c**) 19 March 2017 13:00–19 March 2017 17:00 as seen by the testing stations. Observations are illustrated in black and the three QPE models: random forest (green), step-wise Z–R site specific (blue) and step-wise Z–R Marshall-Palmer (red) at an hourly frequency.

Finally, yet importantly, the purposes of the radars and its QPE derivation along the studies discussed so far are quite different in terms of spatial resolution and coverage. While the great advantage of X-band radar is its high spatial resolution, several studies have limited the coverage of application at about 30 km. Moreover, in the case of 1st generation radars, they also had some advantages compared to CAXX since they were capable for entire 3D scans which allowed the use of the entire rain call, also in the vertical, as inputs for neural networks-based models. However, these were also compromised due to bright band influence at higher altitudes. Thus, as extensively described in [13], the interest of rainfall retrieval close the ground for future applications and the trade-off between high spatial resolution and accessible cost made the implementation of our radar system in the Andean region an optimal decision.

4. Conclusions

An optimized random forest model has been developed for rainfall retrieval of the highest single polarized X-band radar in the world. The performance of the random forest model was compared with the traditional step-wise methodology where reflectivity correction is performed at first in a step-by-step basis (i.e., clutter removal, attenuation correction) and then a Z–R relationship is applied for rainfall derivation. For the latter, two different Z–R relationships were used: Marshall-Palmer and a locally derived Z–R relationship [1]. From the results several conclusions can be drawn:

(i) As expected, the local derived Z–R relationship $Z = 103R^{2.06}$ had a slightly better adjustment than the Marshall-Palmer equation. Nonetheless, both step-wise approaches performed poorly for rainfall retrieval in the study site. It should be pointed out that the very low performance of the model may be an effect of the lack of bias adjustment by means of rain gauge data in the area. However, as emphasized early in the manuscript, the unavailability and uneven distribution of rain gauge data is very often a strong limitation to apply such adjustment. Moreover, it seems that the attenuation influence on the radar signal strongly minimizes the reflectivity signal along the beam and thus the derived rainfall rate is highly underestimated.

(ii) RF greatly outperformed the other models (CC up to 0.83; slope = 1.08) at all testing locations and across different rainfall events. It should be pointed out that the greatest advantage of RF model in comparison with the traditional approach lies on the fact that rain gauge data is not needed for the application of the model in the estimation.

(iii) Radar QPE in mountain regions often undertakes several limitations in terms of range usefulness, as a result of the amount of corrections performed to retrieve rainfall rates. To the contrary, the RF model used in this study has notably reduced the pre-processing and need of correction of radar data. Thus, the useful range may be increased for further analyses.

(iv) An RF machine learning model was successfully applied for radar rainfall retrieval. To the knowledge of the authors, this was also the first time a decision-tree based model was applied using features derived from a single polarized X-band radar data by means of single PPI scans. Despite some under-overestimations of the model, results of this study are promising. Its simplicity in terms of feature derivation, hyper-parameter adjustment and computational cost makes worthwhile its use in radar rainfall retrieval applications.

In summary, this study has explored the high spatial resolution of the CAXX radar imagery and has made the best possible use of the limited data (i.e., single PPI scans of reflectivity) by means of feature derivation and latter implementation of a QPE random forest model. The RF model showed encouraging results as an alternative to overcome the real-time unavailability of rain gauges in mountain regions as the Andes and attenuation issues inherent to X-band radars. Further work will focus on the application of the model in two twin X-band radar of the same radar network in Loja-Ecuador and Piura-Perú. Moreover, the derivation of hourly rainfall maps from the entire radar coverage and its potential use on hydrological and meteorological applications will be pursued. Finally, to better evaluate this machine learning model in rainfall retrieval applications, we highly encourage its accuracy assessment in polarimetric radars.

