Exploitation of X-band weather radar data in the Andes high mountains and its application in hydrology: a machine learning approach

Doctoral Thesis by
Johanna Marlene Orellana Alvear
2020

Philipps Universität Marburg

Front cover: Landscape of the Andean mountain range at Paragüillas peak (4450 m a.s.l.).
Photo by Mario Guallpa, 2016

Back cover: Landscape at the Virgen station (3626 m a.s.l.).
Photo by Galo Carrillo-Rojas, 2016

Exploitation of X-band weather radar data in the Andes high mountains and its application in hydrology: a machine learning approach

kumulative Dissertation
zur
Erlangung des
DOCTORGRADES DER NATURWISSENSCHAFTEN
(Dr. rer. nat.)

dem Fachbereich Geographie
der PHILIPPS-UNIVERSITÄT MARBURG
vorgelegt von

JOHANNA MARLENE ORELLANA ALVEAR

aus Cuenca-Ecuador

Marburg / Lahn, 2020

Vom Fachbereich Geographie
der Philipps-Universität Marburg
als Dissertation am 03.11.2020 angenommen.

Erstgutachter: Prof. Dr. Jörg Bendix
Zweitgutachter: Prof. Dr. Peter Chifflard

Weitere Mitglieder der Prüfungskommission:

Prof. Dr. Maaike Bader
Prof. Dr. Dr. Thomas Brenner
Prof. Dr. Bernhard Seeger

Tag der mündlichen Prüfung am 13.01.2021
Hochschulkennziffer: 1180

Preface and acknowledgments

This journey has taken me throughout several faces of academic life. My doctoral thesis is the final step of this journey that has allowed me to become a better woman scientist and has provided me with challenging and extraordinary experiences in the last years. Finishing this work would not be possible without many people and institutions that support me along this path.

First of all, my heartfelt thanks are due to my mentors. I deeply thank my supervisor Jörg Bendix for guiding me with his invaluable, kind and inspiring mentorship, his constant orientation in our scientific discussions and for granting me the freedom to develop my research while keeping my project on track. My warmest gratitude goes to my co-supervisor Rolando Célleri for encouraging and supporting my involvement into activities far from my own doctoral project so I could grow up as an integral female researcher and for being a trustful mentor whose invaluable advice has helped me in many ways. I deeply thank you both for your time, patience and confidence and for making possible the financial support of my doctoral project along these years. It has been a privilege to work with you. I want to extent my special gratitude to the first mentor in my research life, Esteban Samaniego, who guided my early steps into this amazing adventure in the scientific world.

I want to express my gratitude to my advisor Rütger Rollenbeck for sharing his depth knowledge on radar science and for financially supporting the final stage of my doctoral thesis through his project. Thanks to the "Departamento de Recursos Hídricos y Ciencias Ambientales" (iDRHiCA), the "Dirección de Investigación" (DIUC) of the University of Cuenca, ETAPA EP and partner institutions for the logistics and projects funding as well as to the technical staff for their effort during field campaigns. My special gratitude to the ETAPA EP radar technician, Mario Guallpa, who has permanently greatly supported the operation of the CAXX radar despite the hard circumstances. I am also grateful to the Deutsche Forschungsgemeinschaft (DFG) and to the Laboratory for Climatology and Remote Sensing (LCRS) of the Philipps-Universität Marburg, which supported this study in human, financial and scientific capacities.

My warmest thanks go to my co-authors, colleagues and friends of iDRHiCA who have contributed to a friendly working environment. It has positively influenced in the many years of my doctoral studies. I enjoyed very much our seminars, our after office meetings, our holiday programs and most of all having the opportunity to meet you all in my life. I am equally thankful with my colleagues from the LCRS who made my research stays in Germany more enjoyable away from home. I highly appreaciate your kindness and I will never forget our time in Marburg. My special thanks goes to Birgit for making paperwork easier while traveling to Germany. Your support has been very valuable and I will always treasure your friendship.

I want to thank the "Red Ecuatoriana de Mujeres Científicas", specially to my colleagues and friends, Dani, Xime and Ana Elizabeth, you have always been a source of inspiration, encouragement and sorority during this journey. Your friendship has boosted my energy and giving me a second air on challenging times.The same feeling goes to my dearest friend, Prisci, who has always been checking on me and has undoubtely trust in me for achiving this goal.

My deepest gratitude goes to my family, my mom Ligia, my brothers Esteban and Boris, my sister Marigu, my sisters in law, Ligia and Mónica and my beloved nieces and nephew. You have always remained my major strength and my main motivation for achieving my dreams. I admire you so much.

Finally, I will be eternally thankful to my wonderful husband Galo José, my partner in life and in science. The most beautiful and tough memories of this journey I have shared with you. I am blessed for having you in my life. Thank you for walking with me not only as a supporting husband and colleague but mainly as an unconditional friend and for being my greatest ally on this academic and life challenge. We faced it together as we will face many more in the future. I love you.

Johanna Marlene Orellana Alvear
November 2020.

Summary

Rainfall in the tropical Andes high mountains is paramount for understanding complex hydrological and ecological phenomena that take place in this distinctive area of the world. Here, rainfall drives imminent hazards such as severe floods, rainfall-induced landslides, different types of erosion, among others. Nonetheless, sparse and uneven distributed rain gauge networks as well as low-resolution satellite imagery are not sufficient to capture its high variability and complex dynamics in the irregular topography of high mountains at appropriate temporal and spatial scales. This results in both, a lack of knowledge about rainfall patterns, as well as a poor understanding of rainfall microphysics, which to date are largely underexplored in the tropical Andes.

Therefore, this investigation focuses on the deployment and exploitation of single-polarization (SP) X-band weather radars in the Andean high mountain regions of southern Ecuador, applicable to quantitative precipitation estimation (QPE) and discharge forecasting. This work leverages radar rainfall data by exploring a machine learning (ML) approach. The main aims of the thesis were: (i) The deployment of a first X-band weather radar network in tropical high mountains, (ii) the physically-based QPE of X-band radar retrievals, (iii) the optimization of radar QPE by using a ML-based model and (iv) a discharge forecasting application using a ML-based model and SP X-band radar data.

As a starting point, deployment of the first weather radar network in tropical high mountains was carried out. A complete framework for data transmission was set for communication among the network. The highest radar in the network (4450 m a.s.l.) was selected in this study for exploiting the potential of SP X-band radar data in the Andes. First and foremost, physically-based QPE was performed through the derivation of Z-R relationships. For this, data from three disdrometers at different geographic locations and elevation were used. Several rainfall events were selected in order to perform a classification of rainfall types based on the mean volume diameter (Dm [mm]). Derived Z-R relations confirmed the high variability in their parameters due to different rainfall types in the study area. Afterwards, the optimization of radar QPE was pursued by using a ML approach as an alternative to the common physically-based QPE method by means of the Z-R relation. For

this, radar QPE was tackled by using two different approaches. The first one was conducted by implementing a step-wise approach where reflectivity correction is performed in a step-by-step basis (i.e., clutter removal, attenuation correction). Finally a locally derived Z-R relationship was applied for obtaining radar QPE. Rain gauge-bias adjustment was neglected because the availability of rain gauge data at near-real time is limited and infrequent in the study area. The second one was conducted by an implementation of a radar QPE model that used the Random Forest (RF) algorithm and reflectivity derived features as inputs for the model. Finally, the performances of both models were compared against rain gauge data. The results showed that the ML-based model outperformed the step-wise approach, making it possible to obtain radar QPE without the need of rain gauge data after the model was implemented. It also allowed to extend the useful range of the radar image (i.e., up to 50 km).

Radar QPE can be generally used as input for discharge forecasting models if available. However, one could expect from ML-based models as RF, the ability to map radar data to the target variable (discharge) without any intermediate step (e.g., transformation from reflectivity to rainfall rate). Thus, a comparison for discharge forecasting was performed between RF models that used different input data type. Input data for the relevant models were obtained either from native reflectivity records (i.e., reflectivity corrected from unrealistic measurements) or derived radar-rainfall data (i.e., radar QPE). Results showed that both models performed alike. This proved the suitability of using native radar data (reflectivity) for discharge forecasting in mountain regions. This could be extrapolated in the advantages of deploying radar networks and use their information directly to fed early-warning systems regardless of the availability of rain gauges at ground.

In summary, this investigation (i) participated on the deployment of the first weather radar network in tropical high mountains, (ii) significantly contributed to a deeper understanding of rainfall microphysics and its variability in the high tropical Andes by using disdrometer data and (iii) exploited, for the very first time, the native X-band radar reflectivity as a suitable input for ML-based models for both, optimized radar QPE and discharge forecasting. The latter highlighted the benefits and potentials of using a ML approach in radar hydrology. The research generally accounted for ground monitoring limitations commonly found in mountain regions and provided a promising alternative with leveraging the cost-effective X-band technology in the steep terrain of the Andean Cordillera.

Zusammenfassung

Die Niederschläge im Hochgebirge der tropischen Anden sind für das Verständnis komplexer hydrologischer und ökologischer Phänomene, die sich in dieser besonderen Region der Welt abspielen, von größter Bedeutung. Hier verursachen die Niederschläge drohende Gefahren wie schwere Überschwemmungen, Erdrutsche sowie verschiedene Arten von Erosion. Dennoch reichen dünn und ungleichmäßig verteilte Regenmessnetze sowie Satellitenbilder mit geringer Auflösung nicht aus, um die hohe Variabilität und die komplexe Dynamik vom Niederschlägen in der unregelmäßigen Topographie des Hochgebirges in geeigneten zeitlichen und räumlichen Maßstäben zu erfassen. Dies begründet bis heute ein mangelhaftes Verständnis über Niederschlagsmuster und insbesondere über die Niederschlagsmikrophysik, die in den tropischen Anden bisher weitgehend unerforscht ist.

Daher konzentriert sich die vorliegende Untersuchung auf den Einsatz und die Nutzung von einfach polarisierten (SP) X-Band Wetterradaren in den Hochgebirgsregionen der Anden im Süden Ecuadors, die für die quantitative Niederschlagsschätzung (QPE) und für Abflussvorhersagen anwendbar sind. Die Arbeit nutzt die Radar-Niederschlagsdaten mit Hilfe eines maschinellen Lernansatzes (ML). Die Hauptziele der Arbeit sind: (i) Der Einsatz eines ersten X-Band Wetterradarnetzes in tropischen Hochgebirgen, (ii) die physikalisch-basierte QPE von X-Band Radaradaten, (iii) die Optimierung der Radar-QPE durch Verwendung eines ML-basierten Modells und (iv) eine Anwendung der Daten auf die Abflussvorhersage in einem andinen Einzugsgebiet unter Verwendung eines ML-basierten Modells.

Als Ausgangspunkt wurde des ersten Wetterradarnetzes in einem tropischen Hochgebirge etabliert. Für die Kommunikation im Netzwerk wurde eine netz- und funkbasierte Datenübertragung realisiert. Die vorliegende Studie basiert auf dem höchstgelegenen Radar des Netzwerkes (4450 m ü.d.M.). In einer ersten Analyse wird die physikalisch basierte QPE durch die Ableitung von Z-R-Beziehungen entwickelt. Dazu werden Daten von drei Disdrometern an verschiedenen geographischen Standorten und in unterschiedlicher Höhe verwendet. Mehrere Niederschlagsereignisse wurden ausgewählt, um eine Klassifizierung der Niederschlagsarten auf der Grundlage des mittleren Tropfendurchmessers (Dm [mm]) durchzuführen. Abgeleitete Z-R-Beziehungen zeigen die hohe Variabilität ihrer Param-

eter aufgrund unterschiedlicher Niederschlagsarten im Untersuchungsgebiet.

Anschließend wird die Optimierung der Radar-QPE mit einem ML-Ansatz als Alternative zu der üblichen Z-R-basierten QPE-Methode verfolgt. Hierfür werden zwei verschiedene Ansätze getestet. Der erste Ansatz implementiert zur Radar-QPE eine schrittweise Durchführung der Reflektivitätskorrektur (d.h. Störungsbeseitigung, Dämpfungskorrektur) und die Ableitung einer lokalen Z-R-Beziehung. Eine Bias-Kalibrierung wurde nicht durchgeführt, da die Verfügbarkeit von stationsbasierten Regendaten in Echtzeit v.a. im Untersuchungsgebiet sehr begrenzt ist. Der zweite Ansatz implementiert ein Radar-QPE-Modell auf der Basis des Random Forest (RF)-Algorithmus und der reinen Radarreflektivität (Z). Der Vergleich beider Ansätze mit stationsbasierten Niederschlagsdaten zeigt, dass das ML-basierte Modell die Radar-QPE der schrittweisen Implementierung deutlich verbessert. Das ML-basierte Verfahren erlaubt es auch, die nutzbare Reichweite des Radars zu erweitern (d.h. bis zu 50 km).

Grundsätzlich sind Daten der Radar-QPE als Antrieb für Abflussvorhersagemodelle verwendbar. Ziel der vorliegenden Arbeit ist es aber, mit Hilfe ML-basierter Modelle wie RF die Radardaten ohne Zwischenschritt (Transformation von Reflektivität Z in Regenrate R) auf die Zielvariable (Abfluss) abzubilden. Um die Qualität eines solchen Vorgehens zu testen, wird ein Vergleich verschiedener Abflussvorhersagemodelle durchgeführt, die unterschiedliche Eingabedatentypen des Radarsystems verwenden. Im ersten Fall wird die native Radarreflektivität (die Reflektivität wurde lediglich gegen unrealistische Messungen wie Festzielechos korrigiert), im zweiten Fall Radar-Regendaten (d.h. aus der Radar-QPE) als Antriebsvariable eingesetzt. Die Ergebnisse zeigen, dass beide Modelle gleich gut funktionierten. Somit können native Radardaten (Reflektivität) direkt für Abflussvorhersagen in Bergregionen verwendet werden, ohne sie vorher in Regenraten umrechnen zu müssen. Der Befund ist für ganze Radarnetzwerke gültig, die somit unabhängig von der Verfügbarkeit von stationsgebundenen Regenmessern direkt für Abfluss- Frühwarnsysteme nutzbar sind.

Zusammenfassen hat die vorliegende Arbeit (i) zum Aufbau des ersten Wetterradarnetzes im tropischen Hochgebirge beigetragen, (ii) durch die Verwendung von Disdrometerdaten wesentlich zu einem tieferen Verständnis der Niederschlagsmikrophysik und ihrer Variabilität in den tropischen Hochanden geführt und (iii) zum ersten Mal erfolgreich die native X-Band-Radarreflektivität als geeigneten Input für ML-basierte Modelle sowohl für eine optimierte Radar-QPE als auch für die Abflussvorhersage genutzt. Letzteres zeigt die großen Vorteile und Potentiale der Verwendung eines ML-Ansatzes in der Radarhydrologie peripherer Gebiete mit komplexer Topographie. Die Arbeit überwindet damit die in Bergregionen häufig anzutreffenden Einschränkungen bei der stationsgestützten Überwachung von Niederschlägen und stellte insbesondere bei der

Verwendung der kostengünstigen X-Band-Technologie eine vielversprechende Alternative für Umweltüberwachungssysteme im komplexen Gelände der Andenkordillere dar.

Resumen

La lluvia en las altas montañas de los Andes tropicales es primordial para comprender los complejos fenómenos hidrológicos y ecológicos que tienen lugar en esta particular zona del mundo. Aquí, las lluvias provocan peligros inminentes tales como graves inundaciones, deslizamientos de tierra inducidos por las lluvias, diferentes tipos de erosión, entre otros. Sin embargo, la escasas redes de pluviómetros y su distribución heterogénea, así como las imágenes satelitales de baja resolución no son suficientes para captar su gran variabilidad y su compleja dinámica a escalas temporales y espaciales apropiadas en la irregular topografía montañosa. Esto da lugar tanto a una falta de conocimiento sobre los patrones de las precipitaciones, como a una deficiente comprensión de la microfísica de las mismas, que hasta la fecha han sido muy poco exploradas en los Andes tropicales.

Por lo tanto, esta investigación se centra en la implementación y el aprovechamiento de radares meteorológicos de banda-X con polarización simple (SP por sus siglas en inglés) en las regiones andinas de alta montaña del sur del Ecuador, aplicables a la estimación cuantitativa de la precipitación (QPE por sus siglas en inglés) y al pronóstico de caudales. Este trabajo aprovecha los datos de precipitaciones del radar explorando un enfoque de aprendizaje automático (ML por sus siglas en inglés). Los principales objetivos de la tesis fueron: i) La implementación de una primera red de radares meteorológicos banda-X en las altas montañas tropicales, ii) la estimación cuantitativa de precipitación de radar banda-X basado en ecuaciones físicas, iii) la optimización de la estimación cuantitativa de precipitación del radar utilizando un modelo basado en aprendizaje automático y iv) una aplicación de predicción de caudales utilizando un modelo basado en aprendizaje automático y datos de radar banda-X de polarización simple.

Como punto de partida, se llevó a cabo la implementación de la primera red de radares meteorológicos en las altas montañas tropicales. Se estableció un marco completo de transmisión de datos para la comunicación entre la red. En este estudio se seleccionó el radar más alto de la red (4450 m s.n.m.) para explotar el potencial de los datos de radar banda-X de polarización simple en los Andes. En primer lugar, la estimación cuantitativa de precipitación basada en ecuaciones físicas se realizó a través de la derivación de relaciones Z-R (reflectividad-tasa de lluvia). Para ello, se utilizaron datos

de tres disdrómetros en diferentes ubicaciones geográficas y elevaciones. Se seleccionaron varios eventos de precipitaciones para realizar una clasificación de tipos de lluvia basada en el diámetro de volumen medio (Dm [mm]). Las relaciones Z-R derivadas confirmaron la alta variabilidad de sus parámetros debido a los diferentes tipos de lluvia en la zona de estudio.

Posteriormente se procedió a la optimización de la estimación cuantitativa de precipitación del radar utilizando un enfoque de aprendizaje autmático como alternativa al método común basado en ecuaciones físicas mediante la relación Z-R. Para ello, la estimación cuantitativa de precipitación del radar se abordó utilizando dos enfoques diferentes. El primero se llevó a cabo mediante la aplicación de un enfoque sistemático en el cual la corrección de la reflectividad se realiza paso a paso (i.e., eliminación de ruido, corrección de la atenuación). Luego, se aplicó una relación Z-R derivada localmente para obtener la estimación cuantitativa de precipitación del radar. Se omitió el ajuste de sesgo usando pluviómetros debido a que la disponibilidad de datos de los pluviómetros en tiempo real es limitada y poco frecuente en la zona de estudio. El segundo enfoque se realizó mediante la aplicación de un modelo de estimación cuantitativa de precipitación de radar que utilizó el algoritmo de bosque aleatorio (RF por sus siglas en inglés) y las características derivadas de la reflectividad como entradas para el modelo. Por último, se compararon los resultados de ambos modelos con los datos de los pluviómetros. Los resultados mostraron que el modelo basado en aprendizaje automático superaba el enfoque sistemático, lo que permitía obtener la estimación cuantitativa de precipitación del radar sin necesidad de datos del pluviómetro después de la implementación del modelo. Esto también permitió ampliar el alcance útil de la imagen de radar (i.e., hasta 50 km).

La estimación cuantitativa de precipitación del radar puede utilizarse generalmente como entrada para los modelos de pronóstico de caudales, si se dispone de ellos. Sin embargo, se podría esperar de los modelos basados en aprendizaje automático tales como RF, la capacidad de transformar los datos del radar hacia la variable objetivo (caudal) sin ningún paso intermedio (i.e., transformación de la reflectividad a tasa de lluvia). Así pues, se realizó una comparación del pronóstico de caudales entre modelos RF que utilizaron diferentes tipos de datos de entrada. Los datos de entrada para los modelos se obtuvieron a partir de registros de reflectividad nativa (i.e., reflectividad corregida como resultado de control de calidad de mediciones no realistas) o de datos de las precipitaciones derivadas de radar (i.e., estimación cuantitativa de precipitación de radar — QPE). Los resultados mostraron que ambos modelos funcionaban de manera similar. Esto demostró la idoneidad de utilizar los datos nativos del radar (reflectividad) para el pronóstico de caudales en las regiones montañosas. Esto podría extrapolarse a las ventajas de implementar redes de radar y utilizar su información directamente para alimentar los sistemas de alerta temprana, independientemente de la

disponibilidad de pluviómetros en tierra.

En resumen, esta investigación: i) participó en la implementación de la primera red de radares meteorológicos en las altas montañas tropicales, ii) contribuyó significativamente a una comprensión más profunda de la microfísica de las precipitaciones y su variabilidad en los altos Andes tropicales mediante el uso de datos de disdrómetro y iii) explotó, por primera vez, la reflectividad nativa de un radar banda-X como una fuente de datos de entrada adecuada para los modelos basados en aprendizaje automático, tanto para la optimización de la estimación cuantitativa de precipitación de radar como para el pronóstico de caudales. Este último evidenció las ventajas y posibilidades de utilizar un enfoque de aprendizaje automático en las aplicaciones hidrológicas que hacen uso de datos de radar. La investigación consideró en términos generales las limitaciones de monitoreo en tierra que es común en las regiones montañosas y proporcionó una alternativa prometedora a partir del aprovechamiento y accesibilidad de la tecnología de banda-X en el escarpado terreno de la Cordillera de los Andes.

Table of contents

List of Figures

List of Tables

List of Acronyms and Symbols

ACF	Auto Correlation Function
ANN	Artificial Neural Network
AWS	Automatic Weather Stations
CASA	Collaborative Adaptive Sensing of the Atmosphere
CC	Correlation Coefficient
DEM	Digital Elevation Model
DP	Dual-Polarization
DSD	Drop Size Distribution
Dm	Mean Volume Diameter [mm]
ECMWF	European Centre for Medium-Range Weather Forecasts
ECSF	Estación Científica San Francisco
EN	El Niño
ENSO	El Niño Southern Oscillation
ETAPA EP	Empresa Pública Municipal de Telecomunicaciones, Agua Potable, Saneamiento y Alcantarillado de Cuenca
GIS	Geographic Information System
GPL	Gobierno Provincial de Loja
INAMHI	Instituto Nacional de Meteorología e Hidroloía
iFloodS	Ground Validation Iowa Flood Studies
KGE	Kling-Gupta Efficiency
LN	La Niña
LAWR	Local Area Weather Radar

LPM . . . . . Laser Precipitation Monitor

MARE . . . . . Mean Absolute Relative Error

ML . . . . . Machine Learning

NEXRAD . . . . . Next Generation Weather Radar

NOAA . . . . . National Oceanic and Atmospheric Administration

NSE . . . . . Nash-Sutcliffe Efficiency

ONI . . . . . Oceanic Niño Index

PACF . . . . . Partial Auto Correlation Function

PBIAS . . . . . Percentage Bias

PDO . . . . . Pacific Decadal Oscillation

PIA . . . . . Path Integrated Attenuation

PPI . . . . . Plan Position Indicator

R . . . . . Rainfall Rate

RF . . . . . Random Forest

RMSE . . . . . Root Mean Squared Error

SGR . . . . . Secretaría de Gestión de Riesgos

SLP . . . . . Sea Level Pressure

SP . . . . . Single-Polarization

SST . . . . . Sea Surface Temperature

QPE . . . . . Quantitative Precipitation Estimation

UNESCO . . . . . United Nations Educational, Scientific and Cultural Organization

WMO . . . . . World Meteorological Organization

Z . . . . . Reflectivity

1 Introduction

1.1 Motivation

Rainfall is a complex meteorological phenomenon that directly influences human life. Applications related to ecohydrological services, potable water and risk management highly depend on an adequate rainfall characterization. This is particularly true in regions with irregular topography and steep terrain as those of the Andean Cordillera (Hamel et al., 2018). Nonetheless, despite the importance of studying this phenomenon, rainfall monitoring and thus, scientific knowledge on spatio-temporal rainfall pattern is still insufficient in tropical mountain regions. Several works (Buytaert et al., 2006; Campozano et al., 2016; Contreras et al., 2019; Xie et al., 2011) have revealed deficiencies of sparse rain gauge networks at detecting the spatial variability of rainfall. On the other hand, increasing the density of rain gauge monitoring is both, expensive and logistically challenging because of the inaccessibility to remote mountain areas. Although satellite imagery can partially overcome the issue of obtaining an appropriate rainfall monitoring regarding its spatial distribution, the coarse resolution of satellite images is still a limitating factor for an adequate representation of rainfall in Andean regions (Ballari et al., 2016). It is because the high variability of rainfall (Célleri et al., 2007; Espinoza Villar et al., 2009; Pulwarty et al., 1998; Rollenbeck and Bendix, 2011; Scheel et al., 2011) is related to orographic enhancement and complex rainfall formation processes (Poveda et al., 2005) such as mesoscale convective dynamics particularly important in the tropics, lee confluence, lee divergence, channelization etc. Here, passive remote sensing instruments are found insufficient for representing tropical rainfall in high mountains (Ballari et al., 2018).

As an alternative, ground based weather radars have been deployed worldwide for the last decades. Its high resolution in time and space provides in deep insights about rainfall dynamics and may be of great benefit for hydrometeorological applications. The high spatio-temporal resolution of ground radars allows to retrieve and analyze the variability of rainfall events with great detail (Avolio et al., 2019; Borga et al., 2019; Bronstert et al., 2018). Several radar networks around the world have been used to reinforce hydrometeorological monitoring, generate probabilistic quantitative precipitation

estimations (Kirstetter et al., 2015) and provide high resolution rainfall data as a valuable asset for hydrological modeling (Anagnostou et al., 2018). Nonetheless, most radar networks (C- and S-band) have been deployed mainly in developed countries as a result of their large costs. Fortunately, the repurposing of single-polarization (SP) X-band marine radar (Jensen, 2002) has favored the advent of a new generation of X-band radar ground networks (Lengfeld et al., 2014) which are a more accessible solution that comes with reduced costs but also with a shorter range of coverage. Nonetheless, X-band radars are usually preferred for urban applications (Manola et al., 2020; Ochoa-Rodriguez et al., 2019) because of their high spatial resolution (e.g., 100 m). Implementation of SP X-band radar networks worldwide has been successfully performed such as in Allegretti (2012); Antonini et al. (2017); Lo Conti et al. (2015).

The absence of high-resolution equipment and uneven distribution of rainfall instruments in mountain Andean regions (see Figure 1.1) has caused a lack of understanding of rainfall spatial patterns (Oñate-Valdivieso et al., 2018) and limited knowledge of rain types (Bendix et al., 2006) in this important area of the world. As a result, rainfall in tropical mountain regions as those of the southern Andes of Ecuador have been underexplored (Wohl et al., 2012) despite of the occurrence of important climatic phenomena such as El-Niño Southern Oscillation (ENSO) that constantly affects precipitation (Bendix, 2000; Bendix et al., 2003; Rossel and Cadier, 2009; Vuille et al., 2000). Thereby, the deployment of radar technology would greatly benefit the rainfall monitoring in high mountain Andean regions. Nonetheless, radar rainfall retrievals are obtained after a process that allows converting radar reflectivity into rainfall rate (McKee and Binns, 2016). Approaches for deriving quantitative precipitation estimation (QPE) from radar imagery commonly fit the local geographic conditions making this process an individual tailor-made solution. This can be exacerbated in mountain regions (Germann et al., 2006; Rollenbeck and Bendix, 2006, 2011; Savina, 2011) because the radar beam issues can scale due to the irregular and steep terrain. Thus, more general solutions that can be extrapolated to other regions of the world still need to be addressed.

In addition, early warning systems face several restrictions regarding the feasibility of obtaining accurate water related predictions (e.g., rainfall, discharge) for an appropriate risk management. A better understanding of rainfall dynamics and its exploitation on different applications that provide decision makers with reliable information is urgent and of extreme value. Radar rainfall estimates (McKee and Binns, 2016; Ochoa-Rodriguez et al., 2019) are a powerful source of data for streamflow forecasting models. As radar rainfall retrieval capture the spatial variability of rainfall, it allows discharge models to better represent the rainfall inputs (Emmanuel et al., 2015; Lobligeois et al., 2014). Flood forecasting in mountain Andean catchments may greatly benefit from the

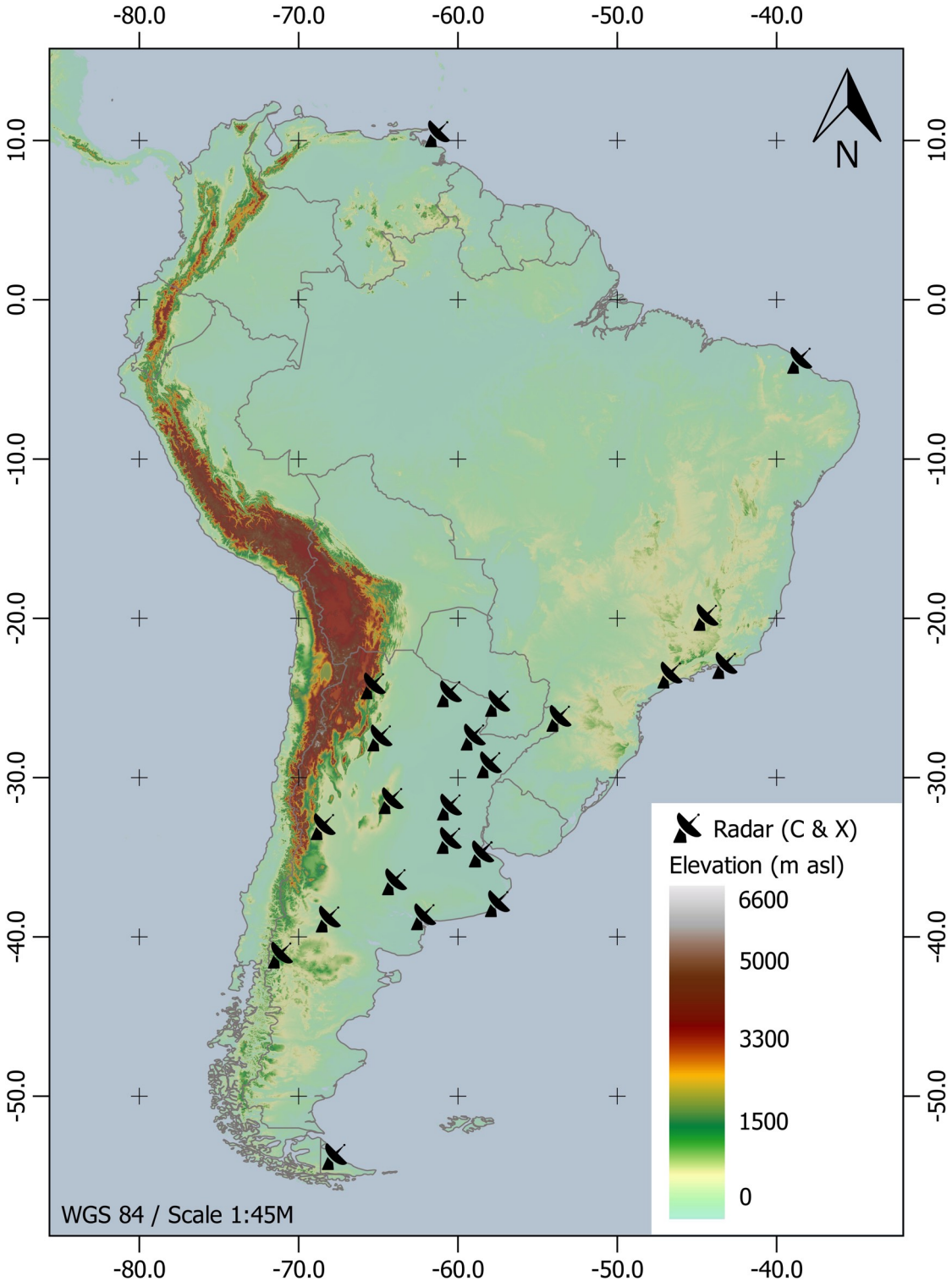

Figure 1.1: C- and X- band ground radar distribution in South America. Adapted from the World Meteorological Organization (WMO) https://wrd.mgm.gov.tr/Home/Wrd

high resolution rainfall from ground radar. This is because scarce and unevenly distributed rain gauge networks restrict an accurate spatial representation of rainfall an thus negatively influence on discharge forecasting models (Sucozhañay and Célleri, 2018).

Altogether, exploration of rainfall at high spatio-temporal resolution in Andean mountain catchments would be a tangible contribution to the hydrometeorological knowledge and a valuable asset for hydrological applications such as discharge forecasting. This would provide stakeholders and decision makers with an important tool for risk management. As a consequence, the use of radar imagery is of great interest and relevance in this geographic location of the world. Therefore, this thesis aims to exploit X-band radar imagery in a key mountainous location of the Andean Cordillera. Derivation of rainfall retrievals and the use of radar rainfall on discharge forecasting, as one of the most common hydrological applications in Andean mountain regions, will be pursued.

1.2 Current use of X-band weather radar

X-band radar technology emerged as an alternative solution for high spatio-temporal monitoring of rainfall. It allows a higher spatial resolution (e.g., 100 m) but also has a smaller range compared to C- or S- band radar. As a result, it has been largely used for urban rainfall monitoring and related hydrological applications (Manola et al., 2020; Ochoa-Rodriguez et al., 2019). Moreover, X-band radar has been a powerful tool for meteorological applications where the observations of mesoscale processes such as convective thunderstorms is of high significance. Therefore, it is especially useful for the tropics where small-scale convection plays a major role in rainfall pattern.

The use of X-band radars worldwide has increased up to 20% according to Büyükbaş (2009). Radar networks have been deployed mainly with dual-polarization technology. For instance, the Collaborative Adaptive Sensing of the Atmosphere (CASA) as in McLaughlin et al. (2009), or the hydrology-oriented Iowa Flood Studies (iFloodS) field campaign as in Mishra et al. (2016). Due to its feasibility as a cost-effective solution for rainfall monitoring, SP X-band radars have also been deployed worldwide. For instance, Local Area Weather Radar (LAWR), has satisfactorily been used in Belgium (Goormans, 2011), Switzerland (Savina, 2011) and Ecuador (Fries et al., 2014; Oñate-Valdivieso et al., 2018; Rollenbeck and Bendix, 2011). However, the major limitation of SP X-band radars is related to attenuation issues (Berne et al., 2006; Frasier et al., 2013; Villarini and Krajewski, 2010) which reduce the range of usability of its data.

1.2.1 Methods on Radar Rainfall Derivation

The reflectivity (Z) variable recorded by SP X-band radars is an indirect measurement of rainfall (R). Therefore, quantitative precipitation estimation (QPE) involves a process that allows converting the radar reflectivity into rainfall rate. The most prevalent approach for radar QPE often includes several steps regarding reflectivity correction and upscaling, the application of a Z-R relationship and finally, a bias adjustment by using rain observations at ground. This is hereafter referred as step-wise approach.

The complex orography of tropical mountain regions challenges the use of this step-wise approach. This is because each correction step (e.g., clutter removal, beam blockage) although can reduce undesired effects on the reflectivity records (Berne et al., 2006; Frasier et al., 2013; McRoberts and Nielsen-Gammon, 2017), it also produces adverse side effects on the radar image (i.e., mistaken changes on reflectivity values in other areas of the radar image; Harrison et al., 2000). Moreover, uncertainties related to the use of a Z-R relationship derived on a different location can negatively impact on the accuracy of radar QPE at the study site of interest.

The Z-R relationship is a power law ($Z=AR^b$) defined by Marshall and Palmer (1948) where A and b are empirical parameters and are referred to as coefficient and exponent, respectively. They are found from a reflectivity time series and the corresponding rainfall at ground. Z-R relations are frequently locally derived in order to account for the rainfall processes on the specific site (Lo Conti et al., 2015). Furthermore, usually tailor-made solutions (Fries et al., 2014; Lo Conti et al., 2015) are designed to ensure a good correlation with rain gauge observations. Therefore, normally a bias adjustment is performed by using dense rain gauge networks (e.g., Morin and Gabella, 2007). Rainfall data at ground is used to generate interpolated maps and afterwards, blending approaches as in Ochoa-Rodriguez et al. (2019) are commonly applied for obtaining more accurate rainfall estimations.

There are several approaches on the Z-R relationship derivation. Some are focused on the use of a unique climatic or local derived relationship that account for the radar rainfall retrieval in the entire study site. Nonetheless, the effects of altitudinal gradients on Z-R relation parameters (Tokay et al., 2009) have not been fully considered in mountain regions. In contrast, other studies focus on the variability of rainfall microphysics (i.e., drop size distribution) and thus provide rain-type Z-R relations under different criteria that define several rainfall classes. Those can be related to rainfall rate thresholds (Cerro et al., 1997; Sen Jaiswal et al., 2009), drop size distribution (Atlas et al., 2000; Caracciolo et al., 2008; Nzeukou and Sauvageot, 2004; Ochou et al., 2011; Rosenfeld et al.,

1993; Uijlenhoet, 2001; Uijlenhoet et al., 2003), etc. As a consequence, plenty of Z-R relations are documented for different rainfall types and locations. However, those of tropical regions are mainly documented for lowlands (Bamba et al., 2014; Kumar et al., 2011; Narayana Rao et al., 2001; Ramli and Tahir, 2011; Russell et al., 2010; Tenório et al., 2012). Regrettably, the derivation of Z-R relations is not standard among studies and has included (i) different scatter regimes and (ii) types of regression. Assumptions regarding the rain drop shape (e.g., Rayleigh and Mie Scattering regimes) influence the calculation of Z and as a result, the derived Z-R relations (Dolan and Rutledge, 2009; Löffler-Mang et al., 1999; Maki et al., 2005; Tokay et al., 2001; Wexler and Atlas, 1963). On the other hand, Z-R derivation has been carried out not only by applying directly a non-linear regression but also performing a logarithmic transformation of Z and R, and a linear regression afterwards.