Abbreviations

The following abbreviations are used in this manuscript:

QPE	Quantitative Precipitation Estimation
RF	Random forest
ANN	Artificial Neural Networks
PPI	Plan Position Indicator
MSE	Mean square error

References

1.　Orellana-Alvear, J.; Célleri, R.; Rollenbeck, R.; Bendix, J. Analysis of Rain Types and Their *Z–R* Relationships at Different Locations in the High Andes of Southern Ecuador. *J. Appl. Meteorol. Climatol.* **2017**, *56*, 3065–3080. [CrossRef]

2.　Anagnostou, M.N.; Nikolopoulos, E.I.; Kalogiros, J.; Anagnostou, E.N.; Marra, F.; Mair, E.; Bertoldi, G.; Tappeiner, U.; Borga, M. Advancing precipitation estimation and streamflow simulations in complex terrain with X-Band dual-polarization radar observations. *Remote Sens.* **2018**, *10*. [CrossRef]

3.　Kirstetter, P.E.; Gourley, J.J.; Hong, Y.; Zhang, J.; Moazamigoodarzi, S.; Langston, C.; Arthur, A. Probabilistic precipitation rate estimates with ground-based radar networks. *Water Resour. Res.* **2015**, *51*, 1422–1442. [CrossRef]

4.　Büyükbas, E. Assess the Current and Potential Capabilities of Weather Radars for the Use in WMO Integrated Global Observing System (WIGOS). In Proceedings of the Joint Meeting of the CIMO Expert Team on Remote Sensing Upper-air Technology and Techniques and CBS Expert Team on Surface Based Remote Sensing, Geneva, Switzerland, 23–27 November 2009.

5.　McLaughlin, D.; Pepyne, D.; Chandrasekar, V.; Philips, B.; Kurose, J.; Zink, M.; Droegemeier, K.; Cruz-Pol, S.; Junyent, F.; Brotzge, J.; et al. Short-wavelength technology and the potential for distributed networks of small radar systems. *Bull. Am. Meteorol. Soc.* **2009**, 1797–1817. [CrossRef]

6.　Mishra, K.V.; Krajewski, W.F.; Goska, R.; Ceynar, D.; Seo, B.C.; Kruger, A.; Niemeier, J.J.; Galvez, P.A.; Thurai, M.; Bringi, V.N.; et al. Deployment and Performance Analyses of High-Resolution Iowa XPOL Radar System during the NASA IFloodS Campaign. *J. Hydrometeor.* **2016**, *17*, 455–479. [CrossRef]

7.　Feng, L.; Xiao, H.; Wen, G.; Li, Z.; Sun, Y.; Tang, Q.; Liu, Y. Rain Attenuation Correction of Reflectivity for X-Band Dual Polarization Radar. *Atmosphere* **2016**, *7*, 164. [CrossRef]

8.　Diederich, M.; Ryzhkov, A.; Simmer, C.; Zhang, P.; Trömel, S. Use of Specific Attenuation for Rainfall Measurement at X-Band Radar Wavelengths. Part II: Rainfall Estimates and Comparison with Rain Gauges. *J. Hydrometeorol.* **2015**, *16*, 503–516. [CrossRef]

9.　Villarini, G.; Krajewski, W.F. Review of the different sources of uncertainty in single polarization radar-based estimates of rainfall. *Surv. Geophys.* **2010**, *31*, 107–129. [CrossRef]

10.　Antonini, A.; Melani, S.; Corongiu, M.; Romanelli, S.; Mazza, A.; Ortolani, A.; Gozzini, B. On the Implementation of a regional X-bandweather radar network. *Atmosphere* **2017**, *8*, 25. [CrossRef]

11.　Lo Conti, F.; Francipane, A.; Pumo, D.; Noto, L.V. Exploring single polarization X-band weather radar potentials for local meteorological and hydrological applications. *J. Hydrol.* **2015**, *531*, 508–522. [CrossRef]

12.　Allegretti, M. X-Band Mini Radar for Observing and Monitoring Rainfall Events. *Atmos. Clim. Sci.* **2012**, *2*, 290–297. [CrossRef]

13.　Bendix, J.; Fries, A.; Zárate, J.; Trachte, K.; Rollenbeck, R.; Pucha-Cofrrep, F.; Paladines, R.; Palacios, I.; Orellana, J.; Oñate-Valdivieso, F.; et al. RadarNet-Sur first weather radar network in tropical high mountains. *Bull. Am. Meteorol. Soc.* **2017**, *98*, 1235–1254. [CrossRef]

14.　Rollenbeck, R.; Bendix, J. Rainfall distribution in the Andes of southern Ecuador derived from blending weather radar data and meteorological field observations. *Atmos. Res.* **2011**, *99*, 277–289. [CrossRef]

15.　Fries, A.; Rollenbeck, R.; Bayer, F.; Gonzalez, V.; Oñate-Valivieso, F.; Peters, T.; Bendix, J. Catchment precipitation processes in the San Francisco valley in southern Ecuador: Combined approach using high-resolution radar images and in situ observations. *Meteorol. Atmos. Phys.* **2014**, *126*, 13–29. [CrossRef]

16.　Oñate-Valdivieso, F.; Fries, A.; Mendoza, K.; Gonzalez-Jaramillo, V.; Pucha-Cofrep, F.; Rollenbeck, R.; Bendix, J. Temporal and spatial analysis of precipitation patterns in an Andean region of southern Ecuador using LAWR weather radar. *Meteorol. Atmos. Phys.* **2018**, *130*, 473–484. [CrossRef]

17.　McRoberts, D.B.; Nielsen-Gammon, J.W. Detecting Beam Blockage in Radar-Based Precipitation Estimates. *J. Atmos. Ocean. Technol.* **2017**, *34*, 1407–1422. [CrossRef]

18.　Berne, A.; Uijlenhoet, R.; Berne, A.; Quantitative, R.U. Quantitative analysis of X-band weather radar attenuation correction accuracy. *Nat. Hazards Earth Syst. Sci.* **2006**, *6*, 419–425. [CrossRef]

19.　Frasier, S.J.; Kabeche, F.; Figueras i Ventura, J.; Al-Sakka, H.; Tabary, P.; Beck, J.; Bousquet, O. In-Place Estimation of Wet Radome Attenuation at X Band. *J. Atmos. Ocean. Technol.* **2013**, *30*, 917–928. [CrossRef]

20. Van de Beek, C.Z.; Leijnse, H.; Stricker, J.N.M.; Uijlenhoet, R.; Russchenberg, H.W.J. Performance of high-resolution X-band radar for rainfall measurement in The Netherlands. *Hydrol. Earth Syst. Sci.* **2010**, *14*, 205–221. [CrossRef]

21. Van de Beek, C.Z.; Leijnse, H.; Hazenberg, P.; Uijlenhoet, R. Close-range radar rainfall estimation and error analysis. *Atmos. Meas. Tech* **2017**, *9*, 3837–3850. [CrossRef]

22. Harrison, D.L.; Driscoll, S.J.; Kitchen, M. Improving precipitation estimates from weather radar using quality control and correction techniques. *Meteorol. Appl.* **2000**, *6*, 135–144. [CrossRef]

23. Morin, E.; Gabella, M. Radar-based quantitative precipitation estimation over Mediterranean and dry climate regimes. *J. Geophys. Res.* **2007**, *112*, 1–13. [CrossRef]

24. Meyer, H.; Kühnlein, M.; Reudenbach, C.; Nauss, T.; Meyer, H.; Kühnlein, M.; Reudenbach, C.; Meyer, H.; Kühnlein, M.; Reudenbach, C.; et al. Revealing the potential of spectral and textural predictor variables in a neural network-based rainfall retrieval technique variables in a neural network-based rainfall retrieval technique. *Remote Sens. Lett.* **2017**, *8*, 647–656. [CrossRef]

25. Yang, T.H.; Feng, L.; Chang, L.Y. Improving radar estimates of rainfall using an input subset of artificial neural networks. *J. Appl. Remote Sens.* **2016**, *10*, 1–12. [CrossRef]

26. Alqudah, A.; Chandrasekar, V.; Le, M. Investigating rainfall estimation from radar measurements using neural networks. *Nat. Hazards Earth Syst. Sci.* **2013**, *13*, 535–544. [CrossRef]

27. Kusiak, A.; Wei, X.; Verma, A.P.; Member, S.; Roz, E. Modeling and Prediction of Rainfall Using Radar Reflectivity Data: A Data-Mining Approach. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 2337–2342. [CrossRef]

28. Teschl, R.; Randeu, W.L.; Teschl, F. Improving weather radar estimates of rainfall using feed-forward neural networks. *Neural Netw.* **2007**, *20*, 519–527. [CrossRef] [PubMed]

29. Orlandini, S.; Morlini, I. Artificial neural networks estimation of rainfall intensity from radar observations. *J. Geophys. Res.* **2000**, *105*, 849–861. [CrossRef]

30. Jing, W.; Zhang, P.; Jiang, H.; Zhao, X. Reconstructing Satellite-Based Monthly Precipitation over Northeast China Using Machine Learning Algorithms. *Remote Sens.* **2017**, *9*, 781. [CrossRef]

31. Yang, X.; Kuang, Q.; Zhang, W.; Zhang, G. A terrain-based weighted random forests method for radar quantitative. *Meteorol. Appl.* **2017**, *414*, 404–414. [CrossRef]