Despite some efforts for providing a common framework for radar QPE by means of specialized radar libraries (Heistermann et al., 2013), tailor-made solutions by using Z-R relationships and locally adapted sub-processes remain the main choice for researchers worldwide. This is understandable although not desirable for a straightforward application of radar QPE methods in other study sites. On the other hand, the suitability of machine learning (ML)-based models for radar QPE has been explored among the last years. This approach takes advantage of ML algorithms for transforming reflectivity values into rainfall without intermediate steps (e.g., rain gauge bias adjustment). This in turn reduces the complexity of the radar rainfall retrieval and provides a solution that neglects the need of rain gauge data after the model has been developed. Some studies have used Artificial Neural Networks (ANNs) for radar QPE by using the vertical profile of rain (Orlandini and Morlini, 2000; Teschl et al., 2007) and have proven to be efficient at short ranges. ANNs have been the most popular ML technique for radar QPE. However, ANNs have some drawbacks related to overfitting, slow convergence and complex tunning. As an alternative, attention has been drawn recently to ensemble methods such as Random Forest (RF). Here the structure of the RF model as a group of binary decision trees facilitates the interpretation of the model. Besides, its reduced number of hyper-parameters to tune simplifies the training process for non-experts. Likewise, RF models for radar QPE have used radar reflectivity of several heights (Yang et al., 2017; Yu et al., 2017).

1.2.2 Methods on Discharge Forecasting using radar rainfall

An accurate radar rainfall derivation has enormous implications on process-based hydrological applications where rainfall estimations can highly influence the results. Discharge forecasting is one

of the most important hydrological applications that strongly rely on accurate rainfall estimations and is greatly benefited from the high spatial resolution of radar imagery (Abon et al., 2016; Chen et al., 2016; He et al., 2013; Hsu et al., 2019; Keblouti et al., 2013). It has been proven that radar rainfall estimations remarkably improve discharge forecasting when rainfall events of high variability occur (He et al., 2013). Thereby, spatial rainfall estimations are a critical data source for distributed and semi-distributed hydrological models. Here, radar rainfall data has been used as the most reliable source because of their high resolution. Nonetheless, for obtaining accurate radar rainfall estimations, usually a dense and evenly distributed rain gauge network is needed for a proper radar bias adjustment. Moreover, these process-based models also need additional geomorphological variables monitored at ground which may be scarce or very difficult to obtain in mountain regions as the Andes. This is because the logistic access to (remote) mountain locations can be extremely difficult and expensive. In contrast, the use of a different approach by means of ML techniques has gradually increased during the last decades with satisfactory results (Dinu et al., 2017a,b; Falck et al., 2018; Ogale and Srivastava, 2019; Ragettli et al., 2017). Regardless of the approach for performing discharge forecasting, studies have continuously used radar rainfall estimates as input for the model which demand an exhaustive radar rainfall derivation process. Therefore, the success of the discharge forecasting models has remained closely related to the quality of radar rainfall estimates.

The use of ML-based models in water resources has mainly be limited to the straightforward application of ML algorithms. Even though promising results have been obtained, the modeling process has lacked deeper consideration of data representation (i.e., feature engineering) which may enhance the performance of ML-based models.

1.3 Study Area

The study area is located in the country of Ecuador that extends from the Pacific Ocean to the Amazon forest in South America. The Andean cordillera crosses the country from north to south and constitutes a meaningful topographic barrier that divides its territory (Morán-Tejeda et al., 2016). This investigation is focused on the southern Andes of Ecuador mainly in the province of Azuay. The study area is depicted in Figure 1.2

The area of investigation comprises the intersection between The Macizo del Cajas, which is part of the United Nations Educational, Scientific and Cultural Organization's (UNESCO's) World Network of Biosphere Reserves, and the Paute River Basin until the Zhurucay Ecohydrological

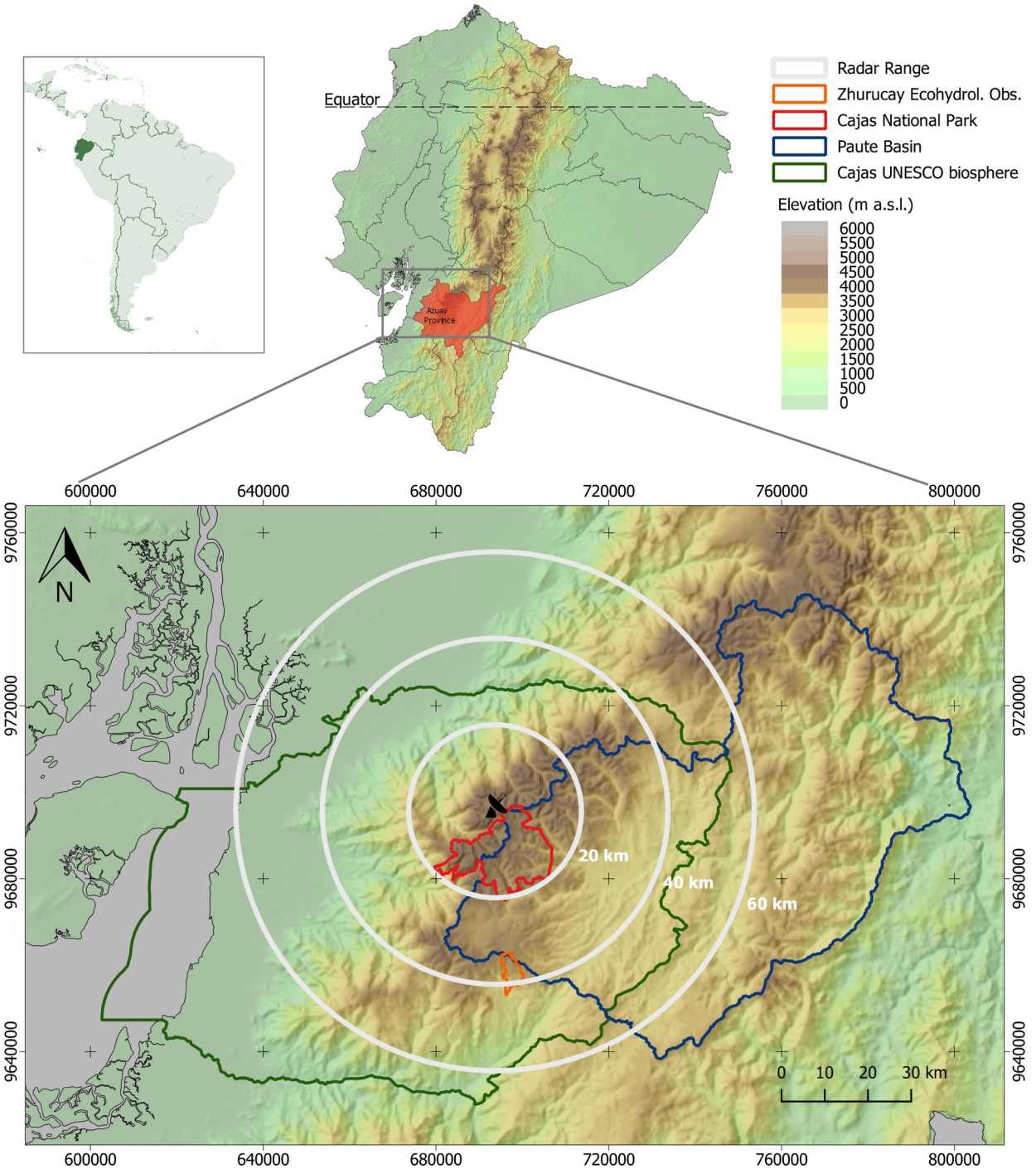

Figure 1.2: The study area of the thesis in the southern Andes of Ecuador is marked by the maximum range (60 km line) of the radar (black symbol) which is installed at the Paragüillas peak at 4450 m a.s.l.

Observatory (ZEO) in the south. The area is characterized by a very complex topography of the Andean Cordillera with elevations ranging from 1600 to 4680 m a.s.l. (Mora and Willems, 2012). The Andean slopes of the study area drain towards the east (i.e., Atlantic Ocean) and the west (i.e., the Pacific Ocean). Mountain catchments in the tropical Andes provide vital watershed services to downstream communities (Hamel et al., 2018), most importantly potable and irrigation water. Despite its high elevations, this tropical mountain region has no permanently snow and ice covered peaks (Célleri et al., 2007). One of the highest summits in the region is the Paragüillas peak (4450 m a.s.l.) at the northern border of the Cajas National Park. While a few others are slightly higher, they are not accessible for installing monitoring equipment. The area comprises altitudinal gradients that extend from the lowlands (inter-andean valleys) passing from a shrub-dominated (sub-páramo) transition zone from upland forest to páramo ecosystem. The latter is the main vegetation coverage of the head water catchments that substantially provide hydrological services to the cities in the lowlands (Célleri et al., 2007).

Precipitation studies have characterized the relevant section of the Paute River basin as a bimodal precipitation regime with the rainy months in the periods March-April and October-November (Campozano et al., 2016; Mora and Willems, 2012). Nonetheless, an area-wide high-resolution precipitation analysis is not yet available for the study area. However, an important contribution of about 10% more rainfall (i.e., drizzle) than previously quantify has been reported by Padrón et al. (2015) at the highest altitudes based on rain gauge and disdrometer comparison. It may account for almost 30% of rainfall in some areas such as the ZEO. Climate is influenced by the Pacific coastal regime from the west, and continental and tropical Amazon air masses from the east (Esquivel-Hernández et al., 2019; Vuille et al., 2000). As a consequence, prevalent convective and orographic cloud formation occurs (Bendix et al., 2005). Temperature in the Andean highlands has been estimated spatially in an Andean catchment within the study area (Córdova et al., 2015) and showed a mean temperature lapse rate of 6.88°C per 1000 m along an altitudinal gradient that ranges from 2600 to 4200 m a.s.l. In addition, Morán-Tejeda et al. (2016) reported a mean annual temperature of 12°C in the Ecuadorian Andean region as derived from data of meteorological stations located from ∼1900 to 3100 m a.s.l.

1.4 Aims and Hypotheses

Up to date, the spatial dynamics of rainfall at tropical high Andean regions have been explored by using satellite data and therefore at spatial resolutions that are not able to capture its high variability.

As a consequence, there is a lack of understanding on the complex rainfall formation processes and its influence on climatological phenomena that occur in this biodiversity hotspot of the world. The first aim of this work therefore is to deploy a ground radar network in the southern Andes of Ecuador that allows monitoring rainfall at higher spatio-temporal resolution which helps to increase the still deficient knowledge about tropical rainfall. On the basis of this implementation, a particular focus is given to the radar located at the highest altitude close to the city of Cuenca (Figure 1.2). The weather radar provides an indirect measurement of rainfall and thus it is necessary to perform a radar QPE process for converting the native variable – radar reflectivity (Z) – into rainfall rate (R). This leads to the second aim of this work: the determination of Z-R relations that particularly account to the specific microphysics of local rainfall types as basis for physically-based radar QPE. Due to the irregular topography in the radar range, the spatial variation of the Z-R relationship needs to be assessed in combination with the corresponding analysis of rainfall types dominant at different altitudes. The Z-R relations found are key for the general radar rainfall retrieval process. It comprises a custom tailored solution of several steps for reflectivity upscaling and correction before the application of the Z-R equation for QPE is possible. Usually an additional step of gauge-based bias adjustment is performed. This is the most common method used for radar QPE and will be referred as a step-wise approach. Nonetheless, as the radar network is expected to grow, a more generalized method that would also be able to improve the performance of the step-wise approach is desirable. This is particularly because rain gauge networks in Andean mountain regions are scarce and unevenly distributed and thus, are of limited use for bias adjustment. Here, the application of ML-based models for radar QPE opens a promising alternative. Therefore, the third aim of the thesis is the optimization of radar QPE by using a ML-based approach. Thus, a RF model that uses reflectivity derived features as inputs of the model is evaluated and compared with the physically-based QPE derived through the application of the step-wise approach that uses a locally derived Z-R relationship. The RF model would compensate for the problems of hardly available rain gauge data at near real-time and thus, a lacking bias correction of the traditional approach. The suitability of the new RF model for hydrological applications is then tested for discharge forecasting in an Andean catchment. Here, one of the main advantages of ML-based models is the possibility to convert several input features (e.g., rainfall, soil moisture) into a target variable (e.g., discharge) without an explicit definition of a physical equation. In this context it should be tested if even the native radar reflectivity might perform well as an input for a discharge forecasting model. This would allow using radar data without performing a previous transformation from reflectivity to rainfall rate which would be highly beneficial in terms of an easy-to-use exploiting of radar

information. In consequence, the fourth aim of this investigation is to evaluate the performance of a RF model for discharge forecasting by using different input data types: native radar data (i.e., corrected reflectivity values from unrealistic measurements) and derived radar rainfall estimates (radar QPE). This part of the thesis will have comprehensive implications for the operation of discharge forecasting models with radar data for early-warning systems, because of its pioneer use of native radar data in hydrology.

In summary, the following research deficits were identified:

(RD1) There is no spatio-temporal rainfall data at high resolution in the mountain Andean regions of southern Ecuador. Operational, ground based imaging C- or X-band radar data are lacking in the entire country, as in the most countries of South America.

(RD2) As a result, there is a lacking knowledge about spatio-temporal rainfall dynamics, which is adverse for rain-driven applications like flood forecast modeling. Regarding radar networks as a solution, also the knowledge on rain types, its microphysical characteristics for tropical mountain rainfall and related Z-R relationships is deficient.

(RD3) Commonly used rainfall retrieval by means of step-wise approach in X-band radars is tailor made and site specific (locally derived Z-R application) with a limited range (e.g., 30 Km) of usability. There is no evidence of ML approaches for radar QPE using SP X-band radar in complex terrain that could improve, optimize and unify the commonly used step-wise approach.

(RD4) Quality of gauge-adjusted radar rainfall data used for discharge forecasting applications highly depends on the rain gauge networks distribution at ground. However, there is no knowledge on the usability of raw radar imagery (i.e., native reflectivity variable) for flood forecasting models when sparse and unevenly distributed rain gauge networks are available, as it is mostly the case in the Andes.

In accordance with these research deficits, the aims of this thesis were formulated as follows:

(A1) Deployment of a radar network in tropical and remote mountains.

(A2) Derive rain-type Z-R relationships at different geographic positions and elevations within the radar coverage.

(A3) Apply the step-wise approach by using the locally derived Z-R relationships (from A2) and compare with a RF model (ML-based) to be used as a radar QPE method in a mountainous area with a scarce and unevenly distributed gauge network.

(A4) Compare the forecasting performance of RF models trained with either native radar data (i.e., corrected reflectivity values from unrealistic measurements) or gauge-adjusted radar rainfall estimates as an alternative to avoid the use of rain gauges for radar rainfall adjustment.

The following three hypotheses were developed in accordance with the research deficits and aims listed above:

H1. Rain-type Z-R relationships derived at different altitudes and geographic locations within the radar coverage highly differ from each other. Thus, a single Z-R relationship is not able to properly convert radar reflectivity into rainfall.

H2. Radar QPE by using a RF model outperforms the commonly used step-wise approach.

H3. Discharge forecasting application using a RF model including native radar data shows similar performance than using gauge-adjusted radar rainfall estimates as model input.

1.5 Thesis Structure

In order to test the aforementioned hypotheses, the following working packages were developed. An overview of the complete thesis structure is given in Figure 1.3

WP1. Setup and physically-based QPE.

WP2. Improvement of QPE: ML-based radar QPE.

WP3. Discharge forecasting using ML and X-band radar data.

WP1 was conceptualized to answer hypothesis H1. An X-band radar network was deployed in the southern Andes of Ecuador that complements the former LAWR located in the city of Loja (Chapter 2). Two twin Rainscanner SP X-band radars (SELEX GmbH) were installed in the provinces of Azuay and Loja (Section 2.3). Moreover, additional equipment (e.g., disdrometers and rain gauges) at ground aimed to provide ancillary variables for the further radar rainfall retrieval process (Section 2.3.1). A complete framework for data transmission was set for communication among the network

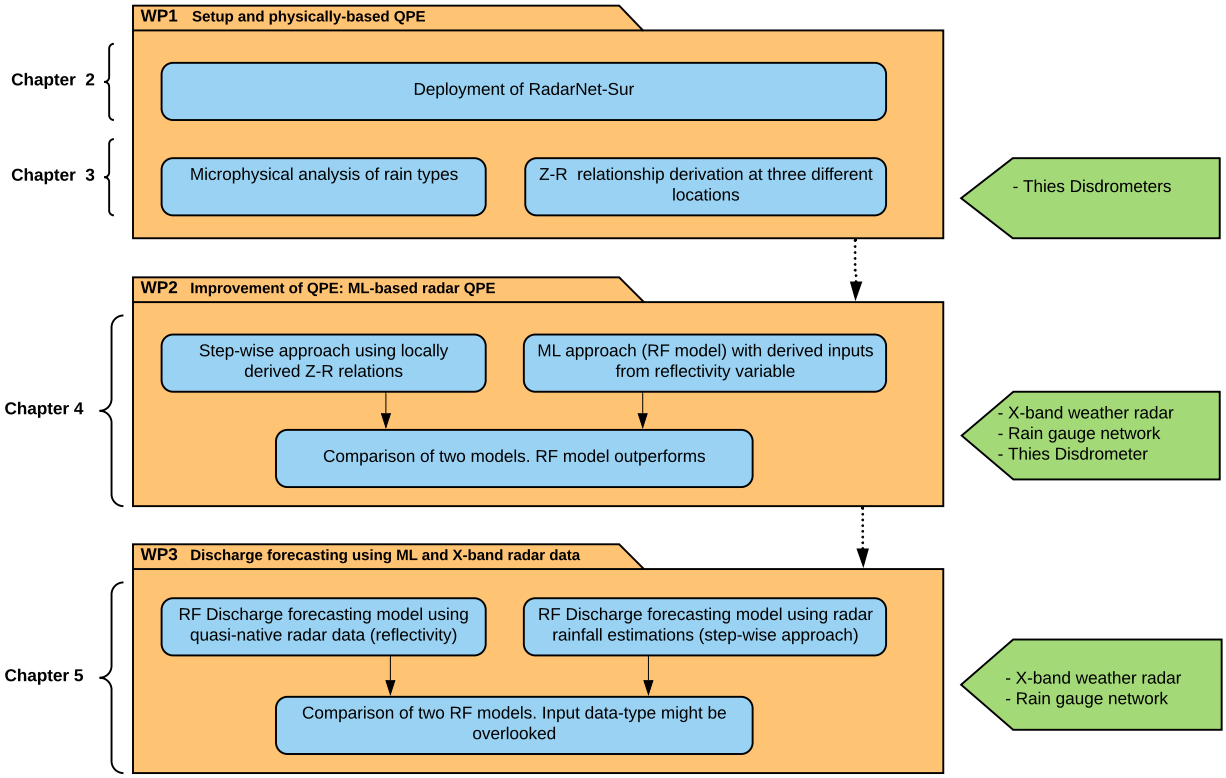

Figure 1.3: Thesis structure of working packages and related chapters.

(Section 2.3.2). The highest radar in the network (4450 m a.s.l.) was selected as study case for exploiting SP X-band radar data over the Andean region. First and foremost, the derivation of Z-R relationships that allow rainfall derivation from the native reflectivity variable was performed (Chapter 3). Although there were plenty of Z-R relations in literature worldwide, this was a necessary step because there was no evidence of Z-R relations in tropical mountain regions. Several rainfall events were detected (Section 3.2.3.3) by using three disdrometer series data recorded at different geographic locations and altitudes in the radar coverage. Thus a rainfall classification based on the mean volume diameter (Dm [mm]) was defined (Section 3.2.3.4). Z-R relations were derived by using both approaches, a linear regression on the log-transformed variables as well as a non-linear regression directly from the reflectivity and rainfall rate variables (Section 3.2.3.5). Z-R relations found towards this chapter would allow to implement the step-wise approach for rainfall retrieval process, by applying reflectivity upscaling and the corresponding transformation from reflectivity to rainfall rate by means of the adequate Z-R relationship. However, strong attenuation as a result of X-band technology, and mountain beam blockage would be of great difficulty to

overcome.

Thus, this brought out the next hypothesis H2 regarding the reliability and performance of the commonly used step-wise approach for radar QPE, where plenty of reflectivity corrections due to clutter, attenuation and beam blockage need to be performed before the application of the corresponding Z-R relationship. Therefore, WP2 was conceptualized to answer hypothesis H2. In WP2 the performance of two radar QPE models were evaluated and compared (Chapter 4). First, the step-wise approach was applied by using the Z-R relationship obtained in WP1. The resulting model incorporated reflectivity corrections regarding static and dynamic clutter, attenuation and finally the application of the locally derived Z-R relationship (Section 4.2.3.1). Likewise, the ML approach was implemented by using a RF model (Section 4.2.3.2). It used several reflectivity derived features from the radar imagery. The RF model satisfactorily mapped reflectivity related inputs into rainfall rate and also outperformed the step-wise approach. As a result of WP2, radar rainfall estimates could be used on different meteorological and hydrological applications.

The positive results obtained with the RF model brought out the possibility of using a comparable strategy when using radar data for discharge forecasting. One could use ML-based models to map discharge directly from the native reflectivity variable itself without the need of exhaustive radar rainfall derivation processes. This would reduce time processing and neglect the need of rain gauge data for applying radar rainfall bias adjustment. Thus WP3 was conceptualized to provide evidence to answer H3. Then, a comparison between RF models that used different input data type (e.g., derived radar rainfall estimates vs. native radar reflectivity) for discharge forecasting was performed (Chapter 5). Input data for the relevant models were obtained from native reflectivity records (i.e., reflectivity corrected from unrealistic measurements) as well as derived radar-rainfall data and as a result different input data-type time series (i.e., reflectivity vs. radar rainfall) were drawn (Section 5.2.3.2). Afterwards, several model configurations that differ on the data-type input and the inclusion or exclusion of a soil moisture proxy were defined (Section 5.2.3.3). Finally, the usefulness of native radar reflectivity for discharge forecasting was proven. It turned out that the models, despite its data-type input, performed alike.

The thesis consists of four manuscripts which comprise core Chapters 2,3,4 and 5, and constitute significant contributions to the deployment, management and exploitation of X-band radar imagery over high tropical mountains. All of them are currently published under peer-review journals. Chapter 6 summarizes the main research findings providing answers to the hypotheses proposed in section 1.4. Finally a brief outlook for further investigation is presented.

References

Abon, C. C., Kneis, D., Crisologo, I., Bronstert, A., Constantino David, C. P., and Heistermann, M. (2016). Evaluating the potential of radar-based rainfall estimates for streamflow and flood simulations in the Philippines. *Geomatics, Natural Hazards and Risk*, 7(4):1390–1405.

Allegretti, M. (2012). X-Band Mini Radar for Observing and Monitoring Rainfall Events. *Atmospheric and Climate Sciences*, 02(03):290–297.

Anagnostou, M. N., Nikolopoulos, E. I., Kalogiros, J., Anagnostou, E. N., Marra, F., Mair, E., Bertoldi, G., Tappeiner, U., and Borga, M. (2018). Advancing precipitation estimation and streamflow simulations in complex terrain with X-Band dual-polarization radar observations. *Remote Sensing*, 10(8).

Antonini, A., Melani, S., Corongiu, M., Romanelli, S., Mazza, A., Ortolani, A., and Gozzini, B. (2017). On the Implementation of a regional X-bandweather radar network. *Atmosphere*, 8(2):1–20.

Atlas, D., Ulbrich, C. W., Marks Jr., F. D., Black, R. A., Amitai, E., Willis, P. T., and Samsury, C. E. (2000). Partitioning tropical oceanic convective and stratiform rains by draft strength. *Journal of Geophysical Research: Atmospheres*, 105(D2):2259–2267.

Avolio, E., Cavalcanti, O., Furnari, L., Senatore, A., and Mendicino, G. (2019). Brief communication: Preliminary hydro-meteorological analysis of the flash flood of 20 August 2018 in Raganello Gorge, southern Italy. *Natural Hazards and Earth System Sciences*, 19(8):1619–1627.

Ballari, D., Castro, E., and Campozano, L. (2016). Validation of satellite precipitation (TRMM 3B43) in Ecuadorian coastal plains, Andean highlands and Amazonian rainforest. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 41(June):305–311.

Ballari, D., Giraldo, R., Campozano, L., and Samaniego, E. (2018). Spatial functional data analysis for regionalizing precipitation seasonality and intensity in a sparsely monitored region: Unveiling the spatio-temporal dependencies of precipitation in Ecuador. *International Journal of Climatology*, 38(8):3337–3354.

Bamba, B., Ochou, A. D., Zahiri, E.-P., and Kacou, M. (2014). Consistency in Z-R Relationship Variability Regardless Precipitating Systems, Climatic Zones Observed from Two Types of Disdrometer. *Atmospheric and Climate Sciences*, 04:941–955.

Bendix, J. (2000). Precipitation dynamics in Ecuador and northern Peru during the 1991/92 El Niño: A remote sensing perspective. *International Journal of Remote Sensing*, 21(3):533–548.

Bendix, J., Rollenbeck, R., Göttlicher, D., and Cermak, J. (2005). Cloud occurrence and cloud properties in Ecuador. *Climate Research*, 30:133–147.

Bendix, J., Rollenbeck, R., and Reudenbach, C. (2006). Diurnal patterns of rainfall in a tropical Andean valley of southern Ecuador as seen by a vertically pointing K-band Doppler radar. *International Journal of Climatology*, 26(6):829–846.

Bendix, J., Stefan, G., Reudenbach, C., and Bendix, A. (2003). A Case Study on Rainfall Dynamics during el Niño/La Niña 1997/99 in Ecuador and surroundings areas as inferred from GOES-8 and TRMM-PR Observations. *Erdkunde*, 57(2):82–93.

Berne, A., Uijlenhoet, R., Berne, A., and Quantitative, R. U. (2006). Quantitative analysis of X-band weather radar attenuation correction accuracy. *Natural Hazards and Earth System Sciences*, 6:419–425.

Borga, M., Comiti, F., Ruin, I., and Marra, F. (2019). Forensic analysis of flash flood response. *Wiley Interdisciplinary Reviews: Water*, 6(2):e1338.

Bronstert, A., Agarwal, A., Boessenkool, B., Crisologo, I., Fischer, M., Heistermann, M., Köhn-Reich, L., López-Tarazón, J. A., Moran, T., Ozturk, U., Reinhardt-Imjela, C., and Wendi, D. (2018). Forensic hydro-meteorological analysis of an extreme flash flood: The 2016-05-29 event in Braunsbach, SW Germany. *Science of the Total Environment*, 630:977–991.

Buytaert, W., Célleri, R., Willems, P., Bièvre, B. D., and Wyseure, G. (2006). Spatial and temporal rainfall variability in mountainous areas: A case study from the south Ecuadorian Andes. *Journal of Hydrology*, 329:413–421.

Büyükbaş, E. (2009). Assess the current and potential capabilities of weather radars for the use in wmo integrated global observing system (wigos). In *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.

Campozano, L., Célleri, R., Trachte, K., Bendix, J., and Samaniego, E. (2016). Rainfall and Cloud Dynamics in the Andes: A Southern Ecuador Case Study. *Advances in Meteorology*, 2016.

Caracciolo, C., Porcù, F., and Prodi, F. (2008). Precipitation classification at mid-latitudes in terms of drop size distribution parameters. *Advance Geosciences*, 16:11–17.

Célleri, R., Willems, P., Buytaert, W., and Feyen, J. (2007). Space-time rainfall variability in the Paute basin, Ecuadorian Andes. *Hydrological Processes*, 21(24):3316–3327.

Cerro, C., Codina, B., Bech, J., and Lorente, J. (1997). Modeling Raindrop Size Distribution and Z(R) Relations in the Western Mediterranean Area. *Journal of Applied Meteorology*, 36:1470–1479.

Chen, B., Wang, J., and Gong, D. (2016). Raindrop size distribution in a midlatitude continental squall line measured by thies optical disdrometers over east China. *Journal of Applied Meteorology and Climatology*, 55(3):621–634.

Contreras, J., Ballari, D., de Bruin, S., and Samaniego, E. (2019). Rainfall monitoring network design using conditioned Latin hypercube sampling and satellite precipitation estimates: An application in the ungauged Ecuadorian Amazon. *International Journal of Climatology*, 39(4):2209–2226.

Córdova, M., Carrillo-Rojas, G., Crespo, P., Wilcox, B., and Célleri, R. (2015). Evaluation of the Penman-Monteith (FAO 56 PM) Method for Calculating Reference Evapotranspiration Using Limited Data. *Mountain Research and Development*, 35:230–239.

Dinu, C., Drobot, R., Pricop, C., and Blidaru, T. V. (2017a). Flash-Flood Modelling with Artificial Neural Networks using Radar Rainfall Estimates. *Mathematical Modelling in Civil Engineering*, 13(3):10–20.

Dinu, C., Drobot, R., Pricop, C., and Blidaru, T. V. (2017b). Genetic Programming Technique applied for Flaash-Flood Modelling using Radar Rainfall Estimates. *Mathematical Modelling in Civil Engineering*, 13(4):27–38.

Dolan, B. and Rutledge, S. A. (2009). A theory-based hydrometeor identification algorithm for X-band polarimetric radars. *Journal of Atmospheric and Oceanic Technology*, 26(10):2071–2088.

Emmanuel, I., Andrieu, H., Leblois, E., Janey, N., and Payrastre, O. (2015). Influence of rainfall spatial variability on rainfall-runoff modelling : Benefit of a simulation approach ? *Journal of Hydrology*.

Espinoza Villar, J. C., Ronchail, J., Guyot, J. L., Cochonneau, G., Naziano, F., Lavado, W., De Oliveira, E., Pombosa, R., and Vauchel, P. (2009). Spatio-temporal rainfall variability in the Amazon basin countries (Brazil, Peru, Bolivia, Colombia, and Ecuador). *International Journal of Climatology*, 29:1574–1594.

Esquivel-Hernández, G., Mosquera, G. M., Sánchez-Murillo, R., Quesada-Román, A., Birkel, C., Crespo, P., Célleri, R., Windhorst, D., Breuer, L., and Boll, J. (2019). Moisture transport and seasonal variations in the stable isotopic composition of rainfall in central american and andean páramo during el niño conditions (2015-2016). *Hydrological Processes*, 33(13):1802–1817.

Falck, A. S., Maggioni, V., Tomasella, J., Diniz, F. L., Mei, Y., Beneti, C. A., Herdies, D. L., Neundorf, R., Caram, R. O., and Rodriguez, D. A. (2018). Improving the use of ground-based radar rainfall data for monitoring and predicting floods in the Iguaçu river basin. *Journal of Hydrology*, 567(October):626–636.

Frasier, S. J., Kabeche, F., Figueras i Ventura, J., Al-Sakka, H., Tabary, P., Beck, J., and Bousquet, O. (2013). In-Place Estimation of Wet Radome Attenuation at X Band. *Journal of Atmospheric and Oceanic Technology*, 30:917–928.

Fries, A., Rollenbeck, R., Bayer, F., Gonzalez, V., Oñate-Valivieso, F., Peters, T., and Bendix, J. (2014). Catchment precipitation processes in the San Francisco valley in southern Ecuador: combined approach using high-resolution radar images and in situ observations. *Meteorology and Atmospheric Physics*, 126(1-2):13–29.

Germann, U., Galli, G., Boscacci, M., and Bolliger, M. (2006). Radar precipitation measurement in a mountainous region. *Quarterly Journal of the Royal Meteorological Society*, 132:1669–1692.

Goormans, T. (2011). *Analysis of local weather radar data in support of sewer system modelling*. PhD thesis, Katholieke Univesiteit Leuven.

Hamel, P., Riveros-Iregui, D., Ballari, D., Browning, T., Célleri, R., Chandler, D., Chun, K. P., Destouni, G., Jacobs, S., Jasechko, S., Johnson, M., Krishnaswamy, J., Poca, M., Pompeu, P. V., and Rocha, H. (2018). Watershed services in the humid tropics: Opportunities from recent advances in ecohydrology. *Ecohydrology*, 11(3).

Harrison, D. L., Driscoll, S. J., and Kitchen, M. (2000). Improving precipitation estimates from weather radar using quality control and correction techniques. *Meteorological Applications*, 6:135–144.

He, X., Sonnenborg, T. O., Refsgaard, J. C., Vejen, F., and Jensen, K. H. (2013). Evaluation of the value of radar QPE data and rain gauge data for hydrological modeling. *Water Resources Research*, 49(September):5989–6005.

Heistermann, M., Jacobi, S., and Pfaff, T. (2013). Technical Note: An open source library for processing weather radar data (wradlib). *Hydrology and Earth System Sciences*, 17(2):863–871.

Hsu, S. Y., Chen, T. B., Du, W. C., Wu, J. H., and Chen, S. C. (2019). IntegrateWeather radar and monitoring devices for urban flooding surveillance. *Sensors (Switzerland)*, 19(4):1–15.

Jensen, N. E. (2002). X-band local area weather radar - preliminary calibration results. *Water Science & Technology*, 45:135–138.

Keblouti, M., Ouerdachi, L., and Berhail, S. (2013). The use of weather radar for rainfall-runoff modeling , case of Seybouse watershed (Algeria). *Arabian Journal of Geosciences*.

Kirstetter, P.-E., Gourley, J. J., Hong, Y., Zhang, J., Moazamigoodarzi, S., Langston, C., and Arthur, A. (2015). Probabilistic precipitation rate estimates with ground-based radar networks. *Water Resources Research*, 51:1422–1442.

Kumar, L. S., Lee, Y. H., Yeo, J. X., and Ong, J. T. (2011). Tropical Rain Classification and Estimation of Rain From Z-R (Reflectivity-Rain Rate) Relationships. *Progress In Electromagnetics Research*, 32:107–127.

Lengfeld, K., Clemens, M., Münster, H., and Ament, F. (2014). Performance of high-resolution X-band weather radar networks — the PATTERN example. *Atmospheric Measurement Techniques*, 7(12):4151–4166.

Lo Conti, F., Francipane, A., Pumo, D., and Noto, L. V. (2015). Exploring single polarization X-band weather radar potentials for local meteorological and hydrological applications. *Journal of Hydrology*, 531:508–522.

Lobligeois, F., Andréassian, V., Perrin, C., Tabary, P., and Loumagne, C. (2014). When does higher spatial resolution rainfall information improve streamflow simulation? An evaluation using 3620 flood events. *Hydrology and Earth System Sciences*, 18(2):575–594.

Löffler-Mang, M., Kunz, M., and Schmid, W. (1999). On the performance of a low-cost K-band Doppler radar for quantitative rain measurements. *Journal of Atmospheric and Oceanic Technology*, 16(2-3):379–387.

Maki, M., Park, S.-G., and Bringi, V. (2005). Effect of Natural Variations in Rain Drop Size Distributions on Rain Rate Estimators of 3 cm Wavelength Polarimetric Radar. *Journal of the Meteorological Society of Japan*, 83(5):871–893.

Manola, I., Steeneveld, G. J., Uijlenhoet, R., and Holtslag, A. A. (2020). Analysis of urban rainfall from hourly to seasonal scales using high-resolution radar observations in the Netherlands. *International Journal of Climatology*, 40(2):822–840.

Marshall, J. S. and Palmer, W. M. K. (1948). The Distribution of Raindrops With Size. *Journal of Meteorology*, 5:165–166.

McKee, J. L. and Binns, A. D. (2016). A review of gauge-radar merging methods for quantitative precipitation estimation in hydrology. *Canadian Water Resources Journal*, 41(1-2):186–203.

McLaughlin, D., Pepyne, D., Chandrasekar, V., Philips, B., Kurose, J., Zink, M., Droegemeier, K., Cruz-Pol, S., Junyent, F., Brotzge, J., Westbrook, D., Bharadwaj, N., Wang, Y., Lyons, E., Hondl, K., Liu, Y., Knapp, E., Xue, M., Hopf, A., Kloesel, K., DeFonzo, A., Kollias, P., Brewster, K., Contreras, R., Dolan, B., Djaferis, T., Insanic, E., Fraiser, S., and Carr, F. (2009). Short-wavelength technology and the potential for distributed networks of small radar systems. *Bulletin of the American Meteorological Society*, 90(12):1797–1817.

McRoberts, D. B. and Nielsen-Gammon, J. W. (2017). Detecting Beam Blockage in Radar-Based Precipitation Estimates. *Journal of Atmospheric and Oceanic Technology*, 34:1407–1422.

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., Tolstoy, L., Kucera, P. A., Petersen, W. A., Grazioli, J., and Pazmany, A. L. (2016). Deployment and Performance Analyses of High-Resolution Iowa XPOL Radar System during the NASA IFloodS Campaign. *Journal of Hydrometeorology*, 17(Febraury):455–479.

Mora, D. E. and Willems, P. (2012). Decadal oscillations in rainfall and air temperature in the Paute River Basin-Southern Andes of Ecuador. *Theoretical and Applied Climatology*, 108(1-2):267–282.

Morán-Tejeda, E., Bazo, J., López-Moreno, J. I., Aguilar, E., Azorín-Molina, C., Sanchez-Lorenzo, A., Martínez, R., Nieto, J. J., Mejía, R., Martín-Hernández, N., and Vicente-Serrano, S. M. (2016). Climate trends and variability in Ecuador (1966-2011). *International Journal of Climatology*, 36(11):3839–3855.

Morin, E. and Gabella, M. (2007). Radar-based quantitative precipitation estimation over Mediterranean and dry climate regimes. *Journal of Geophysical Research*, 112(October):1–13.

Narayana Rao, T., Narayana Rao, D., and Mohan, K. (2001). Classification of tropical precipitating systems and associated Z-R relationships. *Journal of Geophysical Research*, 106:17669–17711.

Nzeukou, A. and Sauvageot, H. (2004). Raindrop size distribution and radar parameters at Cape Verde. *Journal of Applied Meteorology*, 43:90–105.

Ochoa-Rodriguez, S., Wang, L. P., Willems, P., and Onof, C. (2019). A Review of Radar-Rain Gauge Data Merging Methods and Their Potential for Urban Hydrological Applications. *Water Resources Research*, 55(8):6356–6391.

Ochou, A. D., Zahiri, E.-P., Bamba, B., and Koffi, M. (2011). Understanding the Variability of Z-R Relationships Caused by Natural Variations in Raindrop Size Distributions (DSD): Implication of Drop Size and Number. *Atmospheric and Climate Sciences*, 01:147–164.

Ogale, S. and Srivastava, S. (2019). Modelling and short term forecasting of flash floods in an urban environment. *25th National Conference on Communications, NCC 2019*, pages 1–6.

Oñate-Valdivieso, F., Fries, A., Mendoza, K., Gonzalez-Jaramillo, V., Pucha-Cofrep, F., Rollenbeck, R., and Bendix, J. (2018). Temporal and spatial analysis of precipitation patterns in an Andean region of southern Ecuador using LAWR weather radar. *Meteorology and Atmospheric Physics*, 130(4):473–484.

Orlandini, S. and Morlini, I. (2000). Artificial neural networks estimation of rainfall intensity from radar observations. *Journal of Geophysical Research*, 105(D20):849–861.

Padrón, R. S., Wilcox, B. P., Crespo, P., and Célleri, R. (2015). Rainfall in the Andean Páramo: New Insights from High-Resolution Monitoring in Southern Ecuador. *Journal of Hydrometeorology*, 16:985–996.

Poveda, G., Mesa, O. J., Salazar, L. F., Arias, P. a., Moreno, H. a., Vieira, S. C., Agudelo, P. a., Toro, V. G., and Alvarez, J. F. (2005). The Diurnal Cycle of Precipitation in the Tropical Andes of Colombia. *Monthly Weather Review*, 133:228–240.

Pulwarty, R. S., Barry, R. G., Hurst, C. M., Sellinger, K., and Mogollon, L. F. (1998). Precipitation in the Venezuelan Andes in the context of regional climate. *Meteorology and Atmospheric Physics*, 67:217–237.

Ragettli, S., Zhou, J., Wang, H., Liu, C., and Guo, L. (2017). Modeling flash floods in ungauged mountain catchments of China : A decision tree learning approach for parameter regionalization. *Journal of Hydrology*, 555:330–346.

Ramli, S. and Tahir, W. (2011). Radar Hydrology: New Z/R Relationships for Quantitative Precipitation Estimation in Klang River Basin, Malaysia. *International Journal of Environmental Science and Development*, 2:223–227.

Rollenbeck, R. and Bendix, J. (2006). Experimental calibration of a cost-effective X-band weather radar for climate ecological studies in southern Ecuador. *Atmospheric Research*, 79(3-4):296–316.

Rollenbeck, R. and Bendix, J. (2011). Rainfall distribution in the Andes of southern Ecuador derived from blending weather radar data and meteorological field observations. *Atmospheric Research*, 99:277–289.

Rosenfeld, D., Wolff, D. B., and Atlas, D. (1993). General Probability-matched Relations between Radar Reflectivity and Rain Rate. *Journal of Applied Meteorology*, 32(1):50–72.

Rossel, F. and Cadier, E. (2009). El Niño and prediction of anomalous monthly rainfalls in Ecuador. *Hydrological Processes*, 23(22):3253–3260.

Russell, B., Williams, E. R., Gosset, M., Cazenave, F., Descroix, L., Guy, N., Lebel, T., Ali, a., Metayer, F., and Quantin, G. (2010). Radar/rain-gauge comparisons on squall lines in Niamey, Niger for the AMMA. *Quarterly Journal of the Royal Meteorological Society*, 136:289–303.

Savina, M. (2011). *The Use of a Cost-Effective X-band Weather Radar in Alpine Region*. PhD thesis, Swiss Federal Institute of Technology Zurich.

Scheel, M. L. M., Rohrer, M., Huggel, C., Santos Villar, D., Silvestre, E., and Huffman, G. J. (2011). Evaluation of TRMM Multi-satellite Precipitation Analysis (TMPA) performance in the Central Andes region and its dependency on spatial and temporal resolution. *Hydrology and Earth System Sciences*, 15:2649–2663.

Sen Jaiswal, R., Uma, S., and Santhakumaran, A. (2009). Study of Z-R relationship over Gadanki for different rainfall rates. *Indian Journal of Radio and Space Physics*, 38:159–164.

Sucozhañay, A. and Célleri, R. (2018). Impact of Rain Gauges distribution on the runoff simulation of a small mountain catchment in Southern Ecuador. *Water (Switzerland)*, 10(9).

Tenório, S., da Silva Moraes, M. C., and Sauvageot, H. (2012). Raindrop Size Distribution and Radar Parameters in Coastal Tropical Rain Systems of Northeastern Brazil. *Journal of Applied Meteorology and Climatology*, 51:1960–1970.

Teschl, R., Randeu, W. L., and Teschl, F. (2007). Improving weather radar estimates of rainfall using feed-forward neural networks. *Neural Networks*, 20:519–527.

Tokay, A., Hartmann, P., Battaglia, A., Gage, K. S., Clark, W. L., and Williams, C. R. (2009). A field study of reflectivity and Z-R relations using vertically pointing radars and disdrometers. *Journal of Atmospheric and Oceanic Technology*, 26:1120–1134.

Tokay, A., Kruger, A., and Krajewski, W. F. (2001). Comparison of Drop Size Distribution Measurements by Impact and Optical Disdrometers. *Journal of Applied Meteorology*, 40(11):2083–2097.

Uijlenhoet, R. (2001). Raindrop size distributions and radar reflectivity-rain rate relationships for radar hydrology. *Hydrology and Earth System Sciences*, 5(4):615–628.

Uijlenhoet, R., Steiner, M., and Smith, J. A. (2003). Variability of Raindrop Size Distributions in a Squall Line and Implications for Radar Rainfall Estimation. *Journal of Hydrometeorology*, 4:43–61.

Villarini, G. and Krajewski, W. F. (2010). Review of the different sources of uncertainty in single polarization radar-based estimates of rainfall. *Surveys in Geophysics*, 31(1):107–129.

Vuille, M., Bradley, R. S., and Keimig, F. (2000). Climate Variability in the Andes of Ecuador and Its Relation to Tropical Pacific and Atlantic Sea Surface Temperature Anomalies. *Journal of Climate*, 13(14):2520–2535.

Wexler, R. and Atlas, D. (1963). Radar Reflectivity and Attenuation of Rain. *Journal of Applied Meteorology*, 2(2):276–280.

Wohl, E., Barros, A., Brunsell, N., Chappell, N. A., Coe, M., Giambelluca, T., Goldsmith, S., Harmon, R., Hendrickx, J. M. H., Juvik, J., McDonnell, J., and Ogden, F. (2012). The hydrology of the humid tropics. *Nature Climate Change*, 2:655–662.

Xie, H., Zhang, X., Yu, B., and Sharif, H. (2011). Performance evaluation of interpolation methods for incorporating rain gauge measurements into NEXRAD precipitation data: a case study in the Upper Guadalupe River Basin. *Hydrological Processes*, 25(24):3711–3720.

Yang, X., Kuang, Q., Zhang, W., and Zhang, G. (2017). A terrain-based weighted random forests method for radar quantitative. *Meteorological Applications*, 414(May):404–414.

Yu, P.-s., Yang, T.-c., Chen, S.-y., Kuo, C.-m., and Tseng, H.-w. (2017). Comparison of random forests and support vector machine for real-time radar-derived rainfall forecasting. *Journal of Hydrology*, 552:92–104.

2 RadarNet-Sur: first weather radar network in tropical high mountains

This chapter is published in *Bulletin of the American Meteorological Society*, 98, 1235-1254, 2017.
Accepted: 23 October 2016
First published online: 23 June 2017
DOI: https://doi.org/10.1175/BAMS-D-15-00178.1

RadarNet-Sur: first weather radar network in tropical high mountains

Jörg Bendix[1], Andreas Fries[2], Jorge Zárate[3], Katja Trachte[1], Rütger Rollenbeck[1],Franz Pucha-Cofrep[4], Renzo Paladines[10], Ivan Palacios[5], Johanna Orellana[8], Fernando Oñate-Valdivieso[4], Carlos Naranjo[6], Leonardo Mendoza[3], Diego Mejia[7], Mario Guallpa[5], Francisco Gordillo[9], Victor Gonzalez-Jaramillo[4], Maik Dobbermann[1], Rolando Célleri[8], Carlos Carrillo[10], Augusto Araque[7], Sebastian Achilles[1].

[1] Laboratory for Climatology and Remote Sensing, Faculty of Geography, University of Marburg, Marburg, Germany

[2] Laboratory for Climatology and Remote Sensing, Faculty of Geography, University of Marburg, Marburg, Germany, and Departamento de Geologiía y Minas e Ingeniería Civil, Hidrología y Climatología Working Group, Universidad Técnica Particular de Loja, Loja, Ecuador

[3] Departamento de Planificación y Desarrollo Territorial, Gobierno Provincial de Loja, Loja, Ecuador

[4] Departamento de Geología y Minas e Ingeniería Civil, Hidrología y Climatología Working Group, Universidad Técnica Particular de Loja, Loja, Ecuador

[5] Departamento de Gestión Ambiental, Empresa de Telecomunicaciones, Agua Potable y

Alcantarillado de Cuenca, Cuenca, Ecuador
[6] Instituto Nacional de Meteorología e Hidrología, Quito, Ecuador
[7] Instituto Nacional de Meteorología e Hidrología, Loja, Ecuador
[8] Departamento de Recursos Hídricos y Ciencias Ambientales, Facultad de Ingeniería, Facultad de Ciencias Químicas, Universidad de Cuenca, Cuenca, Ecuador
[9] Fondo Regional del Agua, Loja, Ecuador
[10] Naturaleza y Cultura International, Loja, Ecuador

Abstract: Weather radar networks are indispensable tools for forecasting and disaster prevention in industrialized countries. However, they are far less common in the countries of South America, which frequently suffer from an underdeveloped network of meteorological stations. To address this problem in southern Ecuador, this article presents a novel radar network using cost-effective, single-polarization, X-band technology: the RadarNet-Sur. The RadarNet-Sur network is based on three scanning X-band weather radar units that cover approximately 87000 km^2 of southern Ecuador. Several instruments, including five optical disdrometers and two vertically aligned K-band Doppler radar profilers, are used to properly (inter) calibrate the radars. Radar signal processing is a major issue in the high mountains of Ecuador because cost-effective radar technologies typically lack Doppler capabilities. Thus, special procedures were developed for clutter detection and beam blockage correction by integrating ground-based and satelliteborne measurements. To demonstrate practical applications, a map of areas frequently affected by intense rainfall is presented, based on a time series of one radar that has been in operation since 2002. Such information is of vital importance to, for example, infrastructure management because rain-driven landslides are a major issue for road maintenance and safety throughout Ecuador. The presented case study of exceptionally strong rain events during the recent El Niño in March 2015 highlights the system's practicality in weather forecasting related to disaster management. For the first time, RadarNet-Sur warrants a spatial-explicit observation of El Niño-related heavy precipitation in a transect from the coast to the highlands in a spatial resolution of 500 m.

2.1 Introduction

Ecuador is one of the countries in tropical western South America significantly affected by the El Niño-Southern Oscillation (ENSO) phenomenon (Bendix, 2000; Bendix et al., 2003; Rossel and Cadier, 2009; Vuille et al., 2000). Precipitation hazards such as flooding and rainfall-driven landslides during El Niño (EN) events regularly result in severe agricultural losses and infrastructure damage, particularly on the coast and the western slopes of the Andes (e.g., Vos et al., 1999). Here, ENSO rainfall extremes clearly exceed the natural variability of precipitation in normal, non-ENSO years (Bendix, 2004). Table 2.1 shows the complex spatiotemporal nature of heavy ENSO rainfall as measured by the sparse network of long-term operational meteorological stations in southern Ecuador since the Super El Niño of 1982/83 (Figure 2.1). The analysis of rainfall anomalies reveals a general strong positive deviation at the coastal plains in the west of southern Ecuador, while the Andean highlands frequently experience below-average precipitation. Obviously, the positive anomalies at the coast (stations Guayaquil and Milagro) were exceptionally enhanced during the super events in 1983 and 1998 (see also Bendix et al., 2011). However, a high spatial variability of this general tendency is indicated by the fact that the zone of high positive rainfall anomalies is extended eastward into the Andes until the station Celica during EN1983, concomitant with strong westerly wind anomalies (Bendix, 1999). This extension toward the eastern parts of southern Ecuador was not observed at the station Celica during EN1998. The station Cariamanga presents the reverse of strong positive rainfall anomalies during EN1998 but not during EN1983, where orographic effects on rainfall seemed to play an important role in establishing a second maximum at the south Ecuadorian Andean mountain slopes (Bendix et al., 2003). In contrast to previous super events, the strong EN in 2015/16 showed slightly negative anomalies at the coastal stations (e.g., station Milagro), while strong anomalies occurred in the southernmost parts of Ecuador (station Santa Rosa). Noteworthy is that moderate past EN events (1987, 1992/93) revealed very different spatial patterns of positive and negative anomalies in rainfall. The data show that the location of heavy rain during an event depends on the complex interplay of several factors. These include local variations in sea surface temperature (SST) anomalies, mainly in the Niño-1+2 region, a pattern of anomalies in the larger-scale atmospheric circulation, and modifications of the land-sea-breeze system with precipitation effects up to 100 km inland (Bendix and Bendix, 2006; Bendix, 2000).

Compared to EN events, the spatial distribution of heavy rainfall is generally inverted during the La Niña (LN) or Pacific cold phases (Table 2.1). In southern Ecuador, for instance, LN1999 presented strong negative rainfall anomalies at the coastal plains (stations Guayaquil and Milagro),

Table 2.1: Relative anomalies (Δ%) of rainfall during ENSO (El Niño and La Niña) events (1983-2015) from long-term average (1981-2010, column 5) for the main ENSO period of the year (Jan-May). Data are taken from the operational meteorological stations of the Ecuadorian weather service (INAMHI) with long-term availability (see Figure 2.1). Station location: CO = coastal plains (Pacific) and A = Andes mountains. Boldfaced and underlined values indicate anomalies higher than 100%, Boldfaced values anomalies are higher than 50%.

INAMHI station	Zone	Radar	Elevation (m a.s.l.)	Rain (mm)	El Niño						La Niña				
					1983	**1987**	**1992**	**1993**	**1998**	**2015**	**1989**	**1999**	**2008**	**2011**	**2012**
					Δ%	Δ%	Δ%	Δ%	Δ%	Δ%	Δ%	Δ%	Δ%	Δ%	Δ%
Guayaquil	CO	CAXX	5	1057	**199**	**90**	39	29	**208**	-14	-13	63	NA	NA	33
Milagro	CO	CAXX	13	1313	**109**	**68**	**78**	24	**205**	-11	19	-31	**58**	-35	**57**
Paute	A	CAXX	2289	364	28	NA	-29	30	-7	-60	17	38	33	49	16
Cañar	A	CAXX	3038	271	14	5	-32	40	18	-11	5	**57**			
Gualaceo	A	CAXX	2449	377	17	21	-32	7	15	-14	23	**63**	**69**	**55**	38
Santa Rosa	CO	GUAXX	13	283	**65**	NA	NA	NA	NA	**134**	NA	NA	NA	NA	**204**
Celica	A	GUAXX	2700	965	**139**	**-50**	25	**58**	4	33	**149**	-20	**68**	-25	**75**
Saraguro	A	All	2525	457	9	-20	-22	**79**	**1**	-10	**56**	**57**	23	-16	-3
Loja	A	LOXX	2160	520	23	-15	-12	20	-20	25	22	46	47	22	15
Cariamanga	A	LOXX	1960	1084	47	-30	-37	38	**80**	21	36	46	**51**	5	1

while the higher eastern parts of the Andes (stations Saraguro, Loja, Cariamanga, Gualaceo, and Cañar) widely experienced heavy rains. Yet again, the station Celica is an outlier for the lack of consistency in its anomalies. Negative rainfall anomalies occurred during LN1999, while hazardous precipitation was observed during LN1989. The complexity of the situation has grown since the year 2000 because the cold phase of the Pacific decadal oscillation (PDO) has increased the frequency of Modoki events (Bendix et al., 2003). In 2008 and 2012, LN Modokis, a combination of negative and positive SST anomalies in the central Pacific (Niño-3.4 region) and the eastern Pacific (Niño-1+2 region), revealed a specific spatial distribution of rainfall anomalies [for further information on Modoki events, readers may refer to Kao and Yu (2009) as well as to Capotondi et al. (2015)]. Positive rainfall anomalies were observed at the coast (e.g., station Milagro) due to high SSTs in the Niño-1+2 region as well as in different areas of the Andean highlands (stations Paute, Gualaceo, and Celica) triggered by negative SST anomalies in the central Pacific. Thus, LN Modoki events caused different, unexpected changes at individual stations (e.g., station Cariamanga). In summary, a sparse network of rain gauge stations can hardly reveal the complex spatiotemporal dynamics of hazardous rains during ENSO events. Early warning systems urgently need spatial-explicit data to predict resulting flash floods and landslides.

Besides the hazardous rains during ENSO events, water is of utmost importance for irrigation and potable water supply in southern Ecuador, which is mostly fed by the high Andean ecosystems (Páramo). However, changes in land use (Buytaert et al., 2006a, 2007) and climate (Bradley et al., 2006) are threatening the country's water supplies. Approximately half of Ecuador's demand for electricity is met by hydro-electric power generation, making the atmospheric water supply vital to the country (Peláez-Samaniego et al., 2007). Here, the Rio Paute hydroelectric power plant in the south Ecuadorian Andes and its catchment, which is covered by the RadarNet-Sur network, is by far the most important facility in the country (Posso et al., 2015). Therefore, proper spatiotemporal precipitation monitoring is a tangible contribution to these rain-dependent activities in southern Ecuador.

Average precipitation in the Ecuadorian Andes is also characterized by a high spatiotemporal variability due to the dominant influence of mesoscale processes (see, e.g., Bendix et al., 2009; Trachte et al., 2010). The works of Buytaert et al. (2006b); Xie et al. (2011), and Campozano et al. (2016) reveal that spare rain gauge networks are insufficient at detecting variability in rainfall, especially in mountainous regions. Additional research points out that rain in complex terrain cannot properly be predicted with weather forecast models because they still struggle to provide accurate precipitation estimates, especially where convective precipitation is dominant (Wulfmeyer

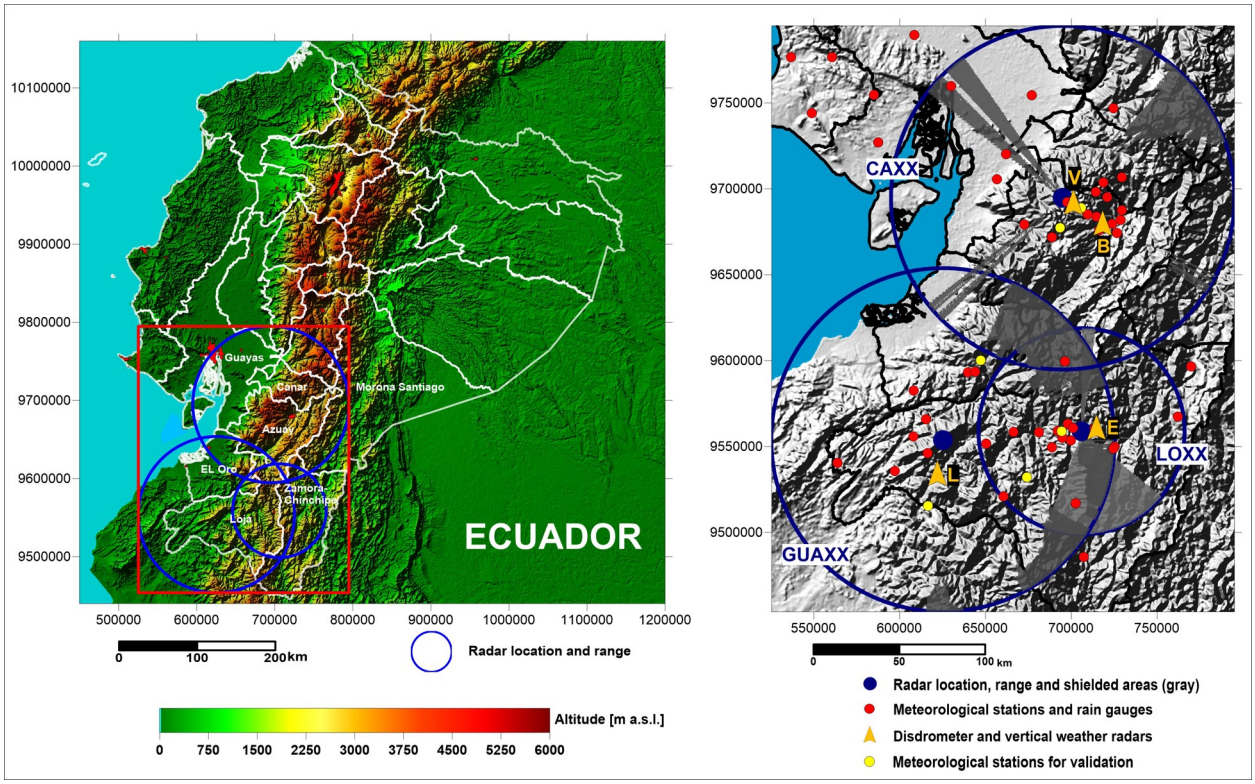

Figure 2.1: (left) Coverage and position of RadarNet-Sur. The meteorological stations are operated by INAMHI, the universities of Cuenca (UC) and Loja (UTPL), ETAPA EP, and the German research program PAK823-825. (right) Disdrometer sites are Laipuna (L) for the GUAXX radar, ECSF (E) for the LOXX radar, and La Virgen (V) as well as Balzay (B) for the CAXX radar. Gray hatching denotes areas with complete beam blockage (radar shadows) due to higher terrain. Coordinates UTM 17S are given in kilometers.

et al., 2008, 2011; Zollo et al., 2016).

2.2 Weather Radar Networks

Weather radar networks deliver high-resolution spatiotemporal information that could offer a viable solution to the problems mentioned above. Despite their indispensable utility for weather forecasting and disaster prevention with regard to spatiotemporal rainfall assessments, radar networks are currently less developed in South America as compared to industrialized countries (Meischner et al., 1997; Serafin and Wilson, 2000). The U.S. Next Generation Weather Radar (NEXRAD) network, a prominent example, was recently upgraded to dual-polarization (DP) technology. DP technology eliminates some problems associated with the previous Doppler systems because it

improves the detection of hydrometeor types, nonweather targets, and heavy rains (Brandes et al., 2002; Hall et al., 1980; Herzegh and Jameson, 1992). In countries like the United States, highly sophisticated mobile radar systems are even deployed in education (Bell et al., 2015). However, DP technology is hardly affordable for most countries, and economically developing parts of the world are nearly devoid of any operational weather radar network. In their stead, sophisticated space missions currently serve to broaden the coverage of precipitation estimates at a global level (Hou et al., 2014).

The repurposing of single-polarization (SP) X-band marine radar (Jensen, 2002; Lengfeld et al., 2014) has recently made it possible to build cost-effective ground radar networks. However, SP X-band systems have a long-standing history of known uncertainty, including (i) calibration, (ii) radar reflectivity (Z)-rain rate (R) relations, (iii) discerning between different types of liquid and solid precipitation particles to derive an adapted Z-R relation, (iv) clutter contamination and beam blockage, and (v) beam attenuation (Wilson and Brandes, 1979, and many thereafter). In addition, mountainous terrain enhances most of these known problems (Delrieu et al., 1999). Therefore, several approaches have been developed, including procedures to minimize the uncertainties by blending the radar signal with rain gauge observations where available (Seo and Breidenbach, 2002), which can partially mitigate these problems (e.g., Germann et al., 2006).

This article summarizes our initial experience with the newly implemented RadarNet-Sur in southern Ecuador. The network's SP X-band radars cost approximately $120,000 (U.S dollars), while more sophisticated C-band radars, used by weather services in economically advanced countries, are in the range of millions of dollars (Rollenbeck and Bendix, 2006, 2011; Selex, 2011). The latter represents an unaffordable radar network solution for most South American countries. RadarNet-Sur is an interdisciplinary effort involving German radar scientists and Ecuadorian stakeholders and serves as a prototype that may be replicable in other mountainous regions of economically developing countries. The network is the result of a knowledge transfer and capacity building project with the goal of bringing weather radar technology into service in southern Ecuador, as a model for a future national network. After a description of the instruments, data, and web infrastructure, our initial work related to data correction and calibration is presented before some examples are given that underpin the potential benefits of the network. The article closes with an outlook for future developments.

2.3 Infraestructure of the radar network

Three scanning SP X-band radars form the basis of the RadarNet-Sur network (Figures. 2.1, 2.2; Table 2.2). The older system, a Local Area Weather Radar (LAWR) purchased from the Danish Hydraulic Institute (DHI), is located at the summit of El Tiro (2850 m a.s.l.; LOXX radar) east of the city of Loja. The LAWR is based on the marine radar Furuno 1525MK3 and transmits at a frequency of 9410 $\pm$ 30 MHz with a bandwidth of 3 MHz and a pulse length of 1.2 μs at a repetition rate of 600 Hz. The corresponding fan beam antenna has a horizontal beamwidth of 0.92° and a vertical beamwidth of 10° (up and down). Internal resampling of the system returns image sizes of 240x240 pixels with a spatial resolution of 500 m at a maximum range of 60 km [for further details refer to Jensen (2002), Rollenbeck and Bendix (2006), and Pedersen et al. (2010)]. Image resolution can be sharpened up to 100 m by reducing the range (Fries et al., 2014).

Table 2.2: Scanning X-band (9410 MHz) radar systems used in RadarNet-Sur.

Radar	Location (° and m a.s.l.)	Type	Manufacturer	Nominal range (km)/resolution (m)
Cajas (CAXX)	2°45'S, 79°16'W; 4450	RS120 Rainscanner	Selex ES	Max 100/500-100
Loja (LOXX)	3°59'S, 79°08'W; 2850	LAWR	DHI	60/500 30/250 15/100
Guachaurco (GUAXX)	4°02'S, 79°52'W; 3100	RS120 Rainscanner	Selex ES	Max 100/500-100

The two newer systems are RS120 RAINSCANNER instruments (RS120) produced by Selex ES GmbH (Germany). They have a transmission frequency of 9410 $\pm$ 30 MHz, a bandwidth of 2.5 MHz, and a pulse length of 1.2-0.5 μs at a repetition rate of 833-1500 Hz. These systems have parabolic antennae with a horizontal and a vertical beamwidth of 2°. The RS120 have a maximum range of 100 km and provide user-defined images with different ranges and radial resolutions (100 to 500 m). The westernmost RS120 system is located on Cerro Guachaurco (3100 m a.s.l.; GUAXX radar) near the city of Celica. The highest radar unit of the network (and to our knowledge the highest worldwide) is situated at the edge of the Cajas National Park on Cerro Paragüillas (4450 m a.s.l.; CAXX radar) near the city of Cuenca, covering the headwaters of the Rio Paute catchment.

Figure 2.2: (top left) X-band scanning instruments of RadarNet-Sur: radar GUAXX RS120 RAINSCANNER on Cerro Guachaurco (3,100 m a.s.l.), (top right) GUAXX radar processing unit with backup power supply, (bottom left) radar LOXX LAWR at El Tiro (2850 m a.s.l.), and (bottom right) radar CAXX RS120 RAINSCANNER on Cerro Paragüillas (4450 m a.s.l.).

2.3.1 Additional instruments for improved rain estimates

To overcome the limitations of operating radar instruments in complex terrain and to improve the quality of the SP X-band weather radar retrievals, several project partners installed networks of

ground-based observation stations (Figures 2.1-2.3). The basic network of Automatic Weather Stations (AWS) and rain gauge observations includes the sparse operational network of the Ecuadorian Weather Service [Instituto Nacional de Meteorología e Hidrología (INAMHI)] and involves equipment from other public institutions. This is complemented by universities and research project AWS networks, which deliver rain gauge data in a high temporal resolution (10 min). The latter also cover remote mountain areas (e.g., Figure 2.3) that are largely not covered by the INAMHI's operational network due to the difficulty in accessing them. For specific data adjustment purposes, such as deriving the Z-R relations for different climate regimes with diverging rainfall characteristics (i.e., the Tumbesian dry forest, the mountain rain forest, and the Andean Páramo highlands), each radar area received one dedicated disdrometer site (Table 2.3).

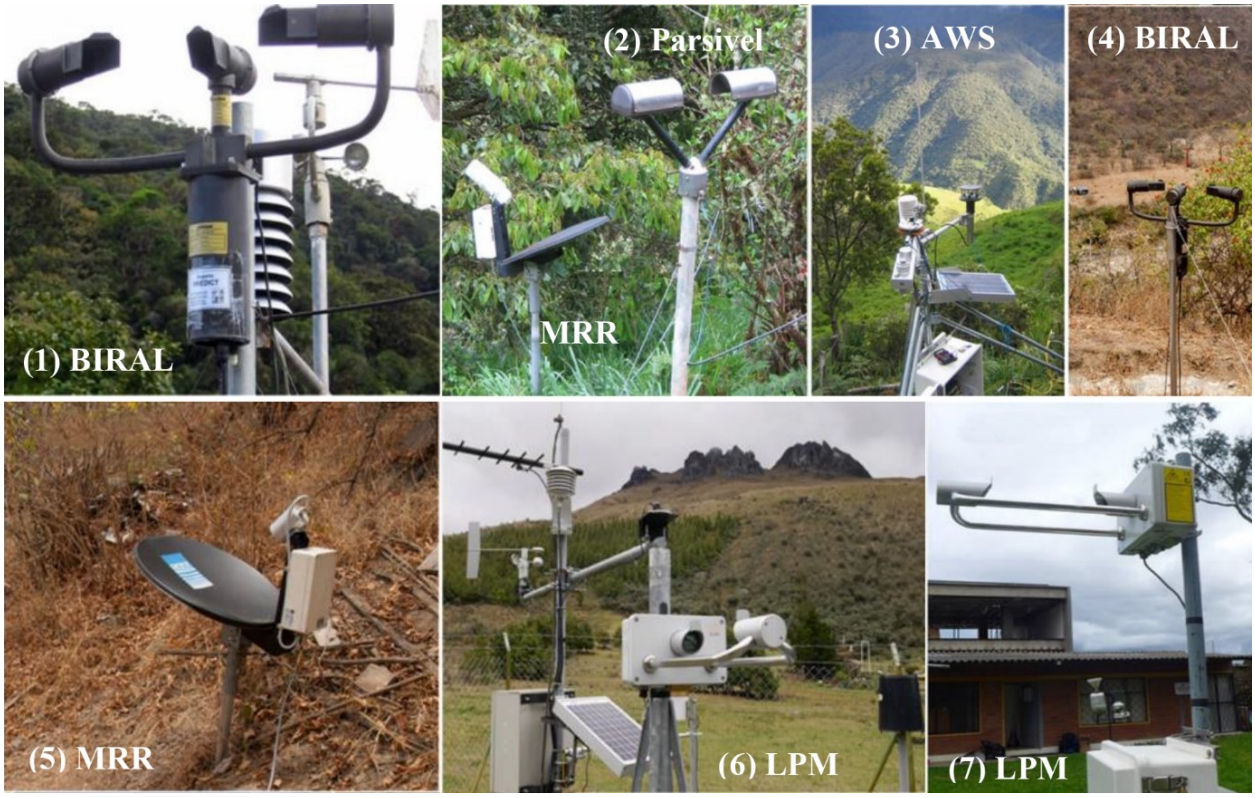

Figure 2.3: Instruments at the disdrometer sites: 1) Biral VPF-730, 2) MRR-2 and OTT Parsivel, and 3) Automatic Weather Station (Campbell) in the ECSF area of the mountain rain forest (for LOXX radar). 4) Biral VPF-730 and 5) MRR-2 at the Laipuna field station in the mountain dry forest (for GUAXX radar). 6) LPM and AWS at La Virgen and 7) LPM at Balzay (for CAXX radar).

The main sites in the LOXX area are the meteorological stations around the Estación Científica San Francisco (ECSF) research station (3°58'S, 79°4'W; Rollenbeck et al. (2007)), where three

Table 2.3: Specific sensors available for calibration.

Radar	Site	Location (° and m a.s.l.)	Instruments	Provided by
CAXX	La Virgen Balzay	2°45'S, 79°16'W; 4450 2°53'S, 79°02'W; 2610	2x Thies Clima LPM 2x AWS	U Cuenca
LOXX	ECSF	3°59'S, 79°08'W; 2850	Biral VPF-730 Metek MRR-2 OTT Parsivel AWS Nubiscope	U Marburg
GUAXX	Laipuna	4°02'S, 79°52'W; 3100	Biral VPF-730 Metek MRR-2 AWS	U Marburg U Erlangen

instruments are installed (see Figure 2.1, E). The first is a Biral VPF-730 present weather sensor, a light scatter instrument capable of measuring precipitation and retrieving a particle size matrix from the amplitude of the intensity and duration of the scintillation signal (Ellis et al., 2006). Because of some uncertainties with oblique particle incidence, the station is complemented by a second instrument, an OTT Parsivel (particle size and velocity) laser optical disdrometer, a light beam occlusion-type instrument that utilizes a horizontal light sheet (Ellis et al., 2006; Löffler-Mang and Jürg, 2000). The third instrument in the array is the vertically pointing frequency-modulated continuous wave (FM-CW) K-band Doppler radar METEK MRR-2 (Bendix et al., 2006; Klugmann et al., 1996).

For the GUAXX radar unit, a Biral VPF-730 and an MRR-2 instrument were installed at the Laipuna research station in the Tumbesian dry forest (see Figure 2.1, L). The CAXX radar area is represented by the station La Virgen at 3626 m a.s.l. in a shrub- dominated (sub-páramo) transition from upland forest to páramo (Hofstede et al., 2014), where a Thies Clima laser precipitation monitor (LPM; Lanzinger, 2006) complements an AWS (see Figure 2.1, V). Like the OTT Parsivel, the LPM precipitation monitor is also a light beam occlusion-type instrument that uses a horizontal light sheet to measure drop size distribution (DSD), drop fall velocities, and rainfall (Ellis et al., 2006). A comparison of data recorded by the LPM and Parsivel revealed a strong consistency in the measurements (de Moraes Frasson et al., 2011). A second LPM instrument is situated in the city of Cuenca (station Balzay; see Figure 2.1, B) at 2610 m a.s.l.

2.3.2 Data transmission and storage

For a radar network to meet stakeholder demands, data must be transmitted in real time, including the publication of radar composite images via a web interface (Cao et al., 2009). Initially, this requires a direct transmission from the remote radar sites (Figures 2.1, 2.2) to the main database, located in the planning department of the provincial government of Loja [Departamento de Planificación y Desarrollo Territorial, Gobierno Provincial de Loja (GPL); Figure 2.4]. For the LOXX radar, a general packet radio service (GPRS; cellular phone) connection was introduced, while a normal Internet connection guarantees real-time data transmission from the GUAXX radar. The CAXX radar uses radio transmission via a directional antenna, connecting the radar to the Intranet of Empresa pública municipal de Telecommunicaciones, Agua Potable, Alcantarillado y saneamiento (ETAPA EP) in Cuenca (http://geo.etapa.net.ec/visorradar/), where a gateway redirects the data stream to the GPL database. Every new image is saved in the database configured as a file system with standardized metadata, which is continuously mirrored to a server at the University of Marburg via FTP for backup reasons. The web page (www.radarnetsur.gob.ec) provides information to the public and includes basic information on the radar units, a simple explanation of the measurement principles in English and Spanish, the radar composite loop for all of southern Ecuador, and separate, more detailed maps of the three radar systems in higher spatial resolution for the last 3 h. Currently, the online radar loop displays radar reflectivity corrected for clutter contamination and beam blocking effects.

2.4 Data Processing

Both X-band weather radar systems, the LAWR and the RS120, require customized data processing strategies to combine them in the RadarNet-Sur. Therefore, new procedures were developed to guarantee a reliable rain-rate retrieval for the complex environment of the south Ecuadorian Andes. Because of the restrictions of the LAWR, specific procedures for (i) attenuation, (ii) beam blockage and occultation, as well as (iii) clutter correction had to be programmed and implemented. The RS120 does not have the same kind of technical limitations as the LAWR, but specific routines for image correction and the conversion of calibrated Z to R had to be developed and applied, too, due to the complex terrain where the radars are operating.

The subsequent paragraphs describe the custom-tailored data processing tools for both radar systems, which extend the proprietary software routines of the manufacturers, and the improvements

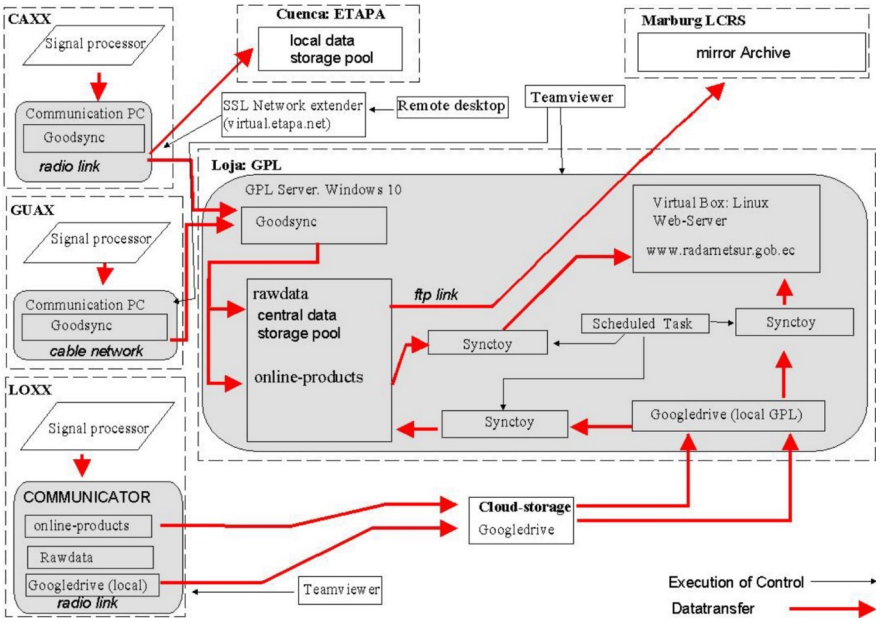

Figure 2.4: Data flow diagram of the three radar instruments. Because of the heterogeneous structure, different tools were employed for execution of control (TeamViewer, remote desktop/SSL network extender) and for data synchronization (GoodSync, Google Drive). All connections are password protected.

in R retrieval gained by the developments. A detailed description of the algorithm and mathematical background is given in Rollenbeck and Bendix (2011) and Fries et al. (2014).

As supplied by the manufacturer, the LAWR only provides the video signal from the Furuno radar monitor as dimensionless 8-bit counts. The radar reflectivity is not available. Thus, an empirical approach was developed for the conversion from raw counts to radar reflectivity and/or rain rate, which encompasses the following modules.

Geometric attenuation (i) is corrected by applying an exponential function dependent on distance and time because of the wear-out of the magnetron, which reduces the emitter power over time. The correction function amplifies the original signal up to a maximum factor of 64, which is further modified by the cumulated scan line totals to correct two-way attenuation caused by rain fields. (ii) A digital elevation model (DEM) is used to calculate beam blockage and occultation correction after passing obstructions like mountain tops. The returned signals are proportionally scaled up to the remaining beam fraction. Parameters for these procedures are derived from maps of cloud frequency, cloud-top altitude, and long-term average rain distribution (Bendix et al., 2004; INAMHI, 1993) to assess a plausible distribution of rain in the LAWR radar range. The routine for clutter removal (iii) is based on radar image time series to determine variance thresholds for affected

areas. Pixels with a variance lower than 50% of the mean signal are flagged as clutter. Bilinear interpolation techniques fill missing values in the radar map, caused by strong permanent beam obstructions and clutter fields, using data from the next available nonocculted/clutter-free points. The RS120 instrument provides readings of radar reflectivity. Data processing is executed by a proprietary software application (Selex, 2011) whose parameters were adapted to the unique local operation conditions in southern Ecuador. The modified standard procedures are as follows below.

Range and attenuation correction (i) applies a $1r^{-2}$ geometric correction for beam expansion (where r is the distance from the emitter), the squared refraction index of the air (0.93), and a transmit and receive path loss of 0.5 dB. The two-way atmospheric attenuation is set to 0.016 dB km^{-1} and the rain-rate-specific attenuation is approximated by a square function giving an attenuation of 0.15 dB km^{-1} at R = 1 mm h^{-1}, 3.0 dB km^{-1} at R = 10 mm h^{-1}, and 62 dB km^{-1} at R = 100 mm h^{-1}. For the occultation correction (ii) a user-defined log threshold of -3 dBZ is applied to suppress random noise. Furthermore, known occultation sectors are assigned manually, setting the start azimuth angle, stop azimuth angle, and start range. Clutter (iii) is identified by a built-in clutter-decision tree, which additionally requires a deeper analysis of radar images captured under rain-free conditions (Selex, 2011). Therefore, user-selected radar images from recent time series are examined to derive the viewing geometry and beam blockage within the radar range. The obtained clutter fields are manually flagged and combined with the clutter fields detected by the built in algorithm. Then, all clutter fields are set to No-Data and submitted to bilinear interpolation. The Selex software allows the operational integration of predefined Z-R relations (iv) based on a gauge- adapted conversion of corrected radar reflectivity to rain rate. However, it was found that the built-in conversion algorithm could not be used because it requires a much higher rain gauge density than available in the study area. Hence, the approach of using continuous relation surfaces (see below) developed for the LAWR was adapted for the RS120 instruments. This empirical approach for quantitative precipitation retrieval (iv) is applied for hourly time steps and establishes a direct relationship between corrected raw counts (LAWR) and Z (RS120) to R.

For each rain gauge location, a daily conversion factor between radar observations (raw counts, respectively, Z) and rainfall is established and then interpolated to form a factor matrix with the resolution of the radar images. The interpolation considers the vertical gradients of rainfall distribution as a drift variable, thus forming anisotropic distance measures between the calibration points. The factor matrix is applied to each hourly radar image to generate the final quantitative precipitation estimate. The images reproduce the spatial distribution of storm cells observed by the radar and the hourly rain rates registered by the rain gauge network. In the future, telemetric rain

gauge data will be included to derive R images in near-real time for applications like nowcasting and short-term alert systems, which is a special challenge in southern Ecuador considering its complex terrain.

Because of the strong influence of the topography in this mountainous region on beam propagation, the developed processing chain yields a clear improvement in the agreement between radar data and independent ground observations [shown later in Figure 2.9 (bottom)]. Uncorrected data from the radar systems (Figure 2.9a) are highly contaminated by beam blockage and clutter. The signal contributions of these disturbances are apparently much higher than the rain signal itself. Eliminating these contaminations from the signal is the crucial step in obtaining stable relations between radar data (raw counts, respectively, Z) and rainfall at the ground station network. Applying the weighted factor matrices, which also account for the local differences in precipitation type and process dynamics of rainfall formation, leads to the strong agreement between calibrated radar rainfall and independent ground rain measurements (Figure 2.9b). Rollenbeck and Bendix (2011) presented further evidence for the success of the developed data processing scheme. Their study revealed increasing correlation coefficients by comparing monthly and daily radar rainfall totals with rain gauge data over longer time scales. By applying the customized processing scheme, the determination coefficients increased from 0.12 to 0.34 to 0.8 to 0.98 for monthly and from 0.05 to 0.11 to 0.69 to 0.9 for daily rainfall totals.

2.4.1 Z-R relations

The Z-R relations (the relation of radar reflectivity Z to rain rate R) are generally derived from disdrometer data to improve the rainfall retrieval by adapting to the local rainfall conditions of the radar areas. In the RadarNet-Sur network, Z-R relations can only be applied to the Selex instruments (CAXX and GUAXX). While the analysis of X-band Z-R relations is still ongoing work, we currently use the standard Z-R relation of $Z = 200R^{1.6}$ as set by Selex. However, we have started to derive C-band Z-R relations from the disdrometer data available for each radar site to gain insights into potential variations among the different rainfall regimes of the network area. For this, the disdrometer data were split into single rainfall events to calculate the A and b parameters of the power law $Z = AR^b$. Then type-specific Z-R relations were calculated for the rainfall categories (e.g., drizzle, rain), which are classified by the present weather-type disdrometer sensors (such as LPM) based on a statistical relationship between particle diameter and fall velocity (Prodi et al., 2011). An example of Z-R relations derived from the two disdrometer sites in the CAXX area is

provided in Table 2.4. The few reported coefficients from the tropics vary between 1.2 and 1.9 for b and 180 and 350 for A (Blanchard, 1953; Wood, 2005). Our results for both high-altitude sites reveal smaller values, particularly for the drizzle categories at the highest site (La Virgen). This corresponds with Z-R relations of orographic rain reported by Blanchard (1953) in Hawaii and drizzle relations as found by Wood (2005). The result and the very high frequency of light rains in the high páramo area compared to midelevation sites (Figure 2.5) underpin the importance of providing site- and rain type-dependent Z-R relation for radar-based rainfall assessment.

Table 2.4: C-band Z-R relations for La Virgen and Balzay sites derived from the LPM instruments for 2015 (CAXX radar).

	La Virgen (3626 m a.s.l.)		**Balzay (2610 m a.s.l.)**	
Rainfall category	**A**	**b**	**A**	**b**
Light drizzle	2.14	0.47	3.03	0.54
Moderate drizzle	14.34	0.87	16.35	0.91
Light rain	33.13	1.02	45.69	1.03
Moderate rain	43.37	1.14	68.80	1.17
Heavy rain	114.45	1.26	98.39	1.28

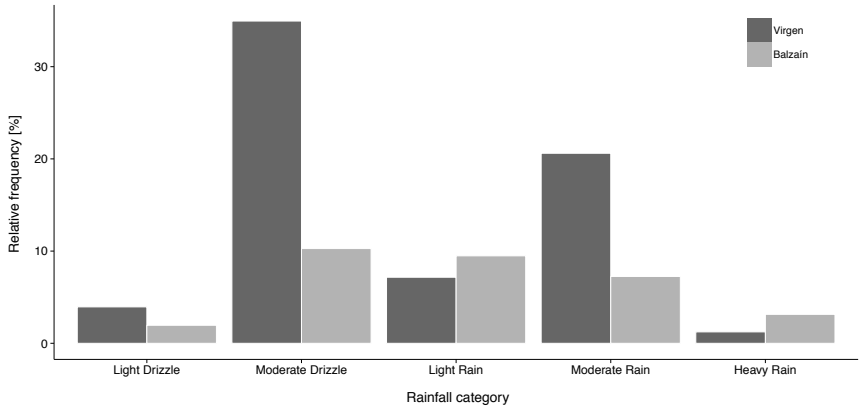

Figure 2.5: Occurrence of rain types for very high elevations (La Virgen station) and the midelevation site (Balzay station) in the CAXX radar area observed with LPM data. Drizzle conditions are much more frequent in the high altitude areas of the Cajas National Park where the prevailing easterly airstream meets the unsheltered east-facing slopes of the western cordillera for the wind situation in easterly exposed crest areas refer to Rollenbeck and Bendix (2011); Wagemann et al. (2015).

2.5 Application Examples

The radar network has several applications that benefit the participating stakeholder groups. Water-related applications, that is, the assessment of rainfall input for hydropower planning, require spatially explicit information about total rainfall over longer periods. Figure 2.6 clearly shows that extrapolation techniques based on sparse operational AWS networks are unable to provide reliable data for long-term rainfall assessment. Rainfall totals are particularly underestimated in the high and remote areas of southeastern Ecuador, not covered by AWS stations. Properly validated weather radar estimates help to overcome this problem (see Figure 2.6; Fries et al., 2014; Rollenbeck and Bendix, 2011).

2.5.1 ENSO (El Niño) rainfall monitoring

Nowcasting heavy rainfall is important for many purposes, such as aviation and disaster prevention. The latter is especially important in southern Ecuador because strong rains are frequently associated with both phases of the ENSO phenomenon: EN (mainly coastal Ecuador and western slopes of the western cordillera of the Andes) and LN (eastern cordillera).

According to the Oceanic Niño index (ONI), evaluated by National Oceanic and Atmospheric Administration (NOAA), the most recent 2015/16 EN event presented values of 2.3 [November-January (NDJ) 2015/16], putting it into the same very strong category as the last Super EN in 1997/98. However, for the remainder of the year 2016, SST anomalies returned to near normal values (+0.5 K) in the Niño-1+2 region, and no further extreme rainfall periods were observed in the coastal areas of southern Ecuador. The most significant heavy rainfall event was closely monitored just a few weeks after the full radar network went operational at the end of March 2015, when large positive SST anomalies shifted westward toward the coast of Ecuador. The first impact of this positive anomaly in rainfall was registered during the period 25-31 March 2015. Two major rain peaks were registered: one on the first and the other on the second-to-last day. The period was characterized by statistically significant above-average SST anomalies of up to 6 K (see hatched areas in Figure 2.7a). The widespread anomalies were associated with positive deviations in the zonal wind field in the lower troposphere (850-h Pa height level; Figure 2.7b) that ranged between +2 and +5 m s^{-1} and below normal sea level pressure (SLP) down to -2 h Pa (Figure 2.7c).

The strongest zonal wind field anomalies occurred in the Niño-3 region west of the Galapagos Islands, but westerly wind gusts also reached the coast of Ecuador. The greatest negative deviations

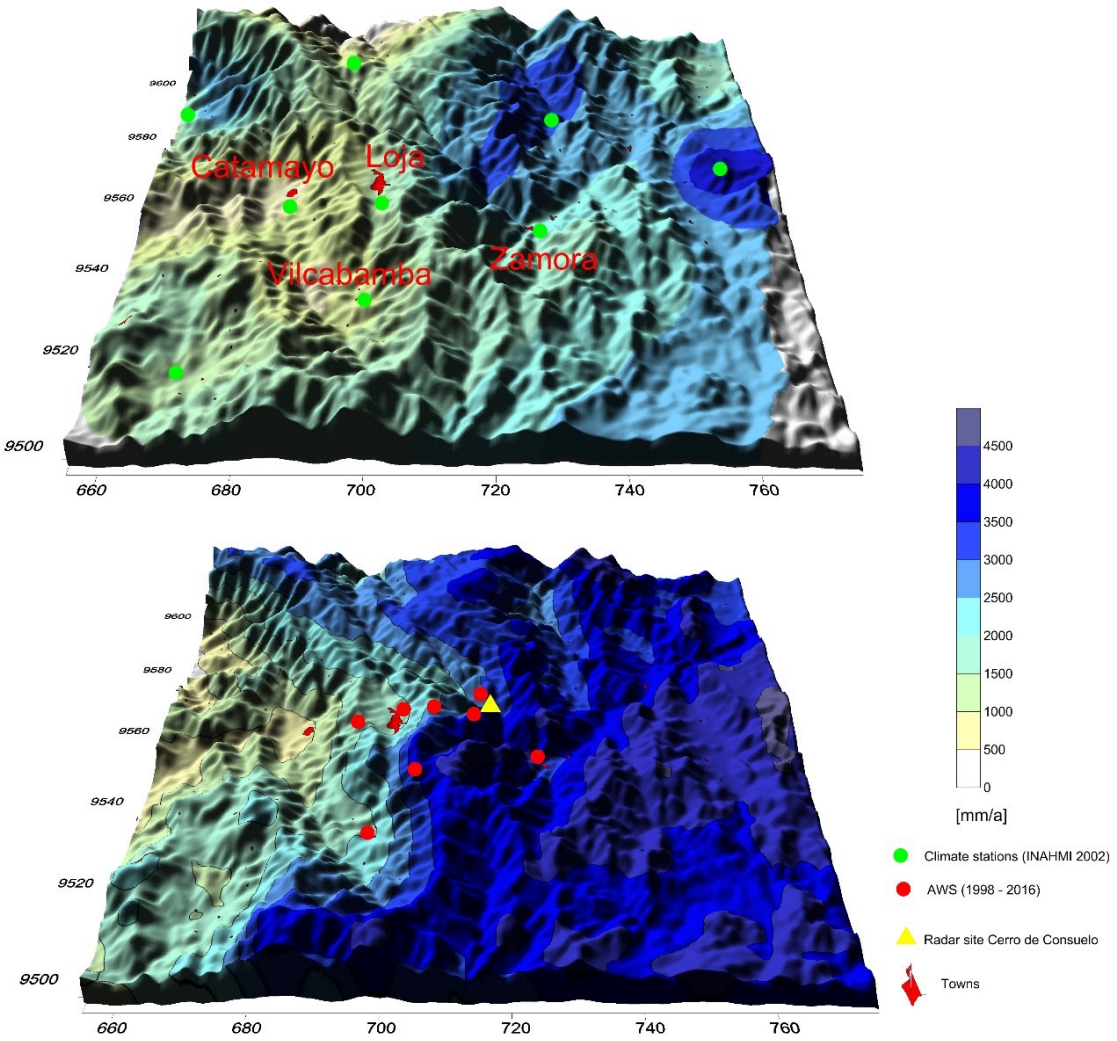

Figure 2.6: Comparison of average annual rainfall totals for the ECSF area based on extrapolated rain gauge data (INAMHI) and radar rain retrievals (LOXX radar), superimposed to a DEM. (top) Official map of INAMHI 1964-2000 and (bottom) LAWR radar observations 1998-2009. LAWR data are taken from the LOXX instrument formerly situated on the Cerro del Consuelo close to the ECSF station [4°00'S, 79°03'W; 3180 m a.s.l.; Rollenbeck and Bendix (2006)]. Significant extrapolation errors occur in the official rainfall map, such as severe underestimation in large swaths of the higher mountain ranges (e.g., between Vilcabamba and Zamora), which are a very important water supply area. Also, the radar data better represents the orographic enhancement of rainfall on the eastern slopes. Coordinates universal transverse Mercator (UTM) zone 17S (x and y axis) are given in kilometers.

of SLP appeared near the coast in the Niño-1+2 region, which corresponded to the anomalous SST pattern. Relative humidity and the wind field in $u - w$ directions taken at 4°S elucidate the

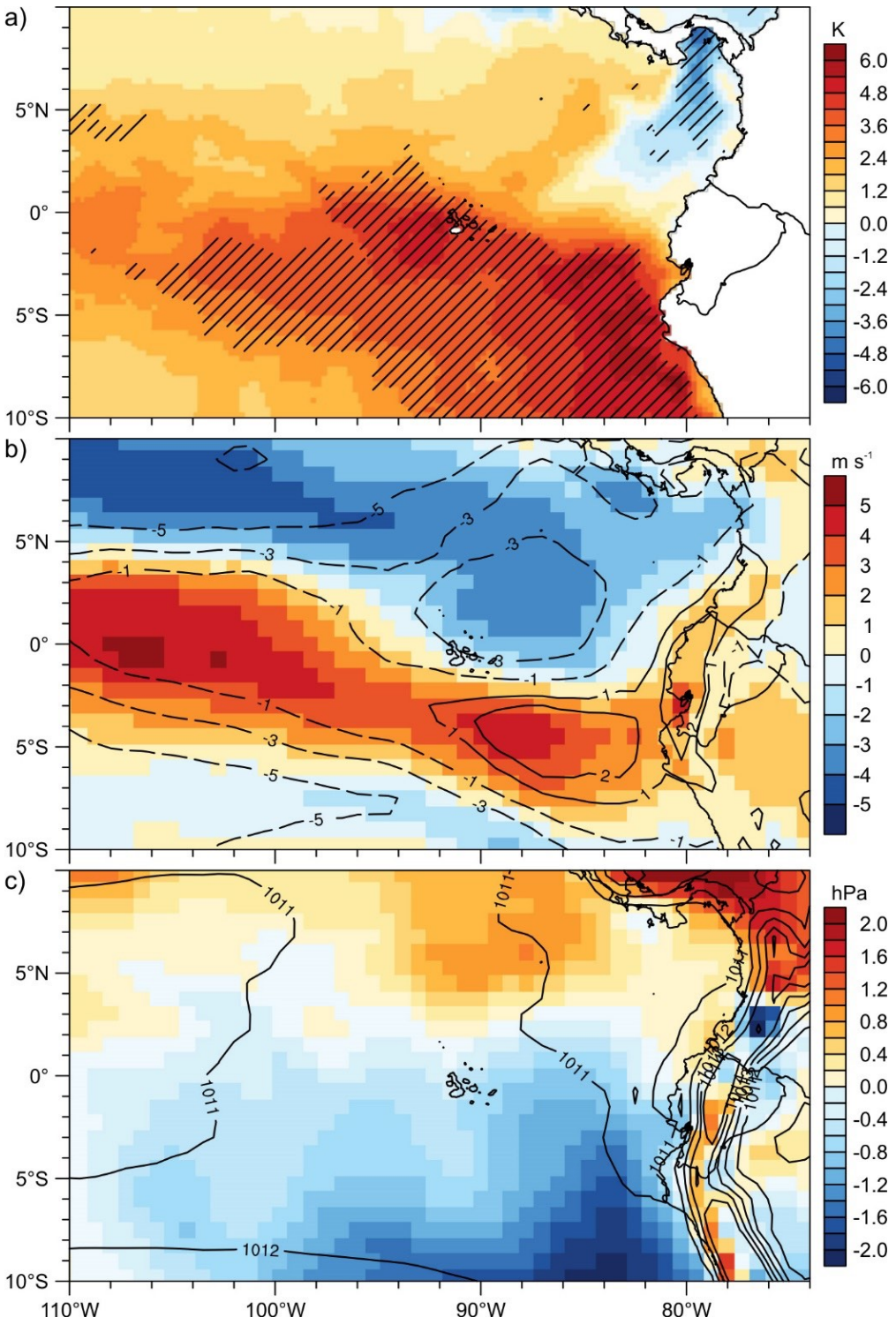

Figure 2.7: (a) Anomalous SST (K, shaded) for 25-30 Mar 2015 using the NOAA high-resolution SST data (Reynolds et al., 2007). Hatched areas indicate regions where the SST is above the 95th (and below the 5th) percentile of the 1981-2010 distribution. (b) Mean and anomalous zonal wind field (m s^{-1}), and (c) sea level pressure (hPa) for 25-30 Mar 2015 using the ERA-Interim data (Dee et al., 2011). Shading indicates anomalies relative to the 1981-2010 base period.

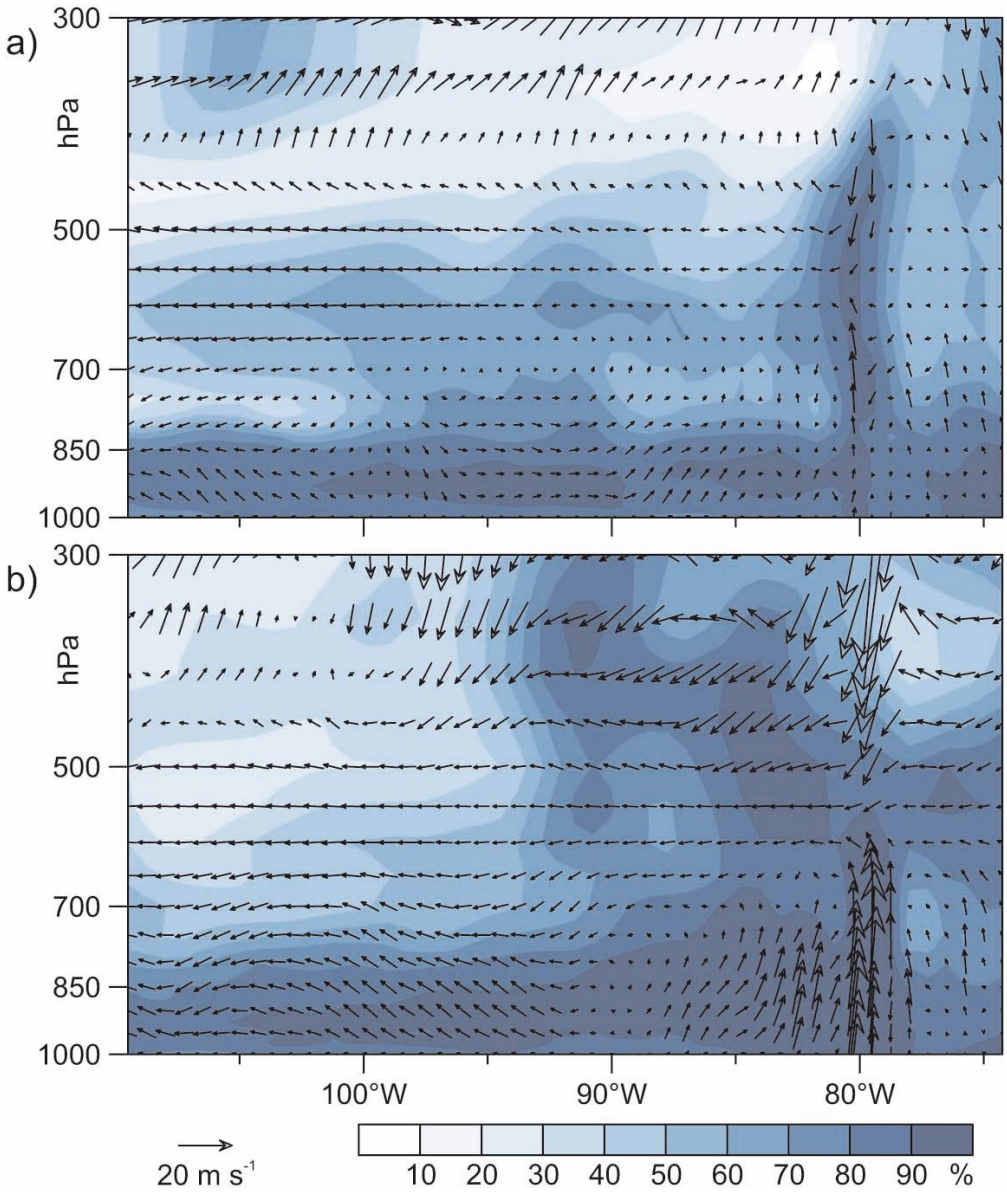

Figure 2.8: The x-z cross section of relative humidity (%, shaded) and wind field in $u - w$ directions (m s^{-1}, vectors) at 4°S for (a) 25 Mar and (b) 30 Mar 2015 using the ERA-Interim data (Dee et al., 2011).

atmospheric conditions during the two major rain peaks (25 and 30 March; Figure 2.8). A westerly component of the airflow was developed in the lower troposphere between 80° and 90°W on both days. The anomalous SST likely induced upward motion in air masses associated with the strongest vertical displacement of moisture near the area of the rain peaks (∼80°W).

The radar network and available station data indicate rainfall totals of >400 mm in the south-

western part of the province of Loja (Canton Zapotillo) close to the city of Celica during the final 7 days of March 2015 (Figure 2.9). This area (station Sabanilla, 701 m a.s.l.) was hit by heavy rainfall twice, on 25 and 30 March (154 and 201 mm day^{-1} respectively; Figure 2.11). Such extremes are exceedingly rare in this dry forest area; in fact, they are near the annual totals normally registered here. March 2015 saw up to 333% of the normal monthly rainfall amounts, 55% of which precipitated during the final 7 days of the month (S. Morocho Quezada, INAMHI Loja, 2016, personal communication). Confirming the European Centre for Medium-Range Weather Forecasts (ECMWF) interim reanalysis (ERA-Interim) data, rainstorm tracking (Figure 2.10) applied to the combined radar data from this period confirms several episodes of westerly cell propagation for the entire southwest. This flow carried humidity from the Pacific via northern Peru to the southern Ecuadorian mountain ridges, where orographic effects and local convergence further intensified rainfall formation (Fries et al., 2014). The short period of extreme rainfall claimed the lives of 17 people (official statement Secretaría de Gestión de Riesgos (SGR); international and local press; e.g., BBC Mundo, 2015; El Universo, 2015), triggered several landslides, and caused considerable damage to agriculture, property, and infrastructure. This extreme rainfall event exemplifies EN's impact on southwestern Ecuador, normally a rather dry region, which is more vulnerable to damage from torrential rain.

Figure 2.11 shows the daily precipitation amounts and natural hazards (floods and landslides) for the selected period. The orographic enhanced precipitation at the southern part of the coastal cordillera is clearly visible during all days. Particularly in the southern part of the network (radars LOXX and GUAXX), the rainfall regionally exceeded 25 mm day^{-1} (darkblue areas in Figure 2.11). Thus, landslides and flooding occurred throughout the entire southern region, especially in the province of Loja and at the upper part of the coastal province of El Oro (see Figure 2.1), caused by the predominate wind direction from the southwest. The slopes of the rural highways are particularly unstable and prone to landslides during heavy precipitation because of the extreme cuts made for road construction. Frequently, reports of landslides also come from the suburbs of towns and bigger cities because of inadequate land development and construction. Urban flooding is very common, particularly in the city of Loja (Figure 2.11, eastern road intersection), because the drainage systems are not designed for such high rainfall amounts. In rural areas, flooding generally occurs at the convergence of larger river systems, whose courses originate in the upper parts of the coastal cordillera. These plane areas are susceptible to inundation of large swaths, which include agricultural land. Most of the reported natural disasters were not triggered by a single rainfall event but rather by the extraordinary precipitation over several days during the study period (M.F. Riofrio

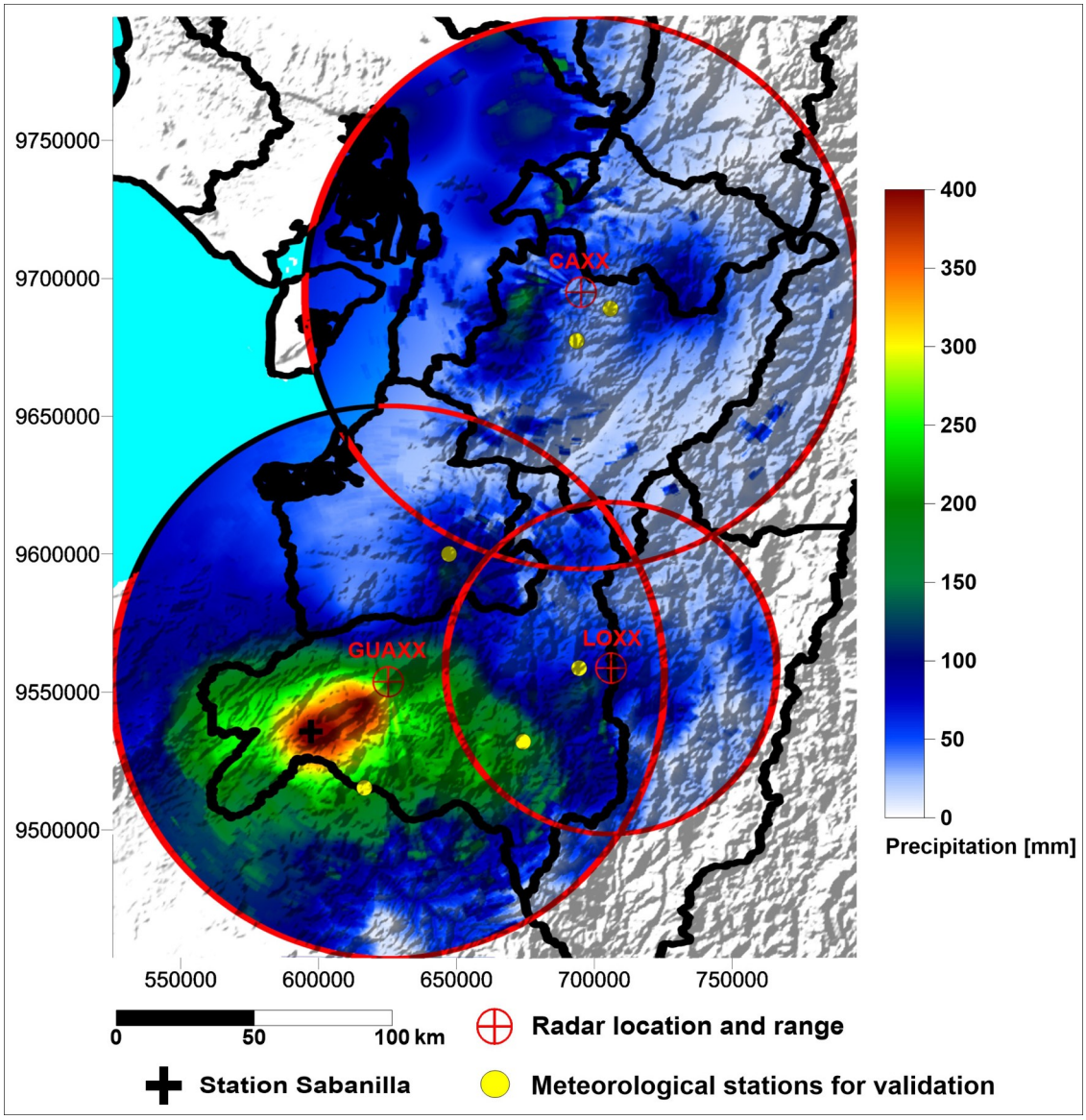

Figure 2.9: (top) Rainfall totals for the EN period 25-31 Mar 2015 as seen by RadarNet-Sur (areas of overlap are bilinearly interpolated and the maximum value set at each grid cell; coordinates UTM 17S are given in kilometers). (bottom) Observed daily rainfall (dBR) of official INAMHI stations (Ayapamba, Chirimachay, ColT, Gonzanama, Izhcayrrumi cabecera, and Macará; randomly chosen), which were not used for radar rain retrieval, compared to (a) uncorrected daily radar rainfall (dBR) and to (b) corrected daily radar rainfall (dBZ) presented in a common logarithmic scale.

Toscano, SGR Zonal 7, 2016, personal communication).

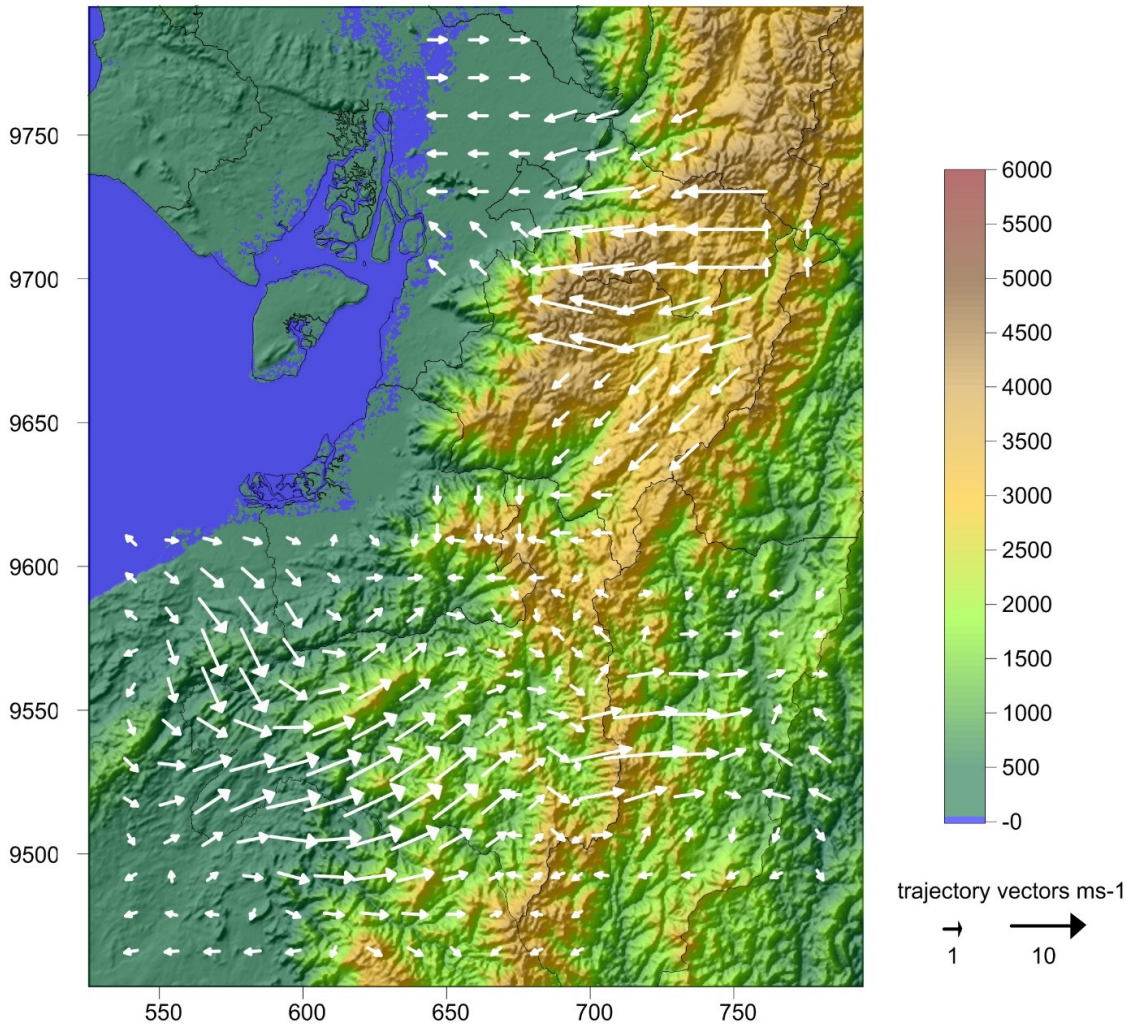

Figure 2.10: Average daily motion vectors derived from rain cell tracking for 25 Mar 2015. Westerly flow dominated the southern parts of the province of Loja and Zamora. Convergent structures are also visible close to where the strongest rainfall was registered on that day. Coordinates UTM 17S are given in kilometers.

2.5.2 Infrastructure planning

Aside from total rainfall, frequency maps of heavy rainfall are also of great importance for infrastructure surveillance and planning. Rain-driven landslides are a major problem in the Andes of southern Ecuador as they endanger travelers and damage the national road system (Bendix et al., 2013; Lozano et al., 2005; Muenchow et al., 2012; Schuster and Fleming, 1986; Solberg et al., 2003). In addition, the huge amount of sediments in rivers during extreme rainfall periods threatens hydropower generation by depositing the material upstream of the dams (Molina et al., 2008).

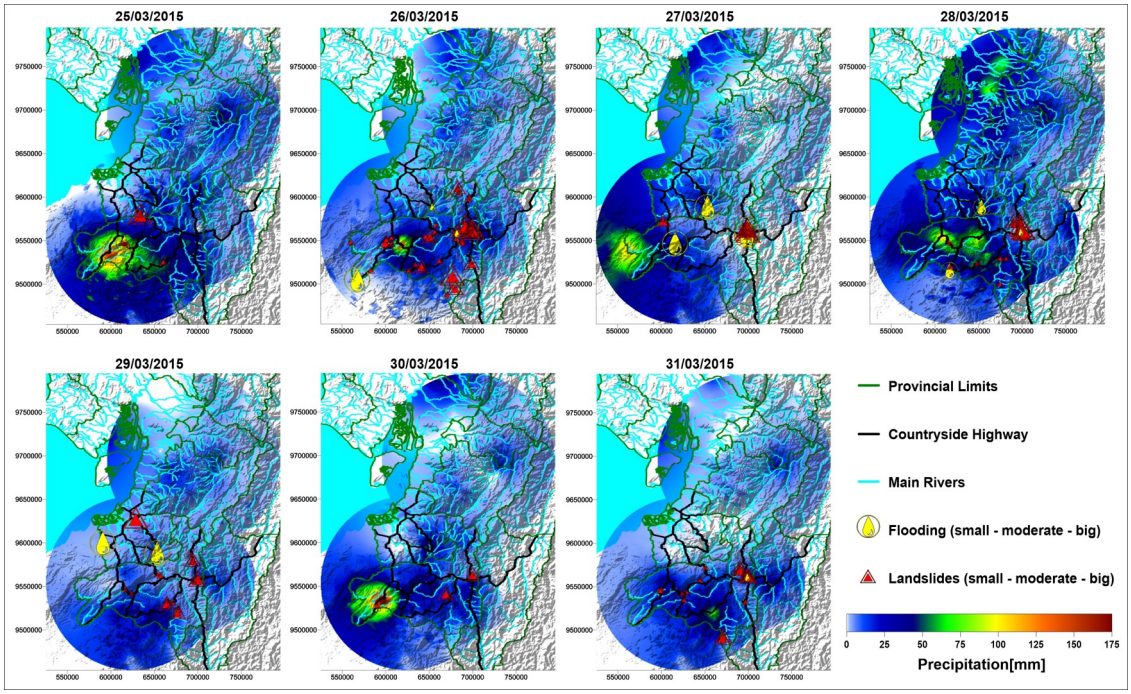

Figure 2.11: Daily precipitation maps for the time period 25-31 Mar 2015, including reported landslides and flooding events in southern Ecuador (region 7 = provinces: El Oro, Loja, and Zamora-Chinchipe; data from SGR). Coordinates UTM 17S are given in kilometers. The size of the marks is based on the strength of the event. (Landslides: small = less than 25-m road or less than 5 people affected; moderate = between 25- and 100-m road or between 5 and 10 people affected; and big = more than 100 m of a road or more than 10 people affected. Flooding: small = less than 10 people or less than 5 buildings affected; moderate = between 10 and 50 people or between 5 and 10 buildings affected; and big = more than 50 people or more than 10 buildings affected.)

Vulnerability to rain-driven landslides can be investigated by blending radar data with other relevant data available in geographic information systems (GIS), such as (i) environmental factors, that is, topography, soil, vegetation, or land cover; (ii) risk elements, that is, road network, settlement, and population data; and (iii) additional triggering data, that is, earthquake information, when applicable. To generate landslide susceptibility maps, trigger factors and environmental GIS data are combined in probabilistic models (according to Bendix et al., 2013, 2009; Fell et al., 2008; Ochoa-Cueva et al., 2015; Remondo et al., 2008).

An initial approach to derive a road vulnerability map is based on radar rainfall data and slope angle (Figure 2.12). Therefore, a DEM is analyzed to identify slopes with gradients of 30° or greater that are adjacent to major roads (<500 m) in combination with hourly precipitation maps. High risk for rain-driven landslides is indicated if the frequency of days with rain rates above 10 mm h^{-1}

exceeds 10 days yr^{-1} in such steep areas (Vorpahl et al., 2012).

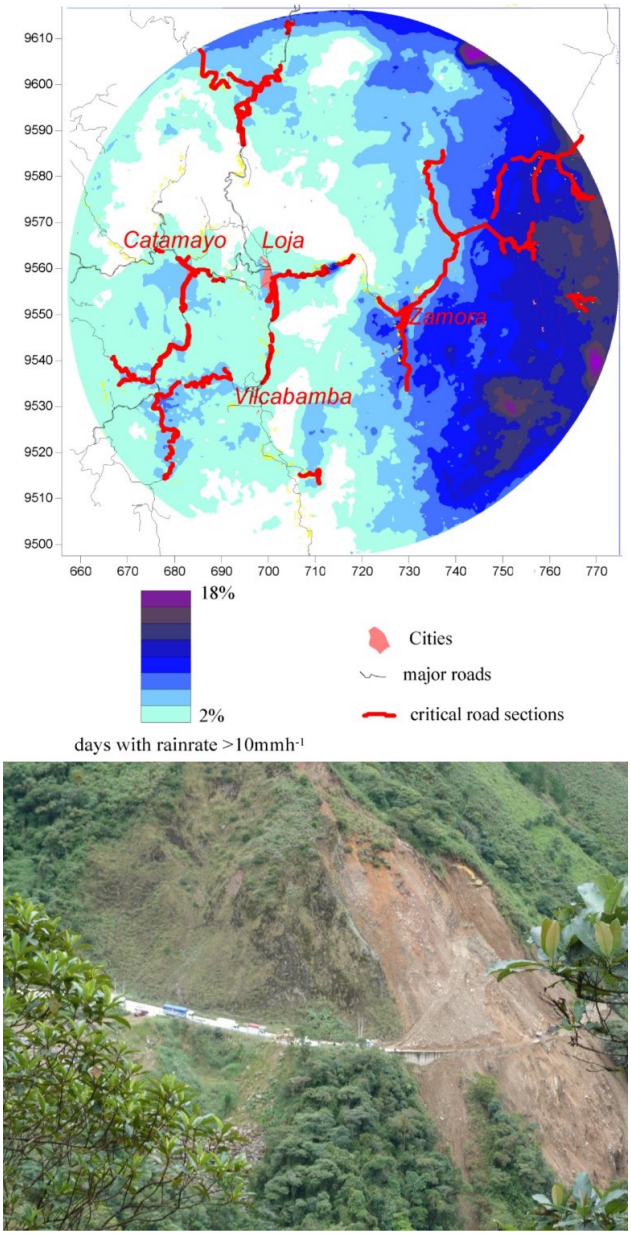

Figure 2.12: Risk assessment map for major roads derived from the DEM and radar data (LOXX instrument 1998 -2009). As in Fig. 2.6, the map area is displaced some 10 km east of the current range of the LOXX radar (former location). (top) Areas of high intensity rainfall are shown in shades of blue, representing the frequency of days with rain rates above 10 mm h^{-1}. Coordinates UTM 17S are given in kilometers. (bottom) Landslide caused by heavy rain on the major road Loja-Zamora.

Figure 2.12 shows two distinct areas of road vulnerability: east of the city of Loja, road density

is lower, but the whole traffic network passing through the steep mountains is threatened by the frequent orographic enhanced advective rainfalls. In contrast, the main risk factors in the western parts are convective showers of short duration but with high rainfall intensity occurring at certain areas (e.g., west of Vilcabamba).

2.6 Summary and Outlook

The RadarNet-Sur is a great step forward in spatial rainfall assessment and nowcasting in remote and mountainous areas of southern Ecuador. The project's installation phase was extremely successful considering the difficult environment facing Ecuador and its complex topography. It would have been impossible to finalize the network without the contribution of the local stakeholders, in particular, setting up the housing infrastructure, including the construction of a new power line for the CAXX radar and operating the web interface (www.radarnetsur.gob.ec; www.radarnetsur.de). After 1 year of operation, however, it has become clear that maintenance of the radar systems is a major issue; the electronic components are exposed to extreme climate conditions, power spikes, and corrosion. If a system fails, maintenance requires several hours of hiking, which, in the case of the CAXX radar, is above 4000 m a.s.l. Thus, maintaining operations will surely be a challenge, especially once the knowledge transfer and capacity building project has been completed. In terms of data processing, the quality of the calibration will be scrutinized in the future. The national weather service (INAMHI) will provide more AWS stations with real-time data transmission and, within the next few years, instantaneous gauge-supported radar image calibration will become possible. A remaining major issue is to derive site-specific X-band Z-R relations for different types of rain. Therefore, a rain-type detection scheme for the radar images will be developed, which allows assignment of the appropriate Z-R to every pixel during a specific rain event, which can be derived from more comprehensive disdrometer data. The web platform will be expanded to provide rain-rate information instead of radar reflectivity and access will be granted to historic data as well. The work on applications, such as the rain cell tracking combined with rainstorm warnings, will also continue. Finally, the consortium endeavors to expand this prototype network into a national system. The project has already received a great level of backing and been endorsed by the most relevant national ministries [Ministerio del Ambiente (MAE), Ministerio de Agricultura, Ganadería, Acuacultura y Pesca (MAGAP), and Secretaría de Gestión de Riesgos (SGR)], which have expressed their willingness to support further developments.

Acknowledgments

The authors thank the German Research Foundation (DFG) for generously funding the knowledge transfer project RadarNet-Sur (BE1780/31-1 and -2). The gratitude is extended to the DFG-funded programs “Platform for Biodiversity and Ecosystem Monitoring and Research in South Ecuador" (PAK823-825), projects C6 (BE1780/38-1) and C12 (BE1780/37-1, TR1201/1-1), and the previous DFG research units FOR402 and FOR816 for their contribution. The authors very much appreciate the significant financial contributions for infrastructural development and operation made by GPL (Gobierno Provincial de Loja), ETAPA EP (Empresa Pública Municipal de Telecomunicaciones, Agua potable, Alcantarillado y Saneamiento de Cuenca, Ecuador), and the Technical University of Loja (Universidad Técnica Particular de Loja; UTPL). Thanks are also due to the University of Marburg (LCRS) for significantly supporting the purchase of one radar unit and the Dirección General de Aviación Civil del Ecuador (DGAC) and the Ecuadorian telecommunications regulatory authority (Agencia de Regulación y Control de las Telecomunicaciones, SENATEL) for the permission to operate RadarNet-Sur. Finally, we thank Gregory Gedeon for text revision.

References

BBC Mundo (2015). Ecuador: Al menos 17 muertos por temporada de lluvias. *BBC Mundo*. Accesed 22 March 2016. [Available online at www.bbc.com/mundo/ultimas_noticias/2015/03/150331_ultnot_ecuador_muertos_temporada_lluvias_lv].

Bell, M. M., Ballard, R. A., Bauman, M., Foerster, A. M., Frambach, A., Kosiba, K. A., Lee, W. C., Rees, S. L., and Wurman, J. (2015). The Hawaiian Educational Radar Opportunity (HERO). *Bulletin of the American Meteorological Society*, 96(12):2167–2181.

Bendix, A. and Bendix, J. (2006). Heavy rainfall episodes in Ecuador during El Niño events and associated regional atmospheric circulation and SST patterns. *Advances in Geosciences*, 6:43–49.

Bendix, J. (1999). A comparative analysis of the major el Niño events in Ecuador and northern Peru over the last two decades. *Zbl. Geol. Paläntol.*, 7(8):1119–1131.

Bendix, J. (2000). Precipitation dynamics in Ecuador and northern Peru during the 1991/92 El Niño: A remote sensing perspective. *International Journal of Remote Sensing*, 21(3):533–548.

Bendix, J. (2004). Extreme events and climate variability in the Andes of Ecuador and Peru (in German). *Geographische Rundschau*, 56:10–16.

Bendix, J., Dislich, C., Huth, A., Huwe, B., Ließ, M., Schröder, B., Thies, B., Vorpahl, P., Wagemann, J., and Wilcke, W. (2013). Natural Landslides Which Impact Current Regulating Services: Environmental Preconditions and Modeling. In Bendix, J., Beck, E., Bräuning, A., Makeschin, F., Mosandl, R., Scheu, S., and Wilcke, W., editors, *Ecosystem Services, Biodiversity and Environmental Change in a Tropical Mountain Ecosystem of South Ecuador*, pages 153–170. Springer Berlin Heidelberg, Berlin, Heidelberg.

Bendix, J., Rollenbeck, R., and Palacios, E. (2004). Cloud detection in the tropics - A suitable tool for climate-ecological studies in the high mountains of Ecuador. *International Journal of Remote Sensing*, 25(21):4521–4540.

Bendix, J., Rollenbeck, R., and Reudenbach, C. (2006). Diurnal patterns of rainfall in a tropical Andean valley of southern Ecuador as seen by a vertically pointing K-band Doppler radar. *International Journal of Climatology*, 26(6):829–846.

Bendix, J., Stefan, G., Reudenbach, C., and Bendix, A. (2003). A Case Study on Rainfall Dynamics during el Niño/La Niña 1997/99 in Ecuador and surroundings areas as inferred from GOES-8 and TRMM-PR Observations. *Erdkunde*, 57(2):82–93.

Bendix, J., Trachte, K., Cermak, J., Rollenbeck, R., and Naub, T. (2009). Formation of convective clouds at the foothills of the tropical eastern Andes (South Ecuador). *Journal of Applied Meteorology and Climatology*, 48(8):1682–1695.

Bendix, J., Trachte, K., Palacios, E., Rollenbeck, R., Göttlicher, D., Nauss, T., and Bendix, A. (2011). El Niño meets La Niña-anomalous rainfall patterns in the "traditional" El Niño region of Southern Ecuador. *Erdkunde*, 65(2):151–167.

Blanchard, D. C. (1953). Raindrop-size Distribution in Hawaiian Rains. *Journal of Meteorology*, 10(6):457–473.

Bradley, R. S., Vuille, M., Diaz, H. F., and Vergara, W. (2006). Threats to water supplies in the tropical Andes. *Science*, 312(5781):1755–1756.

Brandes, E. A., Zhang, G., and Vivekanandan, J. (2002). Experiments in Rainfall Estimation with a Polarimetric Radar in a Subtropical Environment. *Journal of Applied Meteorology*, 41:474–485.

Buytaert, W., Célleri, R., Bièvre, B. D., Cisneros, F., Wyseure, G., Deckers, J., and Hofstede, R. (2006a). Human impact on the hydrology of the Andean páramos. *Earth-Science Reviews*, 79(1):53–72.

Buytaert, W., Célleri, R., Willems, P., Bièvre, B. D., and Wyseure, G. (2006b). Spatial and temporal rainfall variability in mountainous areas: A case study from the south Ecuadorian Andes. *Journal of Hydrology*, 329(3):413–421.

Buytaert, W., Iñiguez, V., and Bièvre, B. D. (2007). The effects of afforestation and cultivation on water yield in the Andean páramo. *Forest Ecology and Management*, 251(1):22–30.

Campozano, L., Célleri, R., Trachte, K., Bendix, J., and Samaniego, E. (2016). Rainfall and Cloud Dynamics in the Andes: A Southern Ecuador Case Study. *Advances in Meteorology*, 2016.

Cao, Y., Yang, C., and Wong, D. W. (2009). An interoperable spatiotemporal weather radar data dissemination system. *International Journal of Remote Sensing*, 30(5):1313–1326.

Capotondi, A., Wittenberg, A. T., Newman, M., Di Lorenzo, E., Yu, J.-Y., Braconnot, P., Cole, J., Dewitte, B., Giese, B., Guilyardi, E., Jin, F.-F., Karnauskas, K., Kirtman, B., Lee, T., Schneider, N., Xue, Y., and Yeh, S.-W. (2015). Understanding enso diversity. *Bulletin of the American Meteorological Society*, 96:921–938.

de Moraes Frasson, R. P., da Cunha, L. K., and Krajewski, W. F. (2011). Assessment of the Thies optical disdrometer performance. *Atmospheric Research*, 101(1):237–255.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kallberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F. (2011). The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, 137(656):553–597.

Delrieu, G., Serrar, S., Guardo, E., and Creutin, J. D. (1999). Rain Measurement in Hilly Terrain with X-Band Weather Radar Systems: Accuracy of Path-Integrated Attenuation Estimates Derived from Mountain Returns. *Journal of Atmospheric and Oceanic Technology*, 16(4):405–416.

El Universo (2015). Deslizamientos y lluvias afectan provincia de loja. *BBC Mundo*. Accesed 22 March 2016. [Available online at www.eluniverso.com/noticias/2015/03/25/nota/4702261/deslizamientos-lluvias-afectan-provincia-loja].

Ellis, R. A., Sandford, A. P., Jones, G. E., Richards, J., Petzing, J., and Coupland, J. M. (2006). New laser technology to determine present weather parameters. *Measurement Science and Technology*, 17(7):1715–1722.

Fell, R., Corominas, J., Bonnard, C., Cascini, L., Leroi, E., and Savage, W. Z. (2008). Guidelines for landslide susceptibility, hazard and risk zoning for land use planning. *Engineering Geology*, 102(3):85–98.

Fries, A., Rollenbeck, R., Bayer, F., Gonzalez, V., Oñate-Valivieso, F., Peters, T., and Bendix, J. (2014). Catchment precipitation processes in the San Francisco valley in southern Ecuador: combined approach using high-resolution radar images and in situ observations. *Meteorology and Atmospheric Physics*, 126(1-2):13–29.

Germann, U., Galli, G., Boscacci, M., and Bolliger, M. (2006). Radar precipitation measurement in a mountainous region. *Quarterly Journal of the Royal Meteorological Society*, 132:1669–1692.

Hall, M. P. M., Cherry, S. M., Goddard, J. W. F., and Kennedy, G. R. (1980). Rain drop sizes and rainfall rate measured by dual-polarization radar. *Nature*, 285(5762):195–198.

Herzegh, P. H. and Jameson, A. R. (1992). Observing Precipitation through Dual-Polarization Radar Measurements. *Bulletin of the American Meteorological Society*, 73(9):1365–1376.

Hofstede, R. G. M., Dickinson, K. J. M., Mark, A. F., and Narváez, E. (2014). A Broad Transition from Cloud Forest to Páramo Characterizes an Undisturbed Treeline in Parque Nacional Llanganates, Ecuador. *Arctic, Antarctic, and Alpine Research*, 46(4):975–986.

Hou, A. Y., Kakar, R. K., Neeck, S., Azarbarzin, A. A., Kummerow, C. D., Kojima, M., Oki, R., Nakamura, K., and Iguchi, T. (2014). The Global Precipitation Measurement Mission. *Bulletin of the American Meteorological Society*, 95(5):701–722.

INAMHI (1993). Anuario Meteorológico No. 33, Edición Especial, Quito-Ecuador 1995. *INAMHI*. [Available online at www.serviciometeorologico.gob.ec/wp-content/uploads/anuarios/meteorologicos/].

Jensen, N. E. (2002). X-band local area weather radar - preliminary calibration results. *Water Science & Technology*, 45:135–138.

Kao, H.-Y. and Yu, J.-Y. (2009). Contrasting Eastern-Pacific and Central-Pacific Types of ENSO. *Journal of Climate*, 22(3):615–632.

Klugmann, D., Heinsohn, K., and Kirtzel, H.-J. (1996). A low cost 24 GHz FM-CW Doppler radar rain profiler. *Contributions to Atmospheric Physics*, 69:247–253.

Lanzinger, E. (2006). Rainfall amount and intensity measured by the Thies laser precipitation monitor. *TECO-2006, Geneva, Switzerland*, pages 1–9.

Lengfeld, K., Clemens, M., Münster, H., and Ament, F. (2014). Performance of high-resolution X-band weather radar networks — the PATTERN example. *Atmospheric Measurement Techniques*, 7(12):4151–4166.

Löffler-Mang, M. and Jürg, J. (2000). An optical disdrometer for measuring size and velocity of hydrometeors. *Journal of Atmospheric and Oceanic Technology*, 17:130–139.

Lozano, P., Bussmann, R. W., and Küppers, M. (2005). Landslides as ecosystem disturbance -their implications and importance in South Ecuador. *Lyonia*, 8(1).

Meischner, P., Collier, C., Illingworth, A., Joss, J., and Randeu, W. (1997). Advanced Weather Radar Systems in Europe: The COST 75 Action. *Bulletin of the American Meteorological Society*, 78(7):1411–1430.

Molina, A., Govers, G., Poesen, J., Hemelryck, H. V., De Bièvre, B., and Vanacker, V. (2008). Environmental factors

controlling spatial variation in sediment yield in a central Andean mountain area. *Geomorphology*, 98(3):176–186.

Muenchow, J., Brenning, A., and Richter, M. (2012). Geomorphic process rates of landslides along a humidity gradient in the tropical Andes. *Geomorphology*, 139-140:271–284.

Ochoa-Cueva, P., Fries, A., Montesinos, P., Rodríguez-Díaz, J. A., and Boll, J. (2015). Spatial Estimation of Soil Erosion Risk by Land-cover Change in the Andes OF Southern Ecuador. *Land Degradation & Development*, 26(6):565–573.

Pedersen, L., Jensen, N. E., and Madsen, H. (2010). Calibration of Local Area Weather Radar-Identifying significant factors affecting the calibration. *Atmospheric Research*, 97(1-2):129–143.

Peláez-Samaniego, M. R., Garcia-Perez, M., Cortez, L. A. B., Oscullo, J., and Olmedo, G. (2007). Energy sector in Ecuador: Current status. *Energy Policy*, 35(8):4177–4189.

Posso, F., Espinoza, J. L., Sánchez, J., and Zalamea, J. (2015). Hydrogen from hydropower in Ecuador: Use and impacts in the transport sector. *International Journal of Hydrogen Energy*, 40(45):15432–15447.

Prodi, F., Caracciolo, C., D'Adderio, L. P., Gnuffi, M., and Lanzinger, E. (2011). Comparative investigation of Pludix disdrometer capability as Present Weather Sensor (PWS) during the Wasserkuppe campaign. *Atmospheric Research*, 99(1):162–173.

Remondo, J., Bonachea, J., and Cendrero, A. (2008). Quantitative landslide risk assessment and mapping on the basis of recent occurrences. *Geomorphology*, 94(3):496–507.

Reynolds, R. W., Smith, T. M., Liu, C., Chelton, D. B., Casey, K. S., and Schlax, M. G. (2007). Daily High-Resolution-Blended Analyses for Sea Surface Temperature. *Journal of Climate*, 20(22):5473–5496.

Rollenbeck, R. and Bendix, J. (2006). Experimental calibration of a cost-effective X-band weather radar for climate ecological studies in southern Ecuador. *Atmospheric Research*, 79(3-4):296–316.

Rollenbeck, R. and Bendix, J. (2011). Rainfall distribution in the Andes of southern Ecuador derived from blending weather radar data and meteorological field observations. *Atmospheric Research*, 99:277–289.

Rollenbeck, R., Bendix, J., Fabian, P., Boy, J., Wilcke, W., Dalitz, H., Oesker, M., and Emck, P. (2007). Comparison of Different Techniques for the Measurement of Precipitation in Tropical Montane Rain Forest Regions. *Journal of Atmospheric and Oceanic Technology*, 24(2):156–168.

Rossel, F. and Cadier, E. (2009). El Niño and prediction of anomalous monthly rainfalls in Ecuador. *Hydrological Processes*, 23(22):3253–3260.

Schuster, R. L. and Fleming, R. W. (1986). Economic Losses and Fatalities Due to Landslides. *Environmental and Engineering Geoscience*, xxiii(1):11–28.

Selex (2011). Instruction manual-RainView Analyzer user guide-Release 5.33.0. Technical report, Selex Systems Integration GmbH Neuss-Rosellen.

Seo, D.-J. and Breidenbach, J. P. (2002). Real-Time Correction of Spatially Nonuniform Bias in Radar Rainfall Data Using Rain Gauge Measurements. *Journal of Hydrometeorology*, 3(2):93–111.

Serafin, R. J. and Wilson, J. W. (2000). Operational Weather Radar in the United States: Progress and Opportunity. *Bulletin of the American Meteorological Society*, 81(3):501–518.

Solberg, S., Hale, D., and Benavides, J. (2003). Natural Disaster Management and the Road Network in Ecuador : Policy Issues and Recommendations. Technical report, Inter-American Development Bank. Sustainable Development Department.

Trachte, K., Rollenbeck, R., and Bendix, J. (2010). Nocturnal convective cloud formation under clear-sky conditions at the eastern Andes of south Ecuador. *Journal of Geophysical Research: Atmospheres*, 115(D24).

Vorpahl, P., Elsenbeer, H., Märker, M., and Schröder, B. (2012). How can statistical models help to determine driving factors of landslides? *Ecological Modelling*, 239:27–39.

Vos, R., Velasco, M., and Labastida, E. (1999). Economic and social effects of "El Niño" in Ecuador, 1997-8. *ISS Working Paper Series/General Series*, 292:1–55.

Vuille, M., Bradley, R. S., and Keimig, F. (2000). Climate Variability in the Andes of Ecuador and Its Relation to Tropical Pacific and Atlantic Sea Surface Temperature Anomalies. *Journal of Climate*, 13(14):2520–2535.

Wagemann, J., Thies, B., Rollenbeck, R., Peters, T., and Bendix, J. (2015). Regionalization of wind-speed data to analyse tree-line wind conditions in the eastern Andes of southern Ecuador. *Erdkunde*, 69(1):3–19.

Wilson, J. W. and Brandes, E. A. (1979). Radar measurement of rainfall– a summary. *Bulletin of the American Meteorological Society*, 60:1048–1058.

Wood, R. (2005). Drizzle in Stratiform Boundary Layer Clouds. Part II: Microphysical Aspects. *Journal of the Atmospheric Sciences*, 62(9):3034–3050.

Wulfmeyer, V., Behrendt, A., Bauer, H. S., Kottmeier, C., Corsmeier, U., Blyth, A., Craig, G., Schumann, U., Hagen, M., Crewell, S., Di Girolamo, P., Flamant, C., Miller, M., Montani, A., Mobbs, S., Richard, E., Rotach, M. W., Arpagaus, M., Russchenberg, H., Schlüssel, P., König, M., Gärtner, V., Steinacker, R., Dorninger, M., Turner, D. D., Weckwerth, T., Hense, A., and Simmer, C. (2008). Research Campaign: The convective and orographically induced precipitation study. *Bulletin of the American Meteorological Society*, 89(10):1477–1486.

Wulfmeyer, V., Behrendt, A., Kottmeier, C., Corsmeier, U., Barthlott, C., Craig, G. C., Hagen, M., Althausen, D., Aoshima, F., Arpagaus, M., Bauer, H.-S., Bennett, L., Blyth, A., Brandau, C., Champollion, C., Crewell, S., Dick, G., Di Girolamo, P., Dorninger, M., Dufournet, Y., Eigenmann, R., Engelmann, R., Flamant, C., Foken, T., Gorgas, T., Grzeschik, M., Handwerker, J., Hauck, C., Höller, H., Junkermann, W., Kalthoff, N., Kiemle, C., Klink, S., König, M., Krauss, L., Long, C. N., Madonna, F., Mobbs, S., Neininger, B., Pal, S., Peters, G., Pigeon, G., Richard, E., Rotach, M. W., Russchenberg, H., Schwitalla, T., Smith, V., Steinacker, R., Trentmann, J., Turner, D. D., van Baelen, J., Vogt, S., Volkert, H., Weckwerth, T., Wernli, H., Wieser, A., and Wirth, M. (2011). The Convective and Orographically-induced Precipitation Study (COPS): the scientific strategy, the field phase, and research highlights. *Quarterly Journal of the Royal Meteorological Society*, 137(S1):3–30.

Xie, H., Zhang, X., Yu, B., and Sharif, H. (2011). Performance evaluation of interpolation methods for incorporating rain gauge measurements into NEXRAD precipitation data: a case study in the Upper Guadalupe River Basin.

Hydrological Processes, 25(24):3711–3720.

Zollo, A. L., Rillo, V., Bucchignani, E., Montesarchio, M., and Mercogliano, P. (2016). Extreme temperature and precipitation events over Italy: assessment of high-resolution simulations with COSMO-CLM and future scenarios. *International Journal of Climatology*, 36(2):987–1004.

3 Analysis of Rain Types and Their Z-R Relationships at Different Locations in the High Andes of Southern Ecuador

This chapter is published in *Journal of Applied Meteorology and Climatology (AMS)*, 36, 3065-3080, 2017.
Received: 10 January 2017
Accepted: 27 July 2017
First published online: 17 November 2017
DOI: https://doi.org/10.1175/JAMC-D-17-0009.1

Analysis of Rain Types and Their Z-R Relationships at Different Locations in the High Andes of Southern Ecuador

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

[1] Laboratory for Climatology and Remote Sensing, Faculty of Geography, University of Marburg, Marburg, Germany

[2] Departamento de Recursos Hídricos y Ciencias Ambientales, and Facultad de Ciencias Químicas, Universidad de Cuenca, Cuenca, Ecuador [3] Departamento de Recursos Hídricos y Ciencias Ambientales, and Facultad de Ingeniería, Universidad de Cuenca, Cuenca, Ecuador

Abstract: Information on the spatiotemporal rainfall occurrence, its microphysical characteristics and its reflectivity-rainfall (Z-R) relations required to provide rainfall mapping based on rain radar data is limited for tropical high mountains. Therefore, this study aims to analyze rainfall types in the Andes cordillera to derive different rain-type Z-R relations using disdrometer observations at three study sites representative for different geographic positions and elevations (2610, 3626 and 3773 m a.s.l.). Rain categorization based on mean drop volume diameter (D_m) thresholds [0.1$<$ D_m(mm)$\leqslant$0.5; 0.5$<$ D_m(mm)$\leqslant$1.0; 1.0$<$ D_m(mm)$\leqslant$2.0] was performed using drop size distribution data at a 5-min time step over an approximate 2-yr period at each location. The findings

are as follows: (i) Rain observations characterized by higher(lower) D_m and rain rates are more frequent at the lower(higher) site. (ii) Because of its geographic position, very light rain (drizzle) is more common at higher altitudes with longer-duration events, whereas rainfall is more convective at the lower range. (iii) The specific spatial exposition regarding cloud and rain formation seems to play an important role for the derivation of the local Z-R relationship. (iv) Low A coefficients ($\leqslant$60) for the first rain type resemble typical characteristics of orographic precipitation. (v) Greater values of A (lowest and highest stations for D_m >1.0 mm) are attributed to transitional rainfall as found in other studies. (vi) Rain-type Z-R relations show a better adjustment in comparison with site-specific Z-R relationships. The study is the first contribution of Z-R relations for tropical rainfall in the high Andes.

3.1 Introduction

Mountain ranges have a major impact on global and regional climate. Their topographic complexity strongly affects rainfall processes and this in turn makes quantitative precipitation estimates in space and time very difficult. Unfortunately, operational rainfall monitoring in highlands is scarce and unevenly distributed, and therefore their rainfall dynamics is significantly underexplored in comparison with other areas of the world (Wohl et al., 2012). This situation particularly holds for the tropical Andes, where high spatiotemporal rainfall variability has been demonstrated (Célleri et al., 2007; Espinoza Villar et al., 2009; Pulwarty et al., 1998; Rollenbeck and Bendix, 2011; Scheel et al., 2011) mainly because of complex rainfall formation processes (Poveda et al., 2005). In addition, rainfall microphysical characteristics [e.g., drop size distribution (DSD)] behind its generation are still unknown. Rainfall has countless implications on human activities, water resource management, biological and ecological processes, just to name a few, and a tremendous impact on the socioeconomic development of the Andean region. Irrigated and rain-fed agricultural areas are still very large (De Bièvre et al., 2003) and spatially distributed, and hydropower generation relies on rainfall and its variability (Célleri and Feyen, 2009). In Ecuador, hydropower generation accounts for approximately 45% of energy production (Buytaert et al., 2006; Célleri and Feyen, 2009) and

is projected to reach a 80% by 2020. In addition, the occurrence of extreme rainfall events (e.g., in the scope of the ENSO phenomenon) may cause hazards for populations and human activities. In consequence, knowledge of the spatiotemporal rainfall distribution is of major importance for economic issues. Therefore, reliable rainfall data and their related microphysical characteristics are greatly needed.

Given the difficulties in deploying rain gauges due to the complex topography of mountain regions, the use of rain radars is highly beneficial in providing spatially well-distributed rainfall data (Yoon et al., 2012). Nevertheless, the hydrological visibility as a result of beam blockage is often a limiting factor (Pellarin et al., 2002). Depending on the technology used, radar implementation might be too expensive (e.g., polarimetric systems as used in Next-Generation Radar (NEXRAD); Cunha et al., 2015). Recently, cost-effective scanning X-band radars with one plan position indicator (PPI) have emerged on the market as a good supplement to sparsely distributed rain gauge networks. This is particularly helpful for developing-countries such as Ecuador, where the Cajas X-band Selex (CAXX) single-polarized weather radar (Rainscanner, SELEX ES GmbH) was installed at 4450 m above sea level (a.s.l.) as part of a novel weather radar network developed in tropical high mountains (RadarNet-Sur; Bendix et al., 2017), making it arguably the highest radar in the world.

Quantitative precipitation estimation (QPE) can be extremely difficult and challenging. In this context, converting native radar variables — normally reflectivity (Z) — into reliable rainfall rates (R) is one of the most important steps, but it is particularly difficult in mountain regions (Germann et al., 2006; Rollenbeck and Bendix, 2006, 2011; Savina, 2011). One possible way to overcome this issue is to use an adapted climatic or, even better, a rain-type-based Z-R relation. For the latter, a quantitative rain-type classification is needed. Approaches for classifying rain types generally rely on (i) an application-oriented discrimination based only on rain rates (Cerro et al., 1997; Sen Jaiswal et al., 2009), (ii) DSD information (Atlas et al., 2000; Caracciolo et al., 2008; Nzeukou and Sauvageot, 2004; Ochou et al., 2011; Rosenfeld et al., 1993; Uijlenhoet, 2001; Uijlenhoet et al., 2003), or (iii) polarimetric rainfall retrieval algorithms, as proposed by L'Ecuyer et al. (2004) and Zhang et al. (2011). Unfortunately, most studies in mountain regions have been carried out without fully considering rainfall spatial variability and, consequently, the effects of altitudinal gradients on Z-R relation parameters (Tokay et al., 2009). Moreover, few Z-R relations adapted to tropical lowlands (Bamba et al., 2014; Kumar et al., 2011; Narayana Rao et al., 2001; Ramli and Tahir, 2011; Russell et al., 2010; Tenório et al., 2012) have been provided thus far and may not represent the rainfall characteristics at high altitudes very well. In consequence, neither rainfall microphysical characteristics such as DSD nor climatic and/or rain-type-based Z-R relations are available to date

for tropical high mountains, such as the south Ecuadorian Andes.

With the objective of improving our understanding of tropical rainfall structure and its spatio-temporal variability (Célleri et al., 2007; Rollenbeck et al., 2011) in the Andean mountain range of south Ecuador, in this study we aim to derive Z-R relations for different rainfall types occurring at three terrain elevations with different topographic location characteristics regarding its synoptic forcing. The relations were derived depending on the drop size distribution through a mean volume diameter [D_m (mm)] characterization that may account for tropical rainfall variability at high altitudes.

3.2 Materials and Methods

3.2.1 Study sites

Figure 3.1 shows the CAXX radar situated at the upper limit of the Cajas National Park (2° 45' S, 79° 16' W) on Cerro Paragüillas and extent of its 100-km maximum measurement range. Three disdrometer sites were selected to represent different altitudinal positions in the CAXX radar range and beyond, in the western Andean cordillera. A summary of study sites information is shown in Table 3.1. These sites were located within the Macizo del Cajas, part of the United Nations Educational, Scientific and Cultural Organization's (UNESCO's) World Network of Biosphere Reserves.

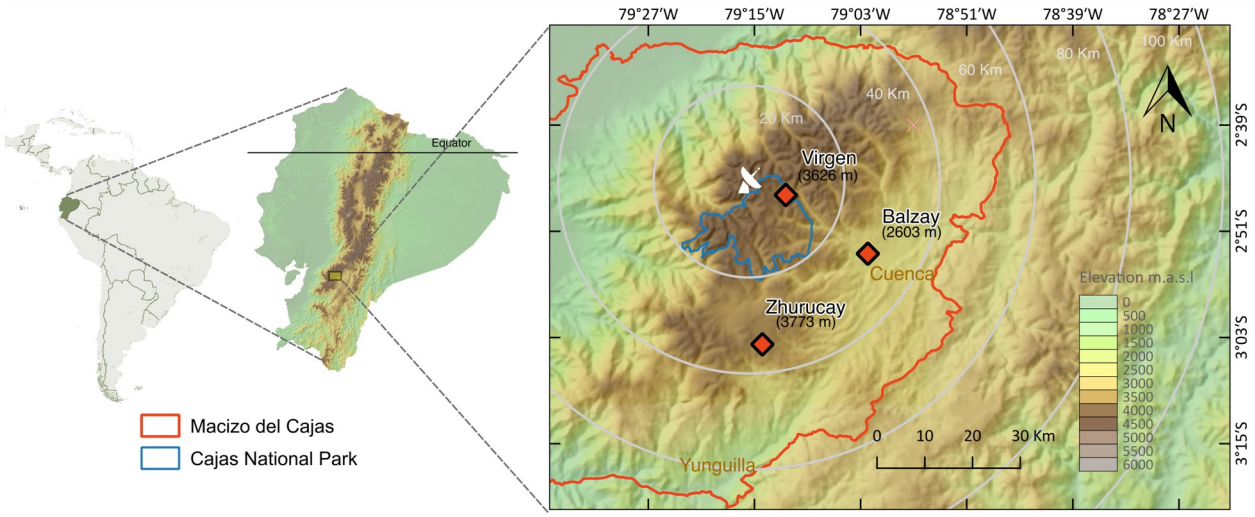

Figure 3.1: Map of study area within the South Ecuadorian Andean region.

Two of the disdrometer study sites, Balzay and Virgen, were located along an altitudinal gradient within the Quinuas Ecohydrological Observatory. The lowest site is the Balzay meteorological station, at 2610 m a.s.l., which is in the Cuenca city outskirts. Climate at Balzay is characterized by a mean annual precipitation of 969 mm and a mean temperature of ∼14°C (Córdova et al., 2016).

The second station, Virgen, is located at 3626 m a.s.l. in a shrub-dominated (sub-páramo) transition zone from upland forest to páramo (Hofstede et al., 2014) within a basin at a maximum elevation of 4400 m a.s.l. with slopes that drain toward the east (i.e., Atlantic Ocean). Average annual precipitation in this area is approximately 998 mm, average air temperature is 6.5°C and relative humidity is ∼90%. For a more detailed climatic characterization of the site, the reader is referred to Carrillo-Rojas et al. (2016) and Córdova et al. (2016).

The third and highest disdrometer site is situated in the páramo ecosystem at an elevation of 3773 m a.s.l. in the Zhurucay Ecohydrological Observatory within a basin at a maximum altitude of 3900 m a.s.l. on the slopes draining to the west (i.e., the Pacific Ocean). Climatologic conditions are characterized by an average temperature of 6°C, a relative humidity of 91% (Córdova et al., 2015), and annual precipitation ranging from 900 to 1600 mm with weak seasonality (Padrón, 2013). The climatology of the area is influenced by the Pacific coastal regime from the West, and continental and tropical Amazon air masses from the east (Vuille et al., 2000). As a consequence, prevalent convective and orographic cloud formation occurs (Bendix et al., 2005). Close-to-freezing temperatures cause moist air condensation resulting in drizzle that accounts for approximately 30% of annual rainfall volume in the area (Padrón et al., 2015).

3.2.2 Instruments and database

The instruments used for this study were two high-resolution disdrometers (Thies Clima laser precipitation monitor (LPM); Thies Clima, 2007) based on a laser sensor that produces a horizontal light strip. When a precipitation particle falls through the light beam, the receiving signal is attenuated. The particle diameter is calculated from the reduction in amplitude. Moreover, the fall speed of the particle is determined from the duration of the reduced signal. The sensor operates at a wavelength of 785 nm and has a reference measurement area of 45.6 cm^2 with a resolution of 0.005 mm h^{-1}. Each drop is assigned to one of 22 size bins and one of 20 velocity bins ranging from 0.125-8 mm in diameter and 0-10 m s^{-1}, respectively. The range of size and velocity classes is finer for smaller and slower particles and coarser for larger and faster ones. A data telegram is transmitted every 60s indicating the numbers of drops detected in each of the total 440 (22x20)

different possible classes. Assessments of the Thies optical disdrometer have been carried out by Frasson et al. (2011) and Sarkar et al. (2015), and many other studies have used this instrument (e.g., Chen et al., 2016; Fernández-Raga et al., 2010; Jameson et al., 2015). Despite its very good performance, the instrument is known to show some difficulties at very high wind speeds (Frasson et al., 2011) which, however, were not present at our sites.

Table 3.1: Summary of study sites and databases.

Station	Lat.	Lon.	Altitude [m]	Data period	Total obs.
Balzay	2°53' S	79°02' W	2610	03.2015-08.2016	5451
Virgen	2°46' S	79°11' W	3626	03.2015-08.2016	17096
Zhurucay	3°03' S	79°14' W	3773	11.2012-08.2014	13582

The first LPM disdrometer (LPM-1) was located at Zhurucay station from 2012 to 2014. After, it was moved to Balzay station while a second Thies disdrometer (LPM-2) was acquired. The LPM-2 was calibrated next to LPM-1 for a one-month period and installed after at the Virgen site. Data used in this study were collected at 5-min sampling resolution over different time periods. Because of the data integration time, a result similar to sequential intensity filtering technique (SIFT) proposed by Lee and Zawadzki (2003) is obtained to overcome the drop sorting effect on the disdrometer database. Data from March 2015 to August 2016 were available at Balzay, from April 2015 to August 2016 at Virgen, and from November 2012 to August 2014 at Zhurucay. Altogether, a total of 8142, 25804, and 18019 sampling observations were obtained at Balzay, Virgen and Zhurucay stations, respectively. A summary of available data is provided in Table 3.1.

3.2.3 Methods

We implemented our methodology systematically as follows (Figure 3.2). First, we derived the rain rate and reflectivity from DSD observations. Then, we split the available disdrometer data from each study site into rainfall events by defining properties of a time series subset. A detailed explanation of these properties is described below [section 3.2.3.3]. However, because of each study area's rainfall variability, we selected a rainfall categorization based on the mean volume diameter (D_m). After we identified the events and classified the 5-min rainfall observations, we calculated the A and b parameters of the power law $Z=AR^b$ for each categorized observation subset. Later, we evaluated predicted rain rates R from reflectivity Z values and type-specific A and b parameter

estimates. Finally, we validated the rain-type based Z-R relations through statistical measures of goodness of fit between observed and predicted rain rates. The complete data processing chain is shown in Figure 3.2 and a more detailed explanation is provided in the subsequent sections.

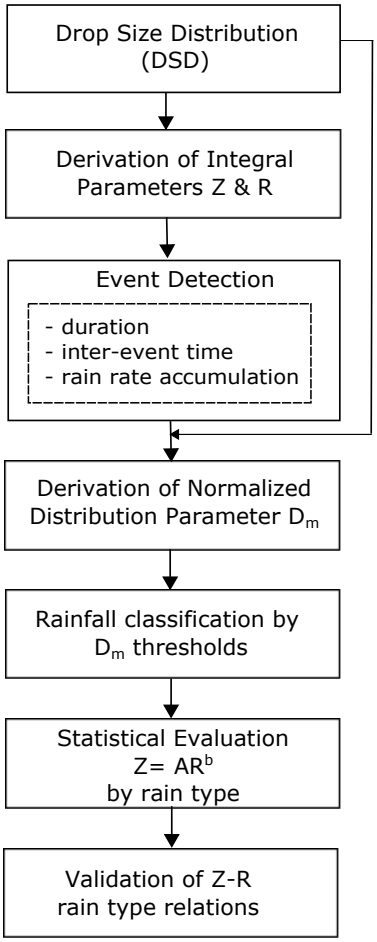

Figure 3.2: Workflow of data processing to obtain rain type *Z-R* relations.

3.2.3.1 Raindrop size distribution and rainfall integral parameters

Integral parameters such as rain rate (R) in (mm h^{-1}) and reflectivity (Z) in (mm^6 m^{-3}; or dBZ= $10\log_{10}(Z)$) are defined as

$$R = \frac{6\pi}{10^4} \int D^3 v(D) N(D) dD \tag{3.1}$$

$$Z = \int D^6 N(D) dD \tag{3.2}$$

where D is the raindrop diameter (mm), $N(D)dD$ is the number of drops per unit volume (mm^{-3}) with diameters between D and $D + dD$, and $v(D)$ is the terminal raindrop fall velocity (m s^{-1}). The number concentration of raindrops per unit volume per unit size $N(D)$ was calculated from the Thies disdrometer counts according to

$$N(D_i) = \sum_{j=1}^{20} \frac{n_{ij}}{AtV_j \Delta D_i} \tag{3.3}$$

here, $N(D_i)$ is the drop size distribution in ith size class (m^{-3} mm^{-1}), D_i is the midsize diameter of the ith class (mm), A is the cross sectional area of the sensor (m^2), t is the measuring time (seconds), n_{ij} is the number of drops within the ith size and jth velocity class, and V_j is the fall speed of the jth velocity class (m s^{-1}) in agreement with Beard (1976). Thus the finite difference equivalences for R (Eq.3.1) and Z (Eq.3.2) were derived from the distribution of all particles over class binning as follows:

$$R = \frac{6\pi}{10^4} \sum_{i=1}^{22} \sum_{j=1}^{20} \frac{n_{ij}}{At} D_i^3 \tag{3.4}$$

$$Z = \sum_{i=1}^{22} \sum_{j=1}^{20} \frac{n_{ij}}{AtV_j} D_i^6 \tag{3.5}$$

Usually for long wavelengths λ (e.g., C-band radars) the Rayleigh approximation is used (as in Eq.3.2), where the effective backscattering cross section of the raindrop is roughly proportional to the sixth power of its diameter. X-band related studies in literature have used both Rayleigh (Lo Conti et al., 2015) and Mie regimes (Delrieu et al., 1997; van de Beek et al., 2009). Nethertheless, several investigations (e.g., Dolan and Rutledge, 2009; Löffler-Mang et al., 1999; Maki et al., 2005; Tokay et al., 2001; Wexler and Atlas, 1963) documented the impact on the derivation of the Z-R relationship parameters due the scattering regime used and highlight the importance of considering the Mie scattering in short wavelenghts such as X-band. Thus, for the purposes of this study, reflectivity was calculated in both Rayleigh and Mie scattering regimes in order to learn about the effect of the drop size spectra on the subsequent Z-R relationship derivation. For this, the effective backscattering cross-section was calculated using the implementation in Matlab by Mätzler (2002).

A detailed explanation of the Mie theory is well described in van de Hulst (1957) and Horvath (2009).

Observed DSDs are commonly represented by theoretical distribution approximations. Two models for the drop size distribution were used to describe and compare the observed DSDs at the study sites: (i) The gamma drop size distribution (Ulbrich, 1983) in the form of $N(D) = N_0 D^{\mu} exp(-\Lambda D)$ where D is the rain drop diameter and N_0, μ and Λ are the intercept, shape and slope parameters respectively. (ii) The Marshall-Palmer (Marshall and Palmer, 1948) distribution defined as $N(D) = N_0 exp(-\Lambda D)$, where N_0= 8000 m^{-3} mm^{-1} and Λ= $41R^{-0.21}$. For the former, the truncated-moment method (Ulbrich and Atlas, 1998) was used for the estimation of the three parameters.

3.2.3.2 Quality control

In a recent study, Padrón et al. (2015) compared the LPM-1 with a corresponding adjacent 0.1-mm-resolution rain gauge at different rain-rate intensities and velocities. Because of the sensitivity of the disdrometer to small drops, differences occurred mainly during low-intensity events, where the rain gauge produced underestimations of approximately 15%. Therefore, we base our data quality on the detailed study of Padrón et al. (2015) regarding LPM-1 operation, which was used afterwards to perform the LPM-2 intercalibration.

In addition, we applied the quality control procedure as described in Friedrich et al. (2013) in order to remove measurement inaccuracies related to: (i) misclassification of particles, (ii) margin fallers and (iii) splashing effects. For a detailed explanation of this procedure the reader may refer to Friedrich et al. (2013). The particle classification scheme (Figure B1 in Friedrich et al., 2013) was adapted by considering changes in fall velocity on high-altitude pressure conditions (Beard, 1976). This quality control process confirmed the very low presence of solid precipitation (i.e., hail, snow, graupel). In our dataset, only four observations showed solid precipitation at the lowest site (Balzay). Neither Virgen nor Zhurucay presented solid precipitation observations. Besides, we compared time series of 30-min rain totals of a collocated Texas Electronics, Inc., TE-525 (tipping-bucket rain gauge (0.1-mm resolution) at each location with its corresponding disdrometer data for the study period. We found correlation coefficients of 0.93, 0.96 and 0.94 for Balzay, Virgen and Zhurucay, respectively. As expected systematic erros were found. Relative bias was 5% (Zhurucay), 18% (Virgen) and -6% (Balzay). It can be largely explained by the climatology at all sites and their rainfall regimes. The differences in relative bias are related to the differences in measurements of

very small (high) rain rates Virgen (Balzay) and a mixture of both at Zhurucay. It should be stressed that we only retained observations with more than 100 droplets and a rain rate higher than 0.1 mm h^{-1} for further analyses to disregard noisy observations. Furthermore, regular maintenance was performed on both LPMs to avoid contamination of the light beam.

3.2.3.3 Rainfall Event Detection

To obtain representative data that accounts for the rainfall regime of each study site while avoiding isolated observations, we first identified rainfall events that occurred during the observation periods. In this context, Dunkerley (2008) provided a well-documented review of rain event properties found in the literature. He concluded that they differ widely from one study to another (e.g., event durations from 3min to 24h were reported). Unfortunately, no standard methodology for rain event selection exists. Nevertheless, the author strongly recommends a more detailed reporting of guidelines used in investigations related to rain events. According to Dunkerley (2008), a rain event is identified by a continuous time interval of rain during which there are no rainless gaps with a duration exceeding the minimum inter-event time. Usually this definition is augmented with some additional criteria that any rain event must meet (e.g., minimum total accumulation and minimum duration). In this study, we used the following three criteria to extract events from 5-min rain-rate time series. (i) Minimum duration: keeping in mind that most convective events are short duration, a rain event is characterized by at least two continuous single rainy records. (ii) Minimum interevent time: we performed several experiments modifying this time criterion while, in general, the results remained unchanged. Therefore, the inter-event time was selected according to previous rainfall disdrometer-related studies (Larsen and Teves, 2015; Padrón et al., 2015; Tokay et al., 2003) that define half an hour — 30 min interevent — as the minimum time to separate one event from another. (iii) Minimum rain-rate accumulation: this value was fixed to 5 mm, providing a good trade-off by reducing the bias of short and extremely light rain events.

3.2.3.4 Rainfall Class Selection

Observations of selected rainfall events were classified according to an adequate tailor-made categorization that accounts for rainfall variability at each site. Some studies have used the most common classification schemes based on nonoverlapping rain-rate (Llasat, 2001; Sen Jaiswal et al., 2009; Varikoden et al., 2011) or reflectivity thresholds. However, the major drawback of these methods is that they do not consider DSD variability of rain (i.e., different DSDs can arise at similar

rain rates). This is the case for drizzle rain type in the Andean highlands, whose occurrence and importance is noted by Padrón et al. (2015), Rollenbeck and Bendix (2011) and Muñoz et al. (2016). Other studies use a DSD parameterization (i.e., assuming a gamma or an exponential distribution) to obtain a rainfall separation criterion as in Tokay and Short (1996). Thus, to objectively compare and account for local rainfall characteristics in the study areas, we selected a rainfall categorization using the mean volume diameter D_m(mm).

$$D_m = \frac{\int_0^\infty N(D)D^4 dD}{\int_0^\infty N(D)D^3 dD} \tag{3.6}$$

defined by Testud et al. (2001) from a normalized distribution, where $N(D)$ is the drop size distribution and D is the particle diameter. For each spectrum, the third and fourth moments are calculated and D_m is subsequently determined. Finally, we defined four rainfall categories according to D_m, which may explain rainfall structure differences: 1) 0.1$<$ D_m (mm)$\leq$0.5; 2) 0.5$<$ D_m (mm)$\leq$1.0; 3) 1.0$<$ D_m (mm)$\leq$2.0; 4) D_m (mm)$>$2.0. This tailor-made classification was based on the findings of Uijlenhoet et al. (2003) in northern Mississippi that showed that mean raindrop sizes $>$1.0 mm characterized convective rainfall, while transitional and stratiform rainfall was mainly associated by mean raindrop sizes of 0.6-0.9 mm. The classification was also supported by the observation of the local D_m frequencies at all study sites as shown in section 3.3.1.

3.2.3.5 Z-R Relation Calculation

After classifying individual rain observations, we calculated the Z-R relation (A and b coefficients) from the classified observation rain rates and corresponding Z values provided by the instruments. For this, we used the well-known Z-R power law relationship defined by Marshall and Palmer (1948) for rain rate estimation based on radar reflectivity records.

$$Z = AR^b \tag{3.7}$$

where Z(mm^6 m^{-3}) is the reflectivity value, R(mm h^{-1}) is rain rate, A and b are empirical parameters and are referred to as coefficient and exponent, respectively. The commonly adopted method of a linear regression of $log(R)$ on $log(Z)$ has been demonstrated to lead to suboptimal parameters (Alfieri et al., 2010). Thus, we performed a nonlinear (power law) regression by fitting least squares with Z as the independent variable and compared both regression approaches in similar

fashion with van de Beek et al. (2016). Values of A and b were calculated by applying Eq. 3.7 to all rainy records (Z,R) as well as to the subset of each rainfall category independently.

3.2.4 Validation of rain type Z-R relations

The validation of rain-type Z-R relations was performed over the 5-min observations corresponding to 20% of rainfall events. These observations were not used for the Z-R parameter derivation process and covered the entire Z reflectivity range. The performance of our type-specific Z-R relations for each study site was assessed using statistical measures of goodness of fit between observed and predicted rain rates. Thus, we calculated the predicted rain rates (R) through the Z-R relation using the obtained A and b parameters per rainfall category and observed reflectivity (Z) records of the 5-min sampling time series as the independent variable. We applied the formula to Z according to the respective class. Statistical measures derived were the coefficient of determination r^2, Root Mean Squared Error (RMSE; mm h^{-1}) and bias (mm h^{-1}).

3.3 Results

3.3.1 Event detection and 5-min rainfall categorization

We found 488, 594 and 887 rainfall events at Balzay, Virgen and Zhurucay, respectively, that met the rainfall event criteria previously defined. They correspond to a total of 5451, 17096, and 13582 sampling observations. Figure 3.3 illustrates the frequency distributions of rainfall event duration, rain rate and D_m for each study site. While the event duration has a minimum of 10 min at all sites, it increases to 745 min (Balzay), 2480 min (Virgen) and 750 min (Zhurucay). We also found that mean event duration varies between stations, yielding 180 min at Balzay, 470 min at Virgen and 225 min at Zhurucay (Figure 3.3), altogether pointing to longer-duration events at higher altitudes. Approximately half of the events are shorter than 125, 335 and 170 minutes.

From those rainfall observations, only 129 (Balzay), 34 (Virgen) and 115 (Zhurucay) rainfall observations were categorized as D_m >2.0. Thus, we disregarded this rainfall category from our study due to lack of representation. Relative frequencies of rainfall categories are illustrated in Figure 3.4. Observations categorized with the smallest D_m are more frequent at the higher sites, Zhurucay and Virgen, than at the lowest site, Balzay, which may be related to more common drizzle rainfall at higher altitudes. The next rain category [0.5< D_m(mm)≤1.0] seems to occur at slightly

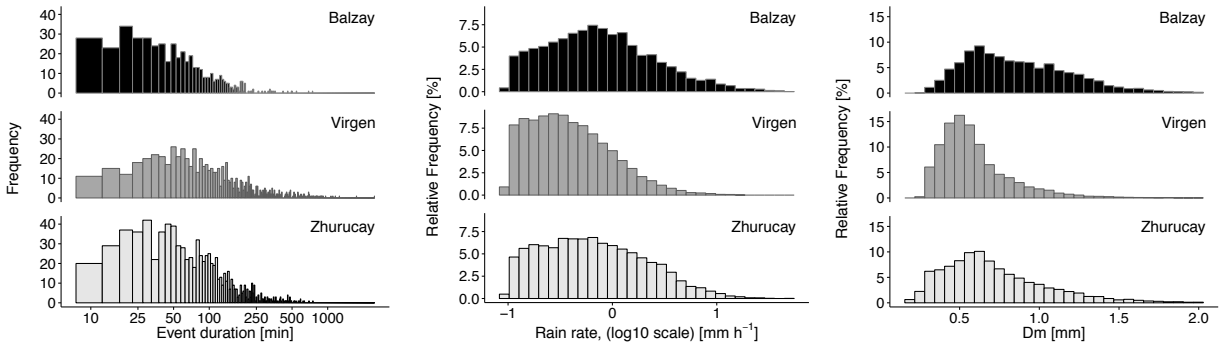

Figure 3.3: Distributions of rainfall event duration, rain rate and mean volume diameter D_m for all disdrometer study sites.

different relative frequencies at all sites. In contrast, rain observations characterized by bigger droplets [1.0$<$ D_m(mm)$\leq$2.0], most related to moderate-heavy rain, are high in number at Balzay and very uncommon at Virgen.

Figure 3.5 illustrates the distribution of rainfall categories at the study sites. The upper limit of rain rates produced by the lowest range of D_m [0.1$<$ D_m(mm)$\leq$0.5] is relatively low, with values of approximately 1.1 mm h^{-1} at Virgen and 3.2 mm h^{-1} at Zhurucay. On the other hand, few low rain rates are characterized by the biggest D_m [1.0$<$ D_m(mm)$\leq$2.0] at all locations in comparison with the preceding classes. At first glance, Figure 3.5 appears to hinder real discrimination between classes. In Figure 3.6 , however, it is shown to what extent the rain categories are overlapping: rain rate and reflectivity systematically increase according to rain type. It should be noticed that the distributions, although roughly similar at a specific rain class, correspond to a different number of observations at each site. Median values of both rain rate and reflectivity are lower for the rain types 0.1$<$ D_m(mm)$\leq$0.5 and 0.5$<$ D_m(mm)$\leq$1.0 at Virgen in comparison to the other sites.

3.3.2 Drop size distribution

The drop size distribution of three different rain rates (1, 10 and 15 mm h^{-1}) is shown in Figure 3.7 for all sites. At the highest rain rate, the observed DSD at the lowest station (Balzay) begins to exceed that of the high-altitude stations, revealing rain-rate-droplet number combinations not observed at higher altitudes. On the other hand, at the highest station (Zhurucay), droplets bigger than 4 mm are registered, which are not present at the other sites. This points to extended droplet growth at lower altitudes leading to higher rain rates at the expense of smaller droplets. The theoretical model of Marshall-Palmer better fits the data at the highest station, Zhurucay, while

Virgen consistently presents a lower number of droplets per diameter. The gamma model (Ulbrich, 1983) seems to be suitable for the lower stations Virgen and Balzay. Figure 3.8 illustrates the average raindrop size spectra at all study sites. The more isolated sub-páramo station, Virgen, shows a lower droplet concentration in comparison with the other sites (Figure 3.8a). When data are divided according the D_m classification, it can be observed that the drop size spectrum is more similar within classes (Figure 3.8b) in comparison with the rain-rate discrimination (Figure 3.7), particularly at low droplet diameters.

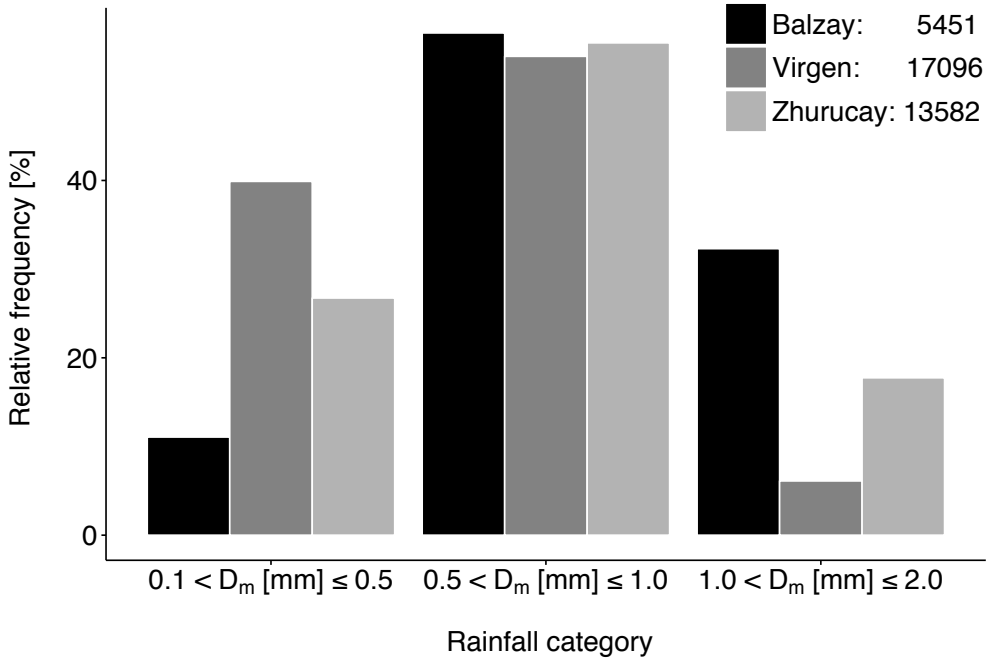

Figure 3.4: Relative frequency of mean volume diameter (D_m) in 5-min interval observations.

3.3.3 Estimation of Z-R parameters

Once we identified the rainfall events, we calculated parameters A and b of the Z-R relation (Eq. 3.7) for all observations at each site. For illustration purposes, Figure 3.9 shows the differences that arise from applying a linear log(R) on log(Z) and nonlinear (power law) regression on the calculation of Z-R parameters as well as the calculation of reflectivity (Z) by using both scattering regimes, Mie and Rayleigh. It can be seen that the nonlinear approach outperforms its counterpart and also departs from the 1-by-1 line at the lowest rain rates. It seems to be an effect of the predominance of least squares fitting on higher rain rates. The scattering regime calculation leads to

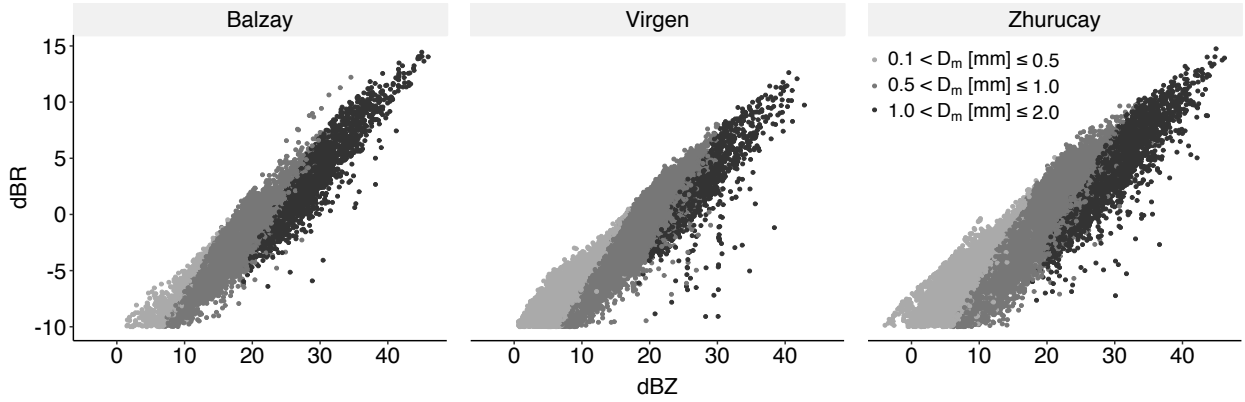

Figure 3.5: Classification of 5-min rainfall observations based on three bins of D_m related to radar reflectivity; dBR=10log$_{10}$R where R is in mm h^{-1} and dBZ=10log$_{10}$Z where Z is in mm^6 m^{-3}.

slight differences in terms of goodness-of-fit statistics. Nevertheless, as only the big droplets are affected by Mie scattering in X-band, as expected, the greatest difference arises from the station with the most frequent high D_m (Balzay) which deviates mostly between Mie and Rayleigh regimes. The opposite occurs in the sub-páramo station (Virgen) where the lower D_m mostly occur. Thus, the Mie scattering regime is used in this article hereinafter. Parameters on these site-specific Z-R relations were similar in terms of exponent b values (2.05<b<2.25); however, coefficient A at the highest station Zhurucay (76) was lower than at the other sites Balzay (103) and Virgen (104).

Afterward, we calculated parameters A and b of the Z-R relation (Eq. 3.7) for each subset of 5-min rainfall observations that belongs to the corresponding rainfall category (Figure 3.10). Generally, the coefficient and exponent increase through the rain categories. The Z-R relation becomes more nonlinear at higher D_m values, namely, moderate-heavy rainfall. Balzay shows the higher (lower) coefficients (exponents) for each rainfall category. Hereafter, rainfall types are referred indistinctly by terms of light and heavy rain, which are directly associated with the D_m range: lower D_m values are related to light rain and higher D_m values are related to heavy rain.

3.3.4 Evaluation of Z-R parameters

For all observed reflectivity values, we calculated the rain rate based on the Z-R relations found per rainfall classification by means of the nonlinear (power law) regression with Z calculated in the Mie scattering regime. The goodness of fit between observed and predicted rain rates is illustrated in Figure 3.10. In addition to the coefficient of determination r^2, statistical values of RMSE and bias are shown. We found that the lowest station (Balzay) showed the highest r^2 in all rain categories.

Rain rates characterized by the highest range of D_m [1.0< D_m(mm)≤2.0] are slightly overestimated (i.e., bias=0.187) at Balzay and underestimated (i.e., bias=-0.198) at Virgen. Finally, we found a stable high $r^2 \sim$ 0.8-0.9 for most rain categories for all sites; however, the lower r^2=0.728 was found at Zhurucay for the lowest D_m range [1.0< D_m(mm)≤2.0].

3.4 Analyses and Discussion

In light of the results, there is clearly a different contribution of rainfall events with 0.1< D_m(mm)≤0.5 and 1.0< D_m(mm)≤2.0 through the study sites, which is directly related to their altitudes and type of cloud formation. Our findings indicate the presence of significant amounts of drizzle, characterized by small droplets, particularly in the high and sub-páramo (stations Zhurucay and Virgen). It is worthwhile to mention that had all rain-rate observations been considered, drizzle occurrence would still have been greater than rain. However, after the threshold used as quality control for rain rate (i.e. $R >$0.1 mm h^{-1}); the number of lower D_m observations decreased remarkably. Low A coefficients in rain type 0.1< D_m(mm)≤0.5 are in good agreement with orographic rain Z-R relations ($Z = 16.6R^{1.58}$, $Z = 31R^{1.71}$) found at high altitudes in Hawaii (Blanchard, 1953).

A noticeable difference in the rainfall regime with largest D_m is observed between the three sites. Heavy rainfall events, here associated with 1.0< D_m(mm)≤2.0, mostly occur at the lower elevation (Balzay) while, in contrast, higher elevation sites have a lower relative frequency of heavy rain. Although Virgen and Zhurucay have similar altitudes, they differ mainly in the A coefficient of the corresponding Z-R relation. In particular, Zhurucay is associated with a higher reflectivity at a given rain rate because of a higher droplet concentration. Both sites are located in different areas of the western Andes Cordillera, with Virgen located on the eastern draining slopes (i.e., toward the Atlantic Ocean) and Zhurucay on the western draining slopes (i.e., toward the Pacific Ocean). Zhurucay is also closer to the lower areas of the adjacent Yunguilla Valley. Advection of moist air masses from the coastal Pacific lowlands through this valley, local orographic uplift and stronger thermal upslope breezes might affect rain formation processes in this area.

As expected, DSDs according to rain rates (Figure 3.7) are diverse at all locations and clearly divergent at a particular rain rate. Differences in drop size and number concentration at all sites (Figure 3.8a) are more evident in the Dm categorization (Figure 3.8b). From the microphysical interpretation of Z-R relations (Figure 3 in Steiner et al., 2004), there is no evidence of number-controlled (i.e., D_m constant) or size-controlled (i.e., drop number density constant) rain types. Thus, it is suggested that rainfall types are controlled by variations of mean drop size and number

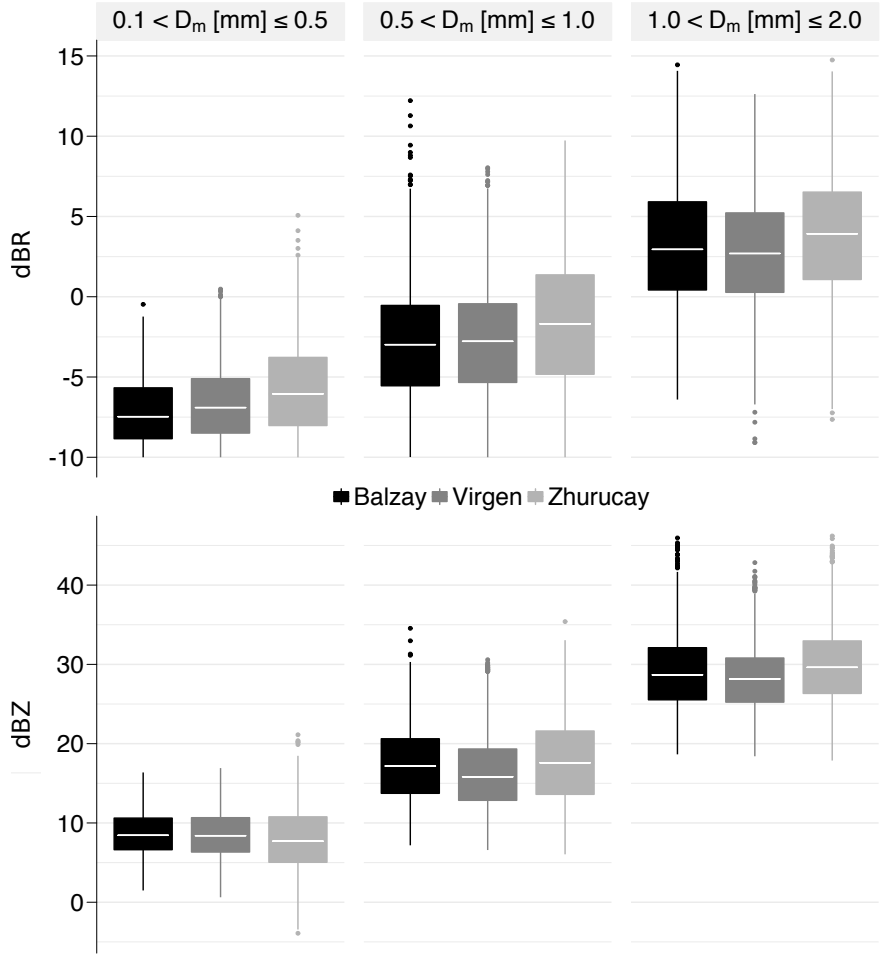

Figure 3.6: Distribution of rain rate and reflectivity values associated with each rainfall category according to D_m binning.

concentration. Regarding the overall and rain-type Z-R relations, the latter clearly outperform the adjustment at all sites (e.g. at Balzay, r^2 increases from 0.729 to 0.867 , RMSE decreases from 1.609 to 0.738 mm h^{-1}, and bias decreases from 0.286 to 0.027 mm h^{-1}). Thus, variations of coefficients and exponents of Z-R relations through the different study sites and rain categories suggest that not only the altitude but also the specific spatial exposition regarding cloud and rain formation seem to play a key role in determining local Z-R relationships.

For comparison purposes, Table 3.2 shows a summary of similar Z-R relation studies performed at equatorial and high altitude locations. It should be stressed that these investigations are diverse in the use of methods to fit Z-R relations to DSD data, which are not described in all of the studies.

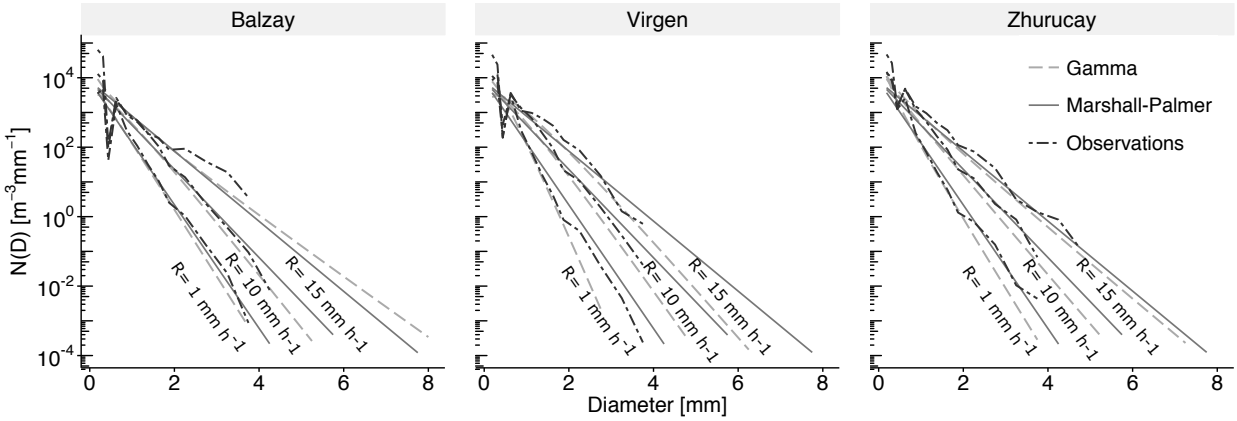

Figure 3.7: Average raindrop size spectra at all study sites for three different rain rates: R=1, 10, 15 mm h^{-1}. Fitting curves of gamma and Marshall-Palmer distributions are also shown.

Our purpose henceforth is to contextualize our results with those of remarkable proximity, similar geographical location and/or influence of complex terrain. For instance, the study conducted by Germann et al. (2006) in an alpine mountainous region only provides a single $Z = 316R^{1.5}$ relation. We obtained lower coefficients and higher exponents for all rainfall categories than the Z-R relation reported there. A single/unique Z-R relation would not consider rainfall variability in the Andean cordillera.

The exponents found in this study are roughly similar within each rain category at all sites (1.27-1.44, 1.34-1.57, and 1.57-1.68). The lowest values per rain-type class were obtained at Balzay. Uijlenhoet et al. (2003) also found that exponent b varied slightly through the rain classes. However, coefficients increase in a different order in our study (i.e., from smaller to larger D_m) and they are also of lower magnitude. This is most likely due to the wide range of rain rates obtained in our results for each rain category, in contrast to other studies that use rain-rate thresholds as well as intra-event mean drop size variability (Narayana Rao et al., 2001; Uijlenhoet et al., 2003). Nevertheless using only the D_m parameter as a rain classification criterion, evaluation on independent data shows good measurements of goodness of fit (r^2 >0.8, RMSE<1.7 mm h^{-1} and bias<0.2 mm h^{-1}) relative to those reported by Ramli and Tahir (2011) in a low-altitude tropical climate. According to our validation, RMSE values vary from 0.038 to 1.656, whereas these values increase from 3.65 to 16.97 in Ramli's study. A similar behavior was presented in the bias statistic. These results indicate that our study is reliable in terms of accuracy of predicted rain rate based on rain-type Z-R relations. Furthermore, indications of an upper limit of 2.0 mm for warm tropical rain, as suggested by Atlas

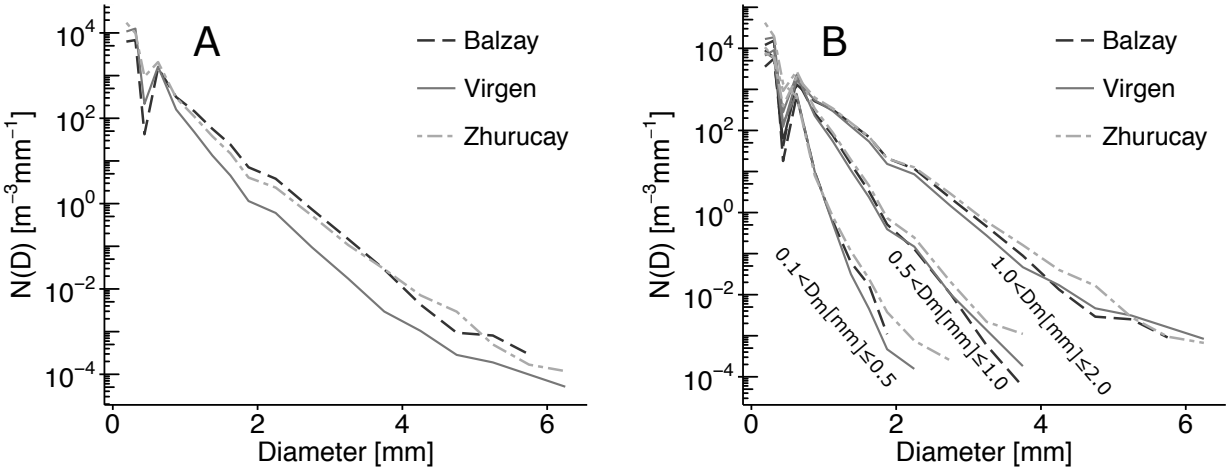

Figure 3.8: Average raindrop size spectra at all study sites, (a) considering the entire observations dataset and (b) considering observations according to the Dm binning: $0.1< D_m \leqslant 0.5$, $0.5< D_m \leqslant 1.0$ and $1.0< D_m \leqslant 2.0$ (mm)

et al. (2000), are in agreement with this study.

There are few studies that have reported similar low magnitudes of A and b (Atlas et al., 2000; Rosenfeld et al., 1993) in lowland tropical equatorial regions (Table 3.2). Those from American equatorial areas seem to be more consistent with our findings. The lower coefficient values ($196< A <230$) reported by Bamba et al. (2014) correspond to the geographically nearest station, located in French Guyana, with respect to our study sites. These values are in better agreement with those calculated in this study for the highest D_m class ($175< A <205$), mainly related to transitional rain types with lower coefficients. The exponents reported by Bamba are lower (1.34 and 1.35) than those calculated in this study ($1.57\leqslant b \leqslant 1.68$) which may compensate differences in the coefficients between both investigations. Interestingly, Tenório et al. (2012) found lower exponent values (1.27 and 1.28) in Northeast Brazil than those usually reported for other equatorial zones. In contrast, our exponent b values are higher in general ($b \sim 1.6$) but associated with lower A coefficient values. We presume that higher values found in the literature (Table 3.2) mainly characterize the monsoon nature, associated with bigger droplets and higher rain-rate intensities, of the West African (Bamba et al., 2014; Russell et al., 2010) and Asia-Australian (Kumar et al., 2011; Narayana Rao et al., 2001) rainfall that would differ from the South American equatorial zone.

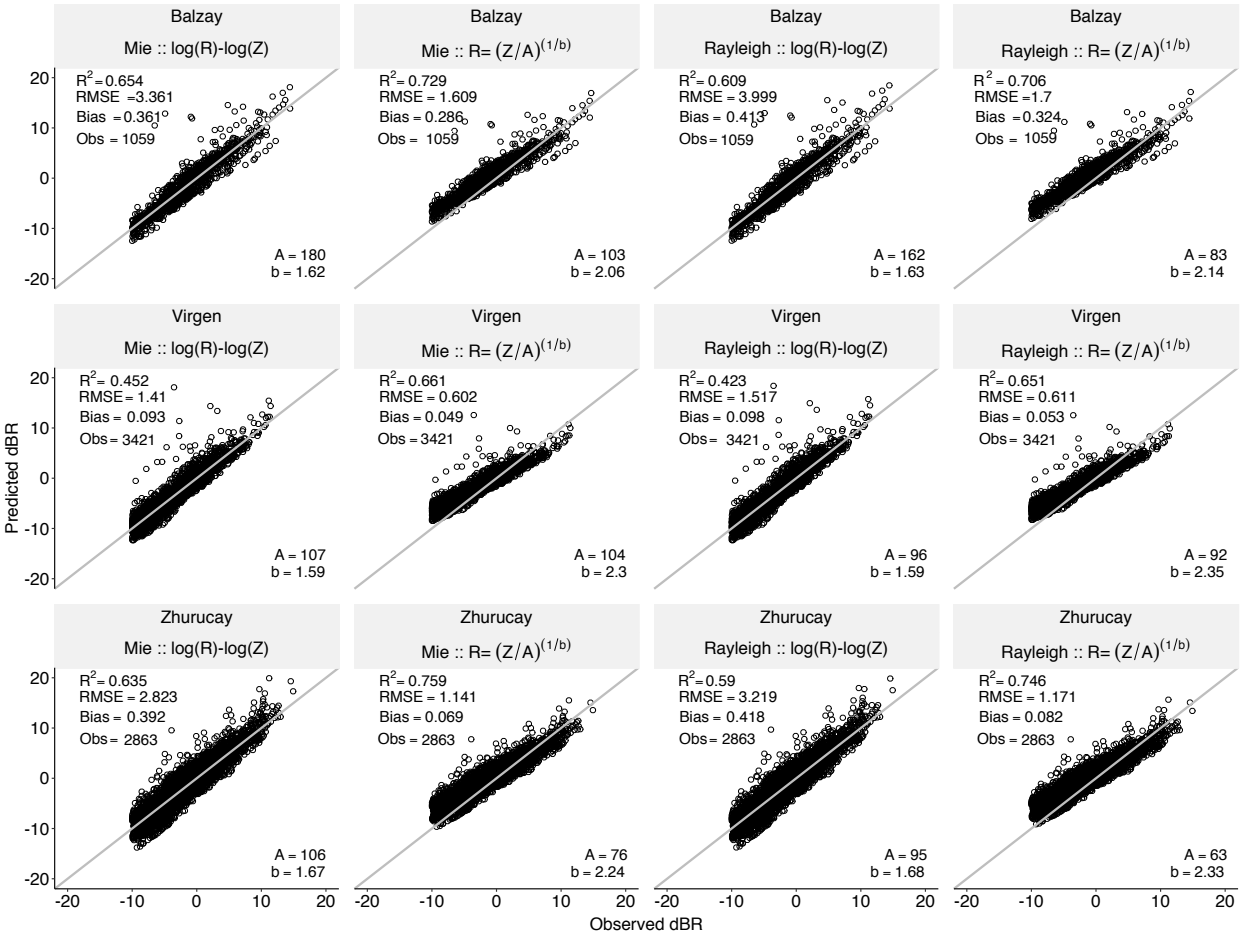

Figure 3.9: Z-R parameters derivation by considering Mie and Rayleigh scattering regime calculation for Z, as well as two regression methods: linear fitting on logarithmic transformation of R and Z [logR-logZ] and nonlinear (power law) fitting [R=(Z/A)$^{1/b}$]. No classification has been performed. Correlation of observed 5-min time step rain rate measurements and predicted rain rate values on the independent validation dataset is illustrated. The bisector line is shown in grey color for all stations.

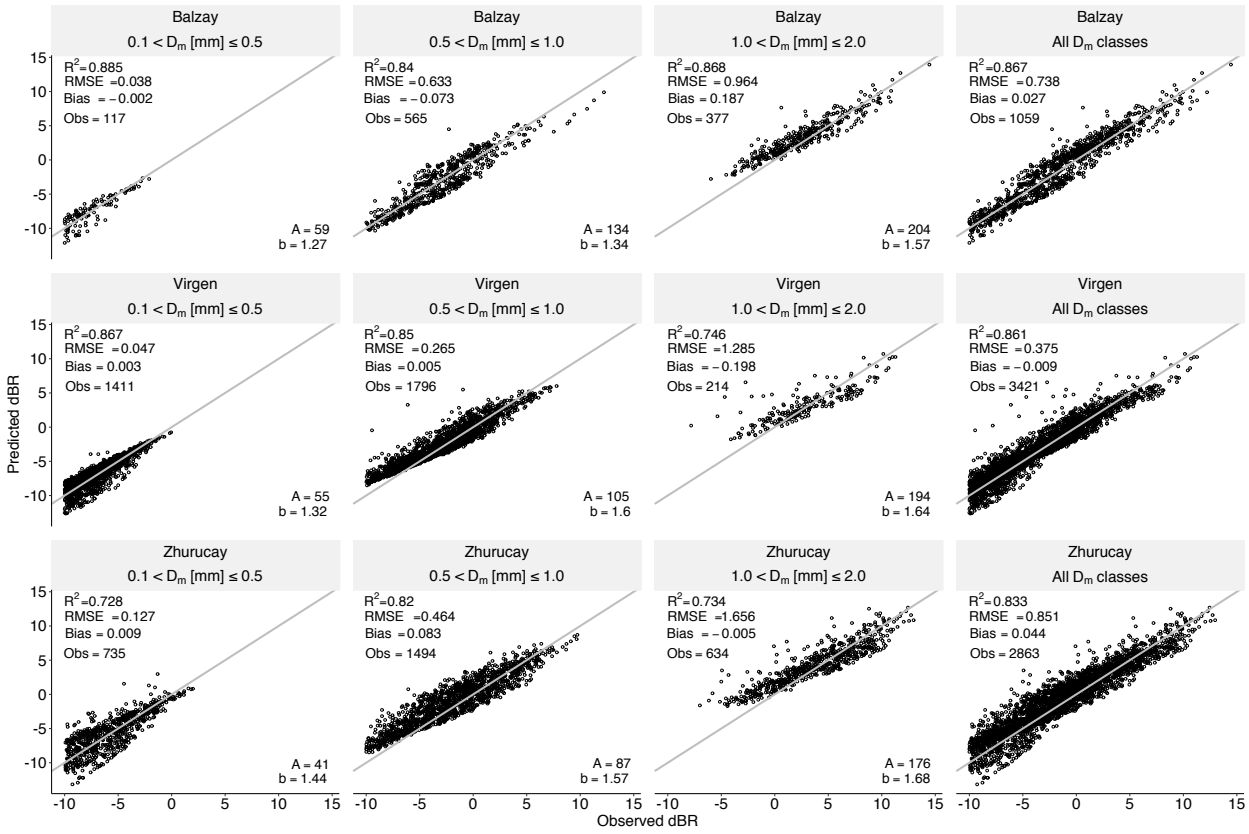

Figure 3.10: Similar to Figure 3.9, but for derivation per rainfall class by considering Mie scattering regime calculation for Z and nonlinear (power law) regression. The given statistics apply to the respective classes.

Table 3.2: Summary of equatorial and high altitude studies of *Z-R* relationships.

Author	Location	Classification	*A*	*b*
Germann et al. (2006)	Switzerland	-	316	1.5
Nzeukou and Sauvageot (2004)	West Africa (Dakar)	-	368	1.24
Russell et al. (2010)	West Africa (Niger)	Convective	21-471	0.98-1.71
		Stratiform	171-330	0.97-1.8
Ochou et al. (2011)	West-Africa (4 countries)	Stratiform to thunderstorms.	319-420	1.28-132
Bamba et al. (2014)	West-Africa (7 countries)	Stratiform and thunderstorms	300-509	1.26-1.42
	America (French Guyana)	Thunderstorms	196	1.35
	America(French Guyana)	Stratiform	230	1.34
Tenório et al. (2012)	Northeast Brazil	maritime rain	146	1.28
		continental rain	256	1.27
Uijlenhoet et al. (2003)	Mississipi	Convective	≈300	≈1.4
		Transition	≈200	≈1.5
		Stratiform	≈450	≈1.4
Narayana Rao et al. (2001)	India	Convective	178	1.51
		Transition	162	1.44
		Stratiform	251	1.48
		NE monsoon	155	1.39
		SW monsoon	407	1.32
Kumar et al. (2011)	Singapore	Convective	328.64	1.29
		Transition	173.24	1.42
		Stratiform	309.2	1.39
Ramli and Tahir (2011)	Malaysia	<10 mm h^{-1}	180	1.9
		$10<x<30$ mm h^{-1}	212	1.9
		>30 mm h^{-1}	262	1.9
Atlas et al. (2000)	Kapingamarangi-Micronesia	-	216	1.25
Rosenfeld et al. (1993)	Darwin-Australia	-	230	1.25

3.5 Conclusions

Specific rain-type Z-R relations were derived for different mountain altitudes and locations, from a lower inter-Andean valley to higher páramo sites, representing the rainfall condition in the Andean cordillera, Ecuador. These Z-R relations were obtained based on the climatic conditions but also included the precipitation variability at each location using a mean volume diameter (D_m) classification analysis. This is the first study of this kind, performed at tropical high altitudes and simultaneously considering rainfall variability. The following conclusions were drawn from the analysis:

(i) The D_m classification approach used in this study points out to a better adjustment of rain-type

Z-R relationships in comparison with a single Z-R relationship. The parameters of the rain-type Z-R relationships increase according to D_m categorization. Our results confirm that at a given rain rate, a light rain category [0.1$<$ D_m(mm)$\leq$0.5] tends to be associated with smaller D_m and larger raindrop concentrations, whereas heavier rainfall categories tend to exhibit larger D_m and smaller droplet concentrations.

(ii) The influence of the sites' spatial exposition on the derivation of rain-type Z-R relationship is confirmed, where both location and elevation play a key role in characterizing the Z-R relationship. Thus, the spatial variability of rainfall in tropical high mountains should be analyzed before only using a single Z-R relation.

(iii) In the Andes of Ecuador, the Z-R relations mainly show a transitional rainfall nature, which differs from most studies in the equatorial zone. This result was particularly found at the lowest station (Balzay) and the station oriented toward the Pacific coastland (Zhurucay) in the highest radius class. Here, the presence of moist air advection, higher thermal breeze and better droplet growth conditions seems to produce more convective activity than in the remote sub-páramo station. The results reveal a clear difference in rainfall types at different terrain expositions, with drizzle being most dominant at the most isolated station, Virgen.

(iv) Tropical high elevation rainfall structure in terms of DSD, under the influence of steep and complex terrain in the high Andes, has been documented for the very fist time. Comparison with theoretical distribution models have shown that the exponential Marshall-Palmer model better fits the data at the highest station Zhurucay, while the gamma distribution is more suitable for the lower stations Virgen and Balzay.

(v) To our knowledge, this study constitutes the first contribution of tropical high mountain radar reflectivity determination by optical disdrometers and related rain-type Z-R relationships.

The Z-R relations found in this study will be used as starting point for rainfall mapping from raw reflectivity data of CAXX radar in southern-Ecuador (refer to Bendix et al., 2017) in areas where the influence attenuation can be overcome. Despite the attenuation issue, the advantages of the higher spatial resolution provided by X-band systems allow for more detailed studies of urban hydrology in the near range as planned for the city of Cuenca (30-km range for CAXX) to support water management. For this purpose, the disdrometers remain operational and will be used in the future. Consequently, further work will focus on exploring techniques to classify each pixel with the proper D_m category by using as inputs to the model the reflectivity values, the distance from the radar and the attenuation along the current bin.

Acknowledgments

The current study was conducted in the scope of two collaborating projects: the knowledge transfer project “RadarNetSur” (http://137.248.191.17:8080/radarnetsur/index.php?radar_id=SE), jointly administered by the University of Marburg and the Provincial Government of Loja (GPL) within the “Platform for Biodiversity and Ecosystem Monitoring and Research in South Ecuador” and the project “Identificación de los procesos hidrometeorológicos que desencadenan crecidas a partir de la información suministrada por un radar de precipitación”. The former was funded by the German Research Foundation DFG (BE1780/31-1 and BE1780/38-1) and the latter by the Research Office of the University of Cuenca (DIUC) and Empresa Pública Municipal de Telecomunicaciones, Agua Potable, Alcantarillado y Saneamiento de Cuenca (ETAPA-EP). Our thanks go to these institutions for their generous funding. We acknowledge the Ministry of Environment of Ecuador (MAE) for providing research permissions. We are grateful to the staff that contributed to the meteorological monitoring.

References

Alfieri, L., Claps, P., and Laio, F. (2010). Time-dependent Z - R relationships for estimating rainfall fields from radar measurements. *Natural Hazards and Earth System Sciences*, 203(1973):149–158.

Atlas, D., Ulbrich, C. W., Marks Jr., F. D., Black, R. A., Amitai, E., Willis, P. T., and Samsury, C. E. (2000). Partitioning tropical oceanic convective and stratiform rains by draft strength. *Journal of Geophysical Research: Atmospheres*, 105(D2):2259–2267.

Bamba, B., Ochou, A. D., Zahiri, E.-P., and Kacou, M. (2014). Consistency in Z-R Relationship Variability Regardless Precipitating Systems, Climatic Zones Observed from Two Types of Disdrometer. *Atmospheric and Climate Sciences*, 04:941–955.

Beard, K. V. (1976). Terminal velocity and shape of cloud and precipitation drops aloft. *Journal of the Atmospheric Sciences*, 33(5):851–864.

Bendix, J., Fries, A., Zárate, J., Trachte, K., Rollenbeck, R., Pucha-Cofrep, F., Paladines, R., Palacios, I., Orellana, J., Oñate-Valdivieso, F., Naranjo, C., Mendoza, L., Mejia, D., Guallpa, M., Gordillo, F., Gonzalez-Jaramillo, V., Dobbermann, M., Célleri, R., Carrillo, C., Araque, A., and Achilles, S. (2017). Radarnet-Sur First weather Radar Network in tropical high mountains. *Bulletin of the American Meteorological Society*, 98(6):1235–1254.

Bendix, J., Rollenbeck, R., Göttlicher, D., and Cermak, J. (2005). Cloud occurrence and cloud properties in Ecuador. *Climate Research*, 30:133–147.

Blanchard, D. C. (1953). Raindrop-size Distribution in Hawaiian Rains. *Journal of Meteorology*, 10(6):457–473.

Buytaert, W., Célleri, R., Willems, P., Bièvre, B. D., and Wyseure, G. (2006). Spatial and temporal rainfall variability in mountainous areas: A case study from the south Ecuadorian Andes. *Journal of Hydrology*, 329:413–421.

Caracciolo, C., Porcù, F., and Prodi, F. (2008). Precipitation classification at mid-latitudes in terms of drop size distribution parameters. *Advance Geosciences*, 16:11–17.

Carrillo-Rojas, G., Silva, B., Córdova, M., Célleri, R., and Bendix, J. (2016). Dynamic mapping of evapotranspiration using an energy balance-based model over an andean páramo catchment of southern ecuador. *Remote Sensing*, 8(2):1–8.

Célleri, R. and Feyen, J. (2009). The Hydrology of Tropical Andean Ecosystems: Importance, Knowledge Status, and Perspectives. *Mountain Research and Development*, 29:350–355.

Célleri, R., Willems, P., Buytaert, W., and Feyen, J. (2007). Space-time rainfall variability in the Paute basin, Ecuadorian Andes. *Hydrological Processes*, 21(24):3316–3327.

Cerro, C., Codina, B., Bech, J., and Lorente, J. (1997). Modeling Raindrop Size Distribution and Z(R) Relations in the Western Mediterranean Area. *Journal of Applied Meteorology*, 36:1470–1479.

Chen, B., Wang, J., and Gong, D. (2016). Raindrop size distribution in a midlatitude continental squall line measured by thies optical disdrometers over east China. *Journal of Applied Meteorology and Climatology*, 55(3):621–634.

Córdova, M., Carrillo-Rojas, G., Crespo, P., Wilcox, B., and Célleri, R. (2015). Evaluation of the Penman-Monteith (FAO 56 PM) Method for Calculating Reference Evapotranspiration Using Limited Data. *Mountain Research and Development*, 35:230–239.

Córdova, M., Célleri, R., Shellito, C. J., Orellana-Alvear, J., Abril, A., and Carrillo-Rojas, G. (2016). Near-surface air temperature lapse rate over complex terrain in the Southern Ecuadorian Andes: implications for temperature mapping. *Arctic, Antarctic and Alpine Reseseach*, 48(4):673–684.

Cunha, L. K., Smith, J. A., Krajewski, W. F., Baeck, M. L., and Seo, B.-C. (2015). NEXRAD NWS Polarimetric Precipitation Product Evaluation for IFloodS. *Journal of Hydrometeorology*, 16:1676–1699.

De Bièvre, B., Alvarado, A., Timbe, L., Célleri, R., and Feyen, J. (2003). Night Irrigation Reduction for Water Saving in Medium-Sized Systems. *ASCE Journal of Irrigation and Drainage Engineering*, 129:108–116.

Delrieu, G., Caoudal, S., and Creutin, J. D. (1997). Feasibility of Using Mountain Return for the Correction of Ground-Based X-Band Weather Radar Data. *Journal of Atmospheric and Oceanic Technology*, 14(3):368–385.

Dolan, B. and Rutledge, S. A. (2009). A theory-based hydrometeor identification algorithm for X-band polarimetric radars. *Journal of Atmospheric and Oceanic Technology*, 26(10):2071–2088.

Dunkerley, D. (2008). Rain event properties in nature and in rainfall simulation experiments: a comparative review with recommendations for increasingly systematic study and reporting. *Hydrological Processes*, 22:4415–4435.

Espinoza Villar, J. C., Ronchail, J., Guyot, J. L., Cochonneau, G., Naziano, F., Lavado, W., De Oliveira, E., Pombosa, R., and Vauchel, P. (2009). Spatio-temporal rainfall variability in the Amazon basin countries (Brazil, Peru, Bolivia, Colombia, and Ecuador). *International Journal of Climatology*, 29:1574–1594.

Fernández-Raga, M., Fraile, R., Keizer, J. J., Varela Teijeiro, M. E., Castro, A., Palencia, C., Calvo, A. I., Koenders, J., and Da Costa Marques, R. L. (2010). The kinetic energy of rain measured with an optical disdrometer: An application to splash erosion. *Atmospheric Research*, 96(2-3):225–240.

Frasson, R. P. d. M., da Cunha, L. K., and Krajewski, W. F. (2011). Assessment of the Thies optical disdrometer performance. *Atmospheric Research*, 101:237–255.

Friedrich, K., Higgins, S., Masters, F. J., and Lopez, C. R. (2013). Articulating and stationary PARSIVEL disdrometer measurements in conditions with strong winds and heavy rainfall. *Journal of Atmospheric and Oceanic Technology*, 30(9):2063–2080.

Germann, U., Galli, G., Boscacci, M., and Bolliger, M. (2006). Radar precipitation measurement in a mountainous region. *Quarterly Journal of the Royal Meteorological Society*, 132:1669–1692.

Hofstede, R. G. M., Dickinson, K. J. M., Mark, A. F., and Narváez, E. (2014). A Broad Transition from Cloud Forest to Páramo Characterizes an Undisturbed Treeline in Parque Nacional Llanganates, Ecuador. *Arctic, Antarctic, and Alpine Research*, 46(4):975–986.

Horvath, H. (2009). Gustav Mie and the scattering and absorption of light by particles: Historic developments and basics. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 110(11):787–799.

Jameson, A. R., Larsen, M. L., and Kostinski, A. B. (2015). On the Variability of Drop Size Distributions over Areas. *Journal of the Atmospheric Sciences*, 72(4):1386–1397.

Kumar, L. S., Lee, Y. H., Yeo, J. X., and Ong, J. T. (2011). Tropical Rain Classification and Estimation of Rain From Z-R (Reflectivity-Rain Rate) Relationships. *Progress In Electromagnetics Research*, 32:107–127.

Larsen, M. L. and Teves, J. B. (2015). Identifying Individual Rain Events with a Dense Disdrometer Network. *Advances in Meteorology*, 2015:1–12.

L'Ecuyer, T. S., Kummerow, C., and Berg, W. (2004). Toward a Global Map of Raindrop Size Distributions. Part I: Rain-Type Classification and Its Implications for Validating Global Rainfall Products. *Journal of Hydrometeorology*, 5:831–849.

Lee, G. and Zawadzki, I. (2003). Sequential Intensity Filtering Technique (SIFT): Filtering out noise tohighlight the physical variability of drop size distributions. In *31st. Conference Radar Meteorology Seattle,WA,American Meteorological Society*, pages 18–21.

Llasat, M.-C. (2001). An objective classification of rainfall events on the basis of their convective features: application to rainfall intensity in the northeast of Spain. *International Journal of Climatology*, 21:1385–1400.

Lo Conti, F., Francipane, A., Pumo, D., and Noto, L. V. (2015). Exploring single polarization X-band weather radar potentials for local meteorological and hydrological applications. *Journal of Hydrology*, 531:508–522.

Löffler-Mang, M., Kunz, M., and Schmid, W. (1999). On the performance of a low-cost K-band Doppler radar for quantitative rain measurements. *Journal of Atmospheric and Oceanic Technology*, 16(2-3):379–387.

Maki, M., Park, S.-G., and Bringi, V. (2005). Effect of Natural Variations in Rain Drop Size Distributions on Rain Rate Estimators of 3 cm Wavelength Polarimetric Radar. *Journal of the Meteorological Society of Japan*, 83(5):871–893.

Marshall, J. S. and Palmer, W. M. K. (1948). The Distribution of Raindrops With Size. *Journal of Meteorology*, 5:165–166.

Mätzler, C. (2002). MATLAB Functions for Mie Scattering and Absorption. Technical report, University of Bern. [Available online at http://arrc.ou.edu/~rockee/NRA_2007_website/Mie-scattering-Matlab.pdf].

Muñoz, P., Célleri, R., and Feyen, J. (2016). Effect of the Resolution of Tipping-Bucket Rain Gauge and Calculation Method on Rainfall Intensities in an Andean Mountain Gradient. *Water*, 8(11):534.

Narayana Rao, T., Narayana Rao, D., and Mohan, K. (2001). Classification of tropical precipitating systems and associated Z-R relationships. *Journal of Geophysical Research*, 106:17669–17711.

Nzeukou, A. and Sauvageot, H. (2004). Raindrop size distribution and radar parameters at Cape Verde. *Journal of Applied Meteorology*, 43:90–105.

Ochou, A. D., Zahiri, E.-P., Bamba, B., and Koffi, M. (2011). Understanding the Variability of Z-R Relationships Caused by Natural Variations in Raindrop Size Distributions (DSD): Implication of Drop Size and Number. *Atmospheric and Climate Sciences*, 01:147–164.

Padrón, R. S. (2013). Análisis de la estructura de la lluvia del páramo. [Available online at http://dspace.ucuenca.edu.

ec/bitstream/123456789/519/1/TESIS.pdf].

Padrón, R. S., Wilcox, B. P., Crespo, P., and Célleri, R. (2015). Rainfall in the Andean Páramo: New Insights from High-Resolution Monitoring in Southern Ecuador. *Journal of Hydrometeorology*, 16:985–996.

Pellarin, T., Delrieu, G., Saulnier, G.-M., Andrieu, H., Vignal, B., and Creutin, J.-D. (2002). Hydrologic Visibility of Weather Radar Systems Operating in Mountainous Regions: Case Study for the Ardèche Catchment (France). *Journal of Hydrometeorology*, 3(5):539–555.

Poveda, G., Mesa, O. J., Salazar, L. F., Arias, P. a., Moreno, H. a., Vieira, S. C., Agudelo, P. a., Toro, V. G., and Alvarez, J. F. (2005). The Diurnal Cycle of Precipitation in the Tropical Andes of Colombia. *Monthly Weather Review*, 133:228–240.

Pulwarty, R. S., Barry, R. G., Hurst, C. M., Sellinger, K., and Mogollon, L. F. (1998). Precipitation in the Venezuelan Andes in the context of regional climate. *Meteorology and Atmospheric Physics*, 67:217–237.

Ramli, S. and Tahir, W. (2011). Radar Hydrology: New Z/R Relationships for Quantitative Precipitation Estimation in Klang River Basin, Malaysia. *International Journal of Environmental Science and Development*, 2:223–227.

Rollenbeck, R. and Bendix, J. (2006). Experimental calibration of a cost-effective X-band weather radar for climate ecological studies in southern Ecuador. *Atmospheric Research*, 79(3-4):296–316.

Rollenbeck, R. and Bendix, J. (2011). Rainfall distribution in the Andes of southern Ecuador derived from blending weather radar data and meteorological field observations. *Atmospheric Research*, 99:277–289.

Rollenbeck, R., Bendix, J., and Fabian, P. (2011). Spatial and temporal dynamics of atmospheric water inputs in tropical mountain forests of South Ecuador. *Hydrological Processes*, 25:344–352.

Rosenfeld, D., Wolff, D. B., and Atlas, D. (1993). General Probability-matched Relations between Radar Reflectivity and Rain Rate. *Journal of Applied Meteorology*, 32(1):50–72.

Russell, B., Williams, E. R., Gosset, M., Cazenave, F., Descroix, L., Guy, N., Lebel, T., Ali, a., Metayer, F., and Quantin, G. (2010). Radar/rain-gauge comparisons on squall lines in Niamey, Niger for the AMMA. *Quarterly Journal of the Royal Meteorological Society*, 136:289–303.

Sarkar, T., Das, S., and Maitra, A. (2015). Assessment of different raindrop size measuring techniques: Inter-comparison of Doppler radar, impact and optical disdrometer. *Atmospheric Research*, 160:15–27.

Savina, M. (2011). *The Use of a Cost-Effective X-band Weather Radar in Alpine Region*. PhD thesis, Swiss Federal Institute of Technology Zurich.

Scheel, M. L. M., Rohrer, M., Huggel, C., Santos Villar, D., Silvestre, E., and Huffman, G. J. (2011). Evaluation of TRMM Multi-satellite Precipitation Analysis (TMPA) performance in the Central Andes region and its dependency on spatial and temporal resolution. *Hydrology and Earth System Sciences*, 15:2649–2663.

Sen Jaiswal, R., Uma, S., and Santhakumaran, A. (2009). Study of Z-R relationship over Gadanki for different rainfall rates. *Indian Journal of Radio and Space Physics*, 38:159–164.

Steiner, M., Smith, J., and Uijlenhoet, R. (2004). A Microphysical Interpretation of Radar Reflectivity-Rain Rate

Relationships. *Journal of the Atmospheric Sciences*, 61(10):1114–1131.

Tenório, S., da Silva Moraes, M. C., and Sauvageot, H. (2012). Raindrop Size Distribution and Radar Parameters in Coastal Tropical Rain Systems of Northeastern Brazil. *Journal of Applied Meteorology and Climatology*, 51:1960–1970.

Testud, J., Oury, S., Black, R. A., Amayenc, P., and Dou, X. (2001). The Concept of "Normalized" Distribution to Describe Raindrop Spectra: A Tool for Cloud Physics and Cloud Remote Sensing. *Journal of Applied Meteorology*, 40(6):1118–1140.

Thies Clima (2007). Instructions for use. Laser Precipitation Monitor 5.4110.xx.x00 V2.2xSTD. Technical report, Thies Clima.

Tokay, A., Hartmann, P., Battaglia, A., Gage, K. S., Clark, W. L., and Williams, C. R. (2009). A field study of reflectivity and Z-R relations using vertically pointing radars and disdrometers. *Journal of Atmospheric and Oceanic Technology*, 26:1120–1134.

Tokay, A., Kruger, A., and Krajewski, W. F. (2001). Comparison of Drop Size Distribution Measurements by Impact and Optical Disdrometers. *Journal of Applied Meteorology*, 40(11):2083–2097.

Tokay, A. and Short, D. A. (1996). Evidence from Tropical Raindrop Spectra of the Origin of Rain from Stratiform versus Convective Clouds. *Journal of Applied Meteorology*, 35:355–371.

Tokay, A., Wolff, D. B., Wolff, K. R., and Bashor, P. (2003). Rain gauge and disdrometer measurements during the Keys Area Microphysics Project (KAMP). *Journal of Atmospheric and Oceanic Technology*, 20:1460–1477.

Uijlenhoet, R. (2001). Raindrop size distributions and radar reflectivity-rain rate relationships for radar hydrology. *Hydrology and Earth System Sciences*, 5(4):615–628.

Uijlenhoet, R., Steiner, M., and Smith, J. A. (2003). Variability of Raindrop Size Distributions in a Squall Line and Implications for Radar Rainfall Estimation. *Journal of Hydrometeorology*, 4:43–61.

Ulbrich, C. W. (1983). Natural Variations in the Analytical Form of the Raindrop Size Distribution. *Journal of Climate and Applied Meteorology*, 22(10):1764–1775.

Ulbrich, C. W. and Atlas, D. (1998). Rainfall Microphysics and Radar Properties: Analysis Methods for Drop Size Spectra. *Journal of Applied Meteorology*, 37(1994):912–923.

van de Beek, C. Z., Leijnse, H., Hazenberg, P., and Uijlenhoet, R. (2016). Close-range radar rainfall estimation and error analysis. *Atmospheric Measurement Techniques*, 9:3837–3850.

van de Beek, C. Z., Leijnse, H., Stricker, J. N. M., Uijlenhoet, R., and Russchenberg, H. W. J. (2009). Performance of high-resolution X-band radar for rainfall measurement in The Netherlands. *Hydrology and Earth System Sciences*, 6(5):6035–6085.

van de Hulst, H. C. (1957). *Light Scattering by Small Particles*. Dover Books on Physics. Dover Publications.

Varikoden, H., Preethi, B., Samah, A. A., and Babu, C. A. (2011). Seasonal variation of rainfall characteristics in different intensity classes over Peninsular Malaysia. *Journal of Hydrology*, 404:99–108.

Vuille, M., Bradley, R. S., and Keimig, F. (2000). Climate Variability in the Andes of Ecuador and Its Relation to Tropical Pacific and Atlantic Sea Surface Temperature Anomalies. *Journal of Climate*, 13(14):2520–2535.

Wexler, R. and Atlas, D. (1963). Radar Reflectivity and Attenuation of Rain. *Journal of Applied Meteorology*, 2(2):276–280.

Wohl, E., Barros, A., Brunsell, N., Chappell, N. A., Coe, M., Giambelluca, T., Goldsmith, S., Harmon, R., Hendrickx, J. M. H., Juvik, J., McDonnell, J., and Ogden, F. (2012). The hydrology of the humid tropics. *Nature Climate Change*, 2:655–662.

Yoon, S.-S., Phuong, T. A., and Bae, D.-H. (2012). Quantitative Comparison of the Spatial Distribution of Radar and Gauge Rainfall Data. *Journal of Hydrometeorology*, 13:1939–1953.

Zhang, J., Howard, K., Langston, C., Vasiloff, S., Kaney, B., Arthur, A., Van Cooten, S., Kelleher, K., Kitzmiller, D., Ding, F., Seo, D. J., Wells, E., and Dempsey, C. (2011). National mosaic and multi-sensor QPE (NMQ) system description, results, and future plans. *Bulletin of the American Meteorological Society*, 92:1321–1338.

4 Optimization of X-band rainfall radar retrieval in the Southern Andes of Ecuador using a random forest model

This chapter is published in *Remote Sensing (MDPI)*, 11(14), 1632, 2019
Received: 30 May 2019
Accepted: 05 July 2019
First published online: 10 July 2019
DOI: https://doi.org/10.3390/rs11141632

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

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.

4.1 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 (Orellana-Alvear et al., 2017). 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 nowdays to dual-polarization devices (Anagnostou et al., 2018; Kirstetter et al., 2015). 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, straighforward 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 (Büyükbaş, 2009). Many of them are part of relevant radar networks used by the Collaborative Adaptive Sensing of the Atmosphere (CASA) Integrated Project 1 (McLaughlin et al., 2009), or the hydrology-oriented Iowa Flood Studies (iFloodS) field campaign (Mishra et al., 2016), just to name a few. The main limitation of X-band radars come from their attenuation issues. Nonetheless, several studies (Anagnostou et al., 2018; Diederich et al., 2015; Feng et al., 2016), 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 (Villarini and Krajewski, 2010) compared to dual-polarization, has also been successfully deployed (e.g., Allegretti, 2012; Antonini et al., 2017; Lo Conti et al., 2015), even in remote areas such as high mountains. Recently, Bendix et al. (2017) reported the first X-band radar network at high altitudes in the southern Andes of Ecuador (Radarnet-Sur). Several studies (Fries et al., 2014; Oñate-Valdivieso et al., 2018; Rollenbeck and Bendix, 2011) 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 (McRoberts and Nielsen-Gammon, 2017), attenuation (Berne et al., 2006), radome attenuation (Frasier et al., 2013) to name the most important ones, and then applies either a generalized or a locally derived Z-R relationship such as in Beek et al. (2010) and van de Beek et al. (2016). However, this process is usually site specific and intricate, as corrections in reflectivity values also add noise and uncertainty to radar images (Harrison et al., 2000). Frequently, an additional step of bias adjustment using rain gauges is performed, as in Morin and Gabella (2007). 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 (Alqudah et al., 2013; Kusiak et al., 2013; Meyer et al., 2017; Yang et al., 2016). 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 (Orlandini and Morlini, 2000; Teschl et al., 2007) 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 (Jing et al., 2017) and radar QPE using reflectivity of several heights (Yang et al., 2017; Yu et al., 2017). 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.

4.2 Materials and Methods

4.2.1 Study site

The study site constitutes the south-east quadrant of the CAXX radar coverage to an extent of 60 km radius (see Figure 4.1 and Section 4.2.2 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 (Célleri et al., 2007). The area extends over valley and mountain landscapes where highest elevations are covered by páramo vegetation (Célleri et al., 2007) 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 (Coltorti and Ollier, 2000). 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 (Campozano et al., 2016).

4.2.2 Instruments and data

4.2.2.1 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; Bendix et al., 2017).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 4.1. 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 4.2). 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

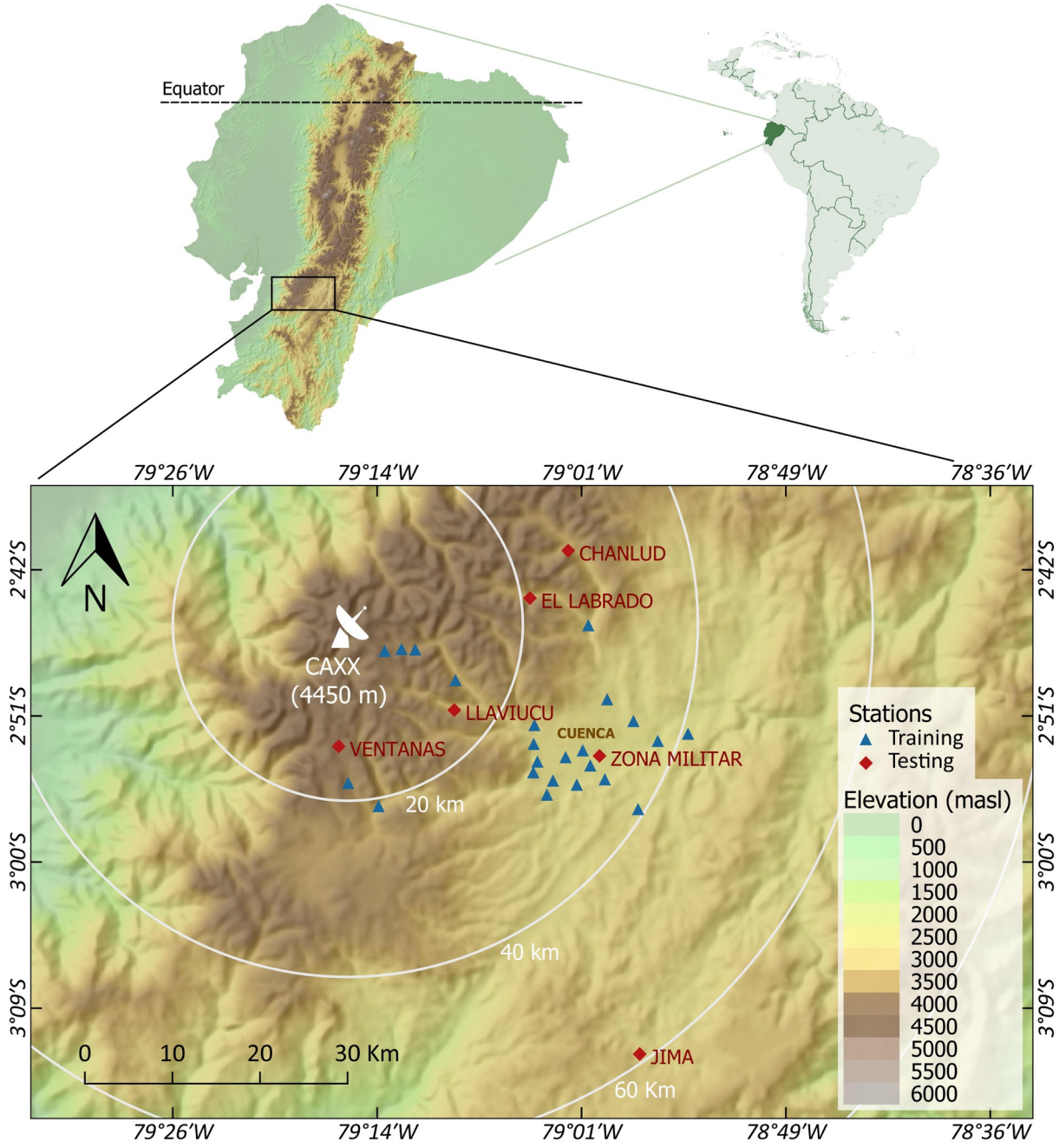

Figure 4.1: Map of the study site and the rain gauge network distribution.

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

Table 4.1: 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
Antena	
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

45 rainfall events that occurred within the period January 2016 to June 2017 were used in this study.

4.2.2.2 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, (Orellana-Alvear et al., 2017) 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 (Orellana-Alvear et al., 2017). Although these data were not used in the current study for modelling purposes, a comparison with the reflectivity records of the

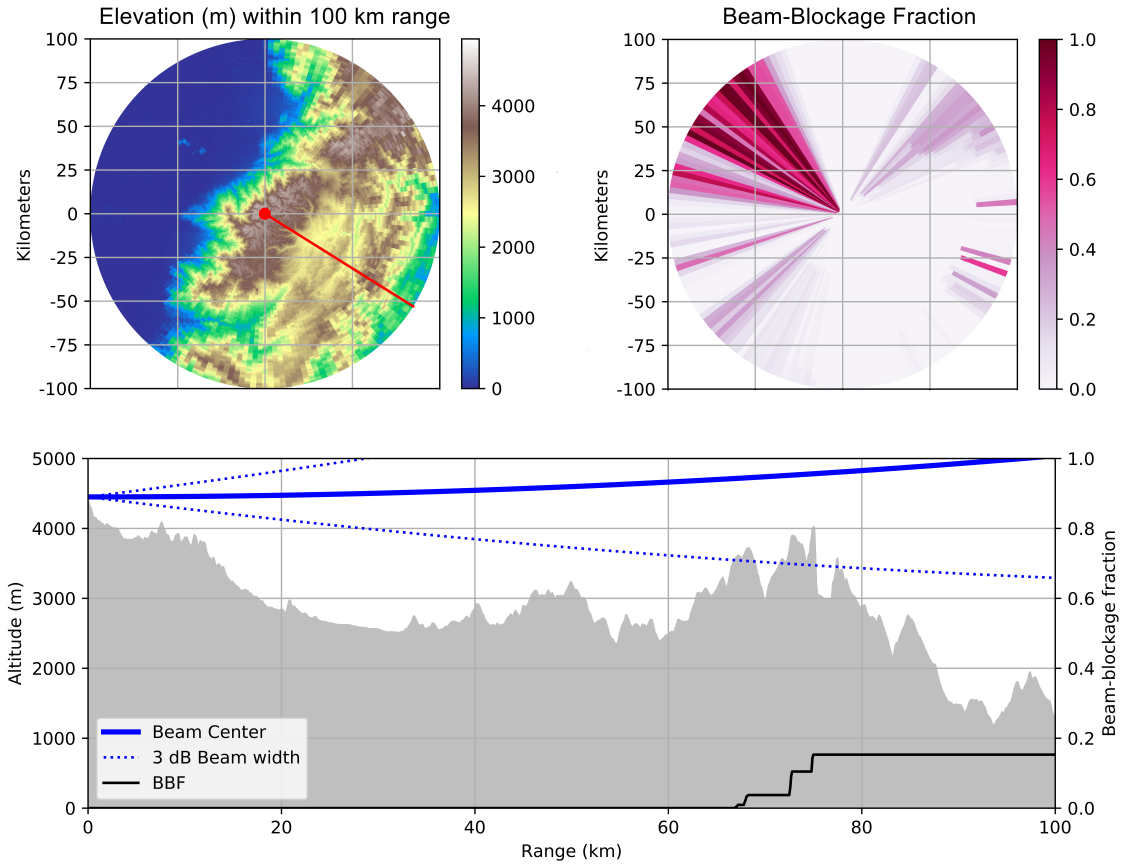

Figure 4.2: Beam-blockage fraction (BBF) calculated for the CAXX coverage. Transversal section for azimuth 122 (Balzay station location) is illustrated. The grey shaded area depicts the elevation of the terrain.

radar was performed. The latter ensures the proper calibration of the CAXX. Figure 4.3 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 2017 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

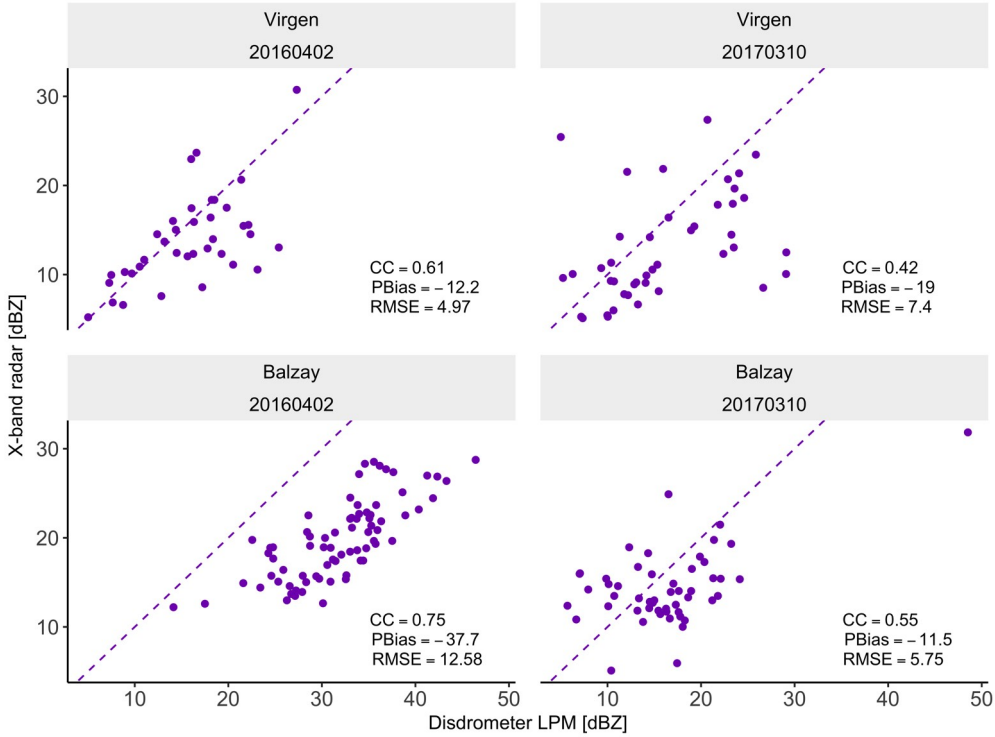

Figure 4.3: Reflectivity [dBZ] as recorded by both disdrometers at Virgen (7 km) and Balzay (30 km) from the radar location in comparison with the X-band CAXX measurements (>= 5 dBZ) for two rainfall events. Pearson's correlation coefficient (CC), percentage bias (PBias) and root mean squared error (RMSE) are illustrated.

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 (Orellana-Alvear et al., 2017).

4.2.2.3 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 4.1). 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 4.2.3.2 and Figure 4.4), rain gauge stations were split for the training-validation and test stages. The former used 23 rain gauges and the latter 6

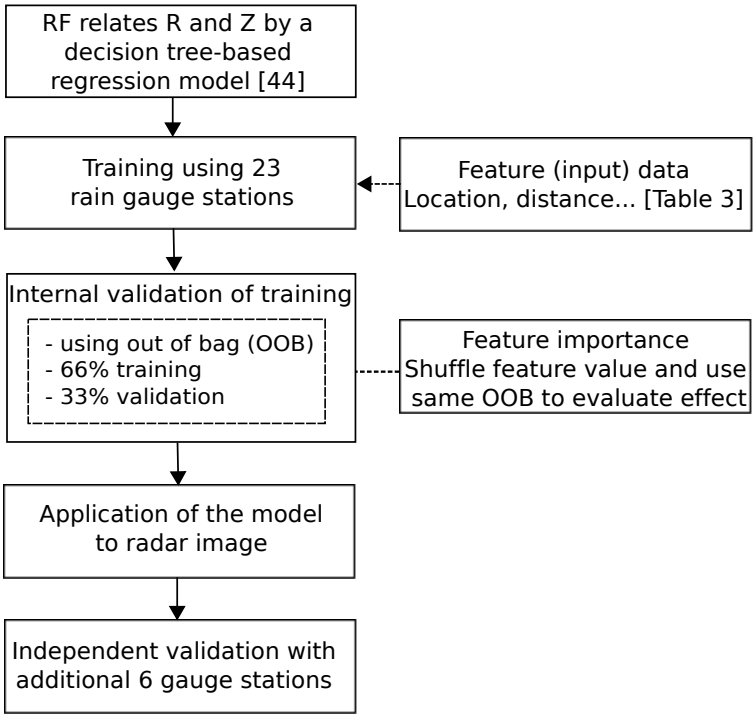

Figure 4.4: Simplified workflow of the development of the random forest (RF) model for the current study.

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 4.2. 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 16000 5-min intervals of positive ($R > 0.1$ mm h^{-1}) rainfall. The remaining 15 rainfall events were used in the test phase.

4.2.3 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.

Table 4.2: (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

4.2.3.1 Step-wise correction approach

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., Lo Conti et al., 2015; Marra and Morin, 2015; van de Beek et al., 2016). To facilitate the reproducibility of this study, we used the methods implemented in the *wradlib* library (Heistermann et al., 2015, 2013) 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 (2015). 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 (Gabella and Notarpietro, 2002), 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 4.1), the main beam is currently zero degrees oriented. Figure 4.2 denotes the beam blockage map calculated based on Bech et al. (Eq.2 and Appendix; Bech et al., 2003). 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 (2009) and evaluated in Jacobi and Heistermann (2016) 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 (2016). 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 $Z=200R^{1.6}$ (Marshall and Palmer, 1948) and b) a site-specific Z-R relationship $Z=103R^{2.06}$ derived by Orellana-Alvear et al. (2017) 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.

4.2.3.2 QPE Random Forest model

Figure 4.4 illustrates a simplified workflow of the random forest model implemented for quantitative precipitation estimation in the current study. Random Forest (RF) (Breiman, 2001) 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, a RF model was trained by using data of 30 selected rainfall events from 23 training rain gauge stations. From these, around $\sim$16000 rain observations were obtained at 5-min frequency and used as target variable. Input features described in Table 4.3 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 a 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 $\sim$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.

4.2.3.2.1 Input Features. — The use of textural features as inputs of data-driven models has proven to be effective on non-coherent radar applications (Hedir and Haddad, 2016). 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 4.3 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 (Breiman, 2001) 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 (Pedregosa et al., 2011). 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).

4.2.4 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. (2013) 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 (1970). Both, ME and RMSE were normalized by the mean of the reference values. Finally, the measurement error was calculated through Equation 4.1. According to Thurai et al. (2017), this metric is related to the radar

Table 4.3: 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, 2002)).
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*.

parameters used in a specific algorithm.

$$\frac{var(R)}{(\bar{R})^2} \tag{4.1}$$

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

4.3 Analyses and Discussion

4.3.1 Step-Wise Correction Model

The step-wise procedure was applied on the radar imagery to derive rainfall rates as explained in section 4.2.3.1. 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 4.2. 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 ($\sim$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 4.5 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: (a) static clutter is concentrated in the close range of the radar ($\sim$5-7 km), (b) 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 Villarini and Krajewski (2010) 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 4.5, region B). A similar effect was obtained in Gabella and Notarpietro (2002) 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.

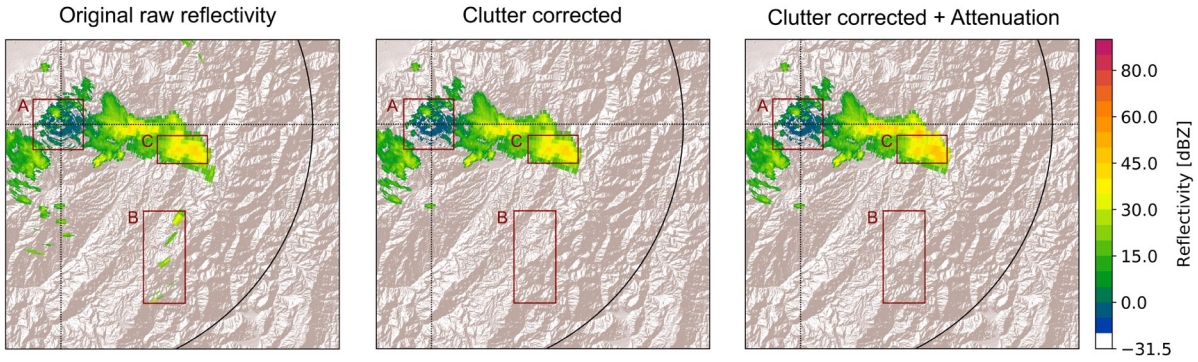

Figure 4.5: Static-Dynamic clutter and attenuation correction applied to the radar scene from 2017.02.10 15:45 local time. From left to right: (i) original uncorrected reflectivity, (ii) upscaled reflectivity after static (region A) and dynamic clutter (region B) removal, and (iii) upscaled reflectivity after clutter removal plus attenuation correction (region C).

4.3.2 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 4.6 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* (3x3 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 (Muñoz et al., 2016), 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.

The correlation of observed and estimated hourly rainfall applying the QPE random forest model to the entire test dataset is shown in Figure 4.7. It can be observed that lower rain rates

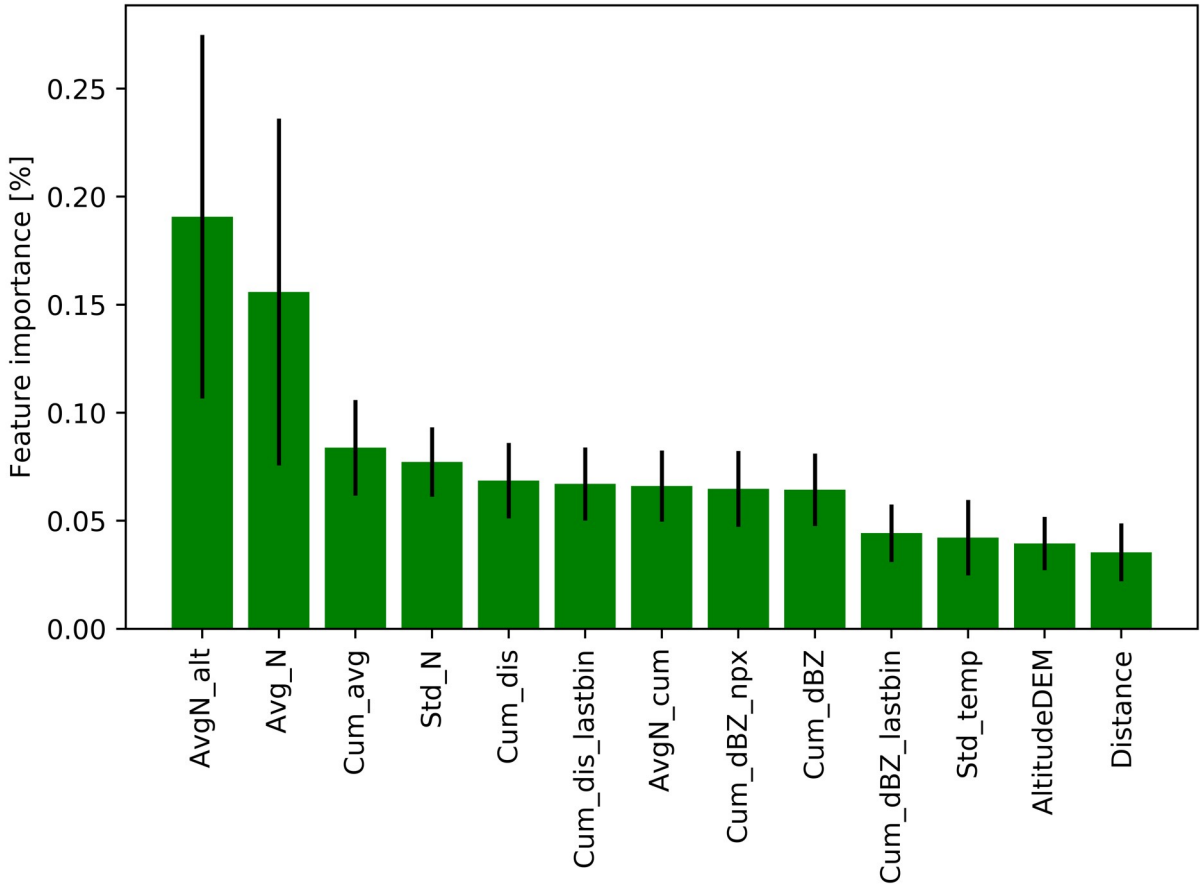

Figure 4.6: Feature importance and its standard deviation on the RF model. Out-of-bag (OOB) error was used to rank the features in the training stage.

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 Beek et al. (2010), but in contrast the authors implemented an exhaustive step-wise approach on the X-band radar imagery under about 15 km coverage. Furthermore, Beek et al. (2010) found that rainfall event totals less than 3 mm h^{-1} were overestimated, while higher total events were mostly underestimated by using a 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 4.7 that rain observations higher than 5 mm h^{-1} are underestimated in all cases. There are possible

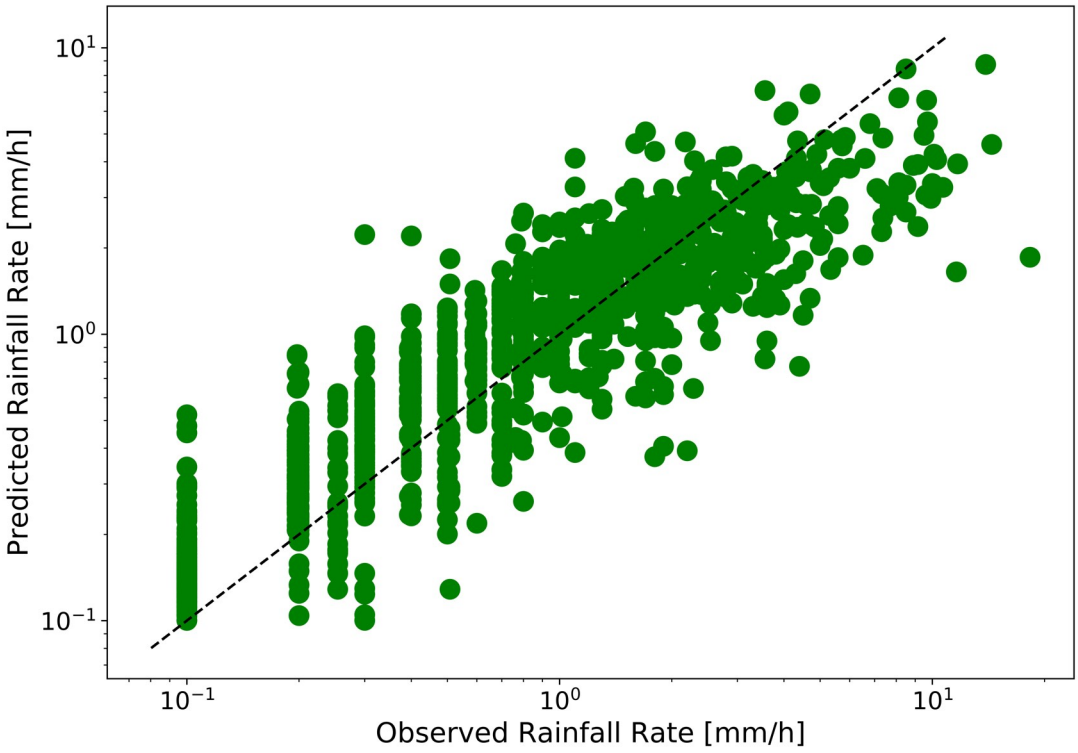

Figure 4.7: Correlation of observed and estimated hourly rainfall using QPE random forest model on the entire test dataset (15 independent events) is illustrated. The 1-1 line is show in black dotted.

explanations for this situation: (i) lack of training samples under similar conditions. As reported by Orellana-Alvear et al. (2017), 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.

4.3.3 Comparison of the Models - Temporal and Spatial Evaluation

The RF model consistently outperformed the other models at all spatial locations as seen in Figure 4.8 and Table 4.4. 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 van de Beek et al. (2016), 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 Morin and Gabella (2007) and Marra and Morin (2015) or vertical reflectivity profiles in the case of machine learning approaches as in Alqudah et al. (2013); Kusiak et al. (2013). The performance of the implemented RF technique in Beek et al. (2010) (correlation coefficient of 0.82) was likewise comparable to our RF model (CC = 0.76) using a short-term dataset. However, it should be notice 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.

Some studies such as Lo Conti et al. (2015) that did not use a rain gauge network for adjustment of the rainfall derivation but accomplished a similar or better model performance where 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 Lo Conti et al. (2015) 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 Lo Conti et al. (2015) that used a unique event-based derived Z-R for the rainfall derivation of the Palermo radar, Orellana-Alvear et al. (2017) 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

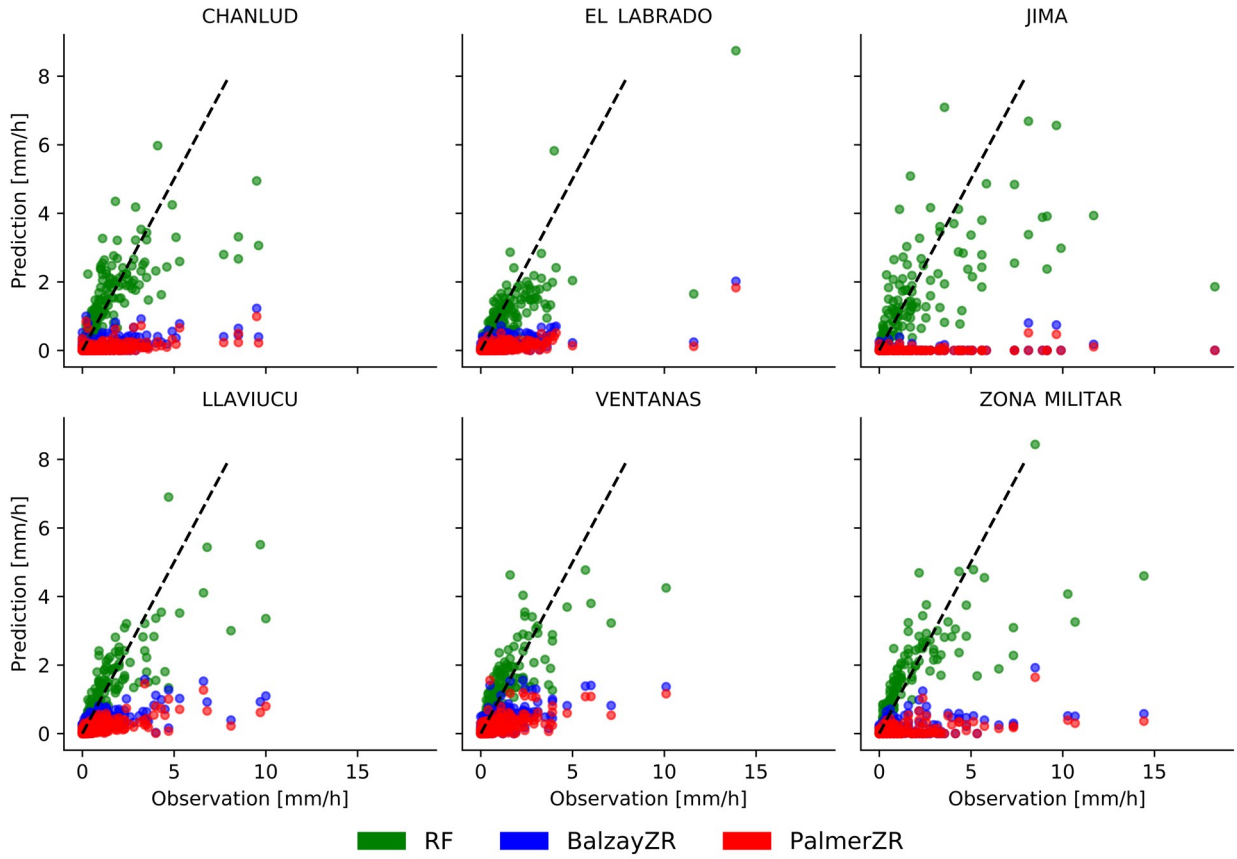

Figure 4.8: Scatter plot of observed rain rate vs. predicted rain rate for hourly accumulated rainfall at all six testing stations using the test data set (15 independent events). The 1-1 line is shown in black dotted. Three models are illustrated: random forest model (green), step-wise model using site-specific Balzay Z-R relation (blue) and step-wise model using Marshall-Palmer Z-R relation (red).

Table 4.4: Statistical measures of the performance of the different rainfall retrieval models at all test locations.

Metric	Model	Chanlud	El Labrado	Jima	Laviucu	Ventanas	Z.Militar
MSE	RF	0.97	0.90	4.30	0.78	0.55	1.85
	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
CC	RF	0.77	0.81	0.66	0.83	0.81	0.78
	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
rME	RF	-0.01	-0.24	-0.27	-0.06	0.09	-0.05
	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
rRMSE	RF	0.91	0.93	1.11	0.85	0.78	0.96
	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
Slope	RF	1.14	1.38	1.26	1.22	1.08	1.31
	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

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 Alqudah et al. (2013); Kusiak et al. (2013); Teschl et al. (2007) (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 $\sim$0.8-0.9) of neural networks applications as in Alqudah et al. (2013); Kusiak et al. (2013). Nonetheless, also a lower performance (correlation coefficient of 0.37) was found by Teschl et al. (2007) 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.

Table 4.5: Measurement errors for all rainfall retrieval models by considering different rain rate intervals.

Rain rate [mm h^{-1}]	RF	BalzayZR	PalmerZR
R < 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
R >=10	0.24	0.72	1.01

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 Anagnostou et al. (2018).

In relation to the measurement errors, it can be seen in Table 4.5, 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.

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 4.2.3.2 as representative for different rainfall formation processes, shows a slight overestimation of drizzle and underestimation of the peaks, which holds for all rainfall events as seen in Figure 4.9. 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.

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

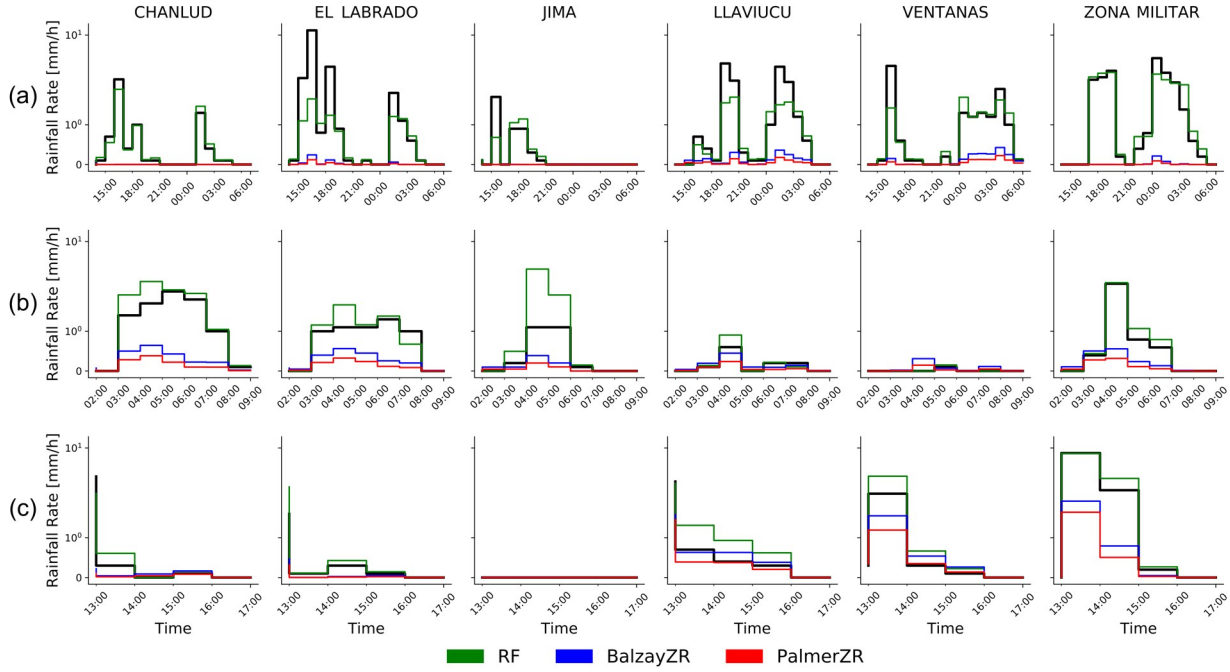

Figure 4.9: Hyetograms of the three representative rainfall events from the independent test set: a) 09 March 2016 14:00 - 10 March 2016 07:30, b) 10 June 2016 02:00 - 10 June 2016 09:00 and c) 19 March 2017 13:00 - 19 March 2017 17:00 as seen by the testing stations. Observations are illustrated in 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.

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 1^{st} 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 Bendix et al. (2017), 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.

4.4 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 (Orellana-Alvear et al., 2017). From the results several conclusions can be drawn:

(i) As expected, the locally 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 emphasize 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 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-Peru. 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.

Acknowledgments

We acknowledge the Ministry of Environment of Ecuador (MAE) for providing research permissions. We are grateful to the technical staff that contributed to the meteorological monitoring and particularly setting-up and operational monitoring of the CAXX radar equipment: Mario Guallpa and Andreas Fries.

References

Allegretti, M. (2012). X-Band Mini Radar for Observing and Monitoring Rainfall Events. *Atmospheric and Climate Sciences*, 02(03):290–297.

Alqudah, A., Chandrasekar, V., and Le, M. (2013). Investigating rainfall estimation from radar measurements using neural networks. *Natural Hazards and Earth System Science*, 13:535–544.

Anagnostou, M. N., Kalogiros, J., Marzano, F. S., Anagnostou, E. N., Montopoli, M., and Picciotti, E. (2013). Performance Evaluation of a New Dual-Polarization Microphysical Algorithm Based on Long-Term X-Band Radar and Disdrometer Observations. *Journal of Hydrometeorology*, 14(April):560–576.

Anagnostou, M. N., Nikolopoulos, E. I., Kalogiros, J., Anagnostou, E. N., Marra, F., Mair, E., Bertoldi, G., Tappeiner, U., and Borga, M. (2018). Advancing precipitation estimation and streamflow simulations in complex terrain with X-Band dual-polarization radar observations. *Remote Sensing*, 10(8).

Antonini, A., Melani, S., Corongiu, M., Romanelli, S., Mazza, A., Ortolani, A., and Gozzini, B. (2017). On the Implementation of a regional X-bandweather radar network. *Atmosphere*, 8(2):1–20.

Bech, J., Codina, B., Lorente, J., and Bebbington, D. (2003). The sensitivity of single polarization weather radar beam blockage correction to variability in the vertical refractivity gradient. *Journal of Atmospheric and Oceanic Technology*, 20(6):845–855.

Beek, C. Z. V. D., Leijnse, H., Stricker, J. N. M., Uijlenhoet, R., and Russchenberg, H. W. J. (2010). Performance of high-resolution X-band radar for rainfall measurement in The Netherlands. *Hydrology and Earth System Sciences*, 14:205–221.

Bendix, J., Fries, A., Zárate, J., Trachte, K., Rollenbeck, R., Pucha-Cofrep, F., Paladines, R., Palacios, I., Orellana, J., Oñate-Valdivieso, F., Naranjo, C., Mendoza, L., Mejia, D., Guallpa, M., Gordillo, F., Gonzalez-Jaramillo, V., Dobbermann, M., Célleri, R., Carrillo, C., Araque, A., and Achilles, S. (2017). Radarnet-Sur First weather Radar Network in tropical high mountains. *Bulletin of the American Meteorological Society*, 98(6):1235–1254.

Berne, A., Uijlenhoet, R., Berne, A., and Quantitative, R. U. (2006). Quantitative analysis of X-band weather radar attenuation correction accuracy. *Natural Hazards and Earth System Sciences*, 6:419–425.

Breiman, L. (2001). Random forests. *Machine learning*, 45(1):5–32.

Büyükbaş, E. (2009). Assess the current and potential capabilities of weather radars for the use in wmo integrated global observing system (wigos). In *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.

Campozano, L., Célleri, R., Trachte, K