32. Yu, P.S.; Yang, T.C.; Chen, S.Y.; Kuo, C.M.; Tseng, H.W. Comparison of random forests and support vector machine for real-time radar-derived rainfall forecasting. *J. Hydrol.* **2017**, *552*, 92–104. [CrossRef]

33. Celleri, R.; Willems, P.; Buytaert, W.; Feyen, J. Space-time rainfall variability in the Paute Basin, Ecuadorian Andes. *Hydrol. Process.* **2007**, *21*, 3316–3327. [CrossRef]

34. Coltorti, M.; Ollier, C.D. Geomorphic and tectonic evolution of the Ecuadorian Andes. *Geomorphology* **2000**, *32*, 1–19. [CrossRef]

35. Campozano, L.; Célleri, R.; Trachte, K.; Bendix, J.; Samaniego, E. Rainfall and Cloud Dynamics in the Andes: A Southern Ecuador Case Study. *Adv. Meteorol.* **2016**, *2016*. [CrossRef]

36. Marra, F.; Morin, E. Use of radar QPE for the derivation of Intensity-Duration-Frequency curves in a range of climatic regimes. *J. Hydrol.* **2015**, *531*, 427–440. [CrossRef]

37. Heistermann, M.; Jacobi, S.; Pfaff, T. Technical Note: An open source library for processing weather radar data (wradlib). *Hydrol. Earth Syst. Sci.* **2013**, *17*, 863–871. [CrossRef]

38. Heistermann, M.; Collis, S.; Dixon, M.J.; Giangrande, S.; Helmus, J.J.; Kelley, B.; Koistinen, J.; Michelson, D.B.; Peura, M.; Pfaff, T.; et al. The emergence of open-source software for the weather radar community. *Bull. Am. Meteorol. Soc.* **2015**, *96*, 117–128. [CrossRef]

39. Gabella, M.; Notarpietro, R. Ground clutter characterization and elimination in mountainous terrain. In Proceedings of the European Conference on Radar Meteorology and Hydrology (ERAD) 2012, Toulouse, France, 24–29 June 2012; pp. 305–311.

40. Bech, J.; Codina, B.; Lorente, J.; Bebbington, D. The sensitivity of single polarization weather radar beam blockage correction to variability in the vertical refractivity gradient. *J. Atmos. Ocean. Technol.* **2003**, *20*, 845–855. [CrossRef]

41. Krämer, S.; Verworn, H.R. Improved radar data processing algorithms for quantitative rainfall estimation in real time. *Water Sci. Technol.* **2009**, *60*, 175–184. [CrossRef]

42. Jacobi, S.; Heistermann, M. Benchmarking attenuation correction procedures for six years of single-polarized C-band weather radar observations in South-West Germany. *Geomat. Nat. Hazards Risk* **2016**, *5705*. [CrossRef]

43. Marshall, J.S.; Palmer, W.M.K. The distribution of raindrops with size. *J. Meteorol.* **1948**, *5*, 165–166. [CrossRef]

44. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
45. Hedir, M.; Haddad, B. Automatic system for radar echoes filtering based on textural features and artificial intelligence. *Meteorol. Atmos. Phys.* **2016**. [CrossRef]
46. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michael, 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.
47. Anagnostou, M.N.; Kalogiros, J.; Marzano, F.S.; Anagnostou, E.N.; Montopoli, M.; Picciotti, E. Performance Evaluation of a New Dual-Polarization Microphysical Algorithm Based on Long-Term X-Band Radar and Disdrometer Observations. *J. Hydrometeorol.* **2013**, *14*, 560–576. [CrossRef]
48. Nash, J.E.; Sutcliffe, J.V. River flow forecasting through conceptual models part I—A discussion of principles. *J. Hydrol.* **1970**, *10*, 282–290. [CrossRef]
49. Thurai, M.; Mishra, K.V.; Bringi, V.N.; Krajewski, W.F. Initial Results of a New Composite-Weighted Algorithm for Dual-Polarized X-Band Rainfall Estimation. *J. Hydrometeorol.* **2017**, *18*, 1081–1100. [CrossRef]
50. Muñoz, P.; Célleri, R.; Feyen, J. Effect of the Resolution of Tipping-Bucket Rain Gauge and Calculation Method on Rainfall Intensities in an Andean Mountain Gradient. *Water* **2016**, *8*, 534. [CrossRef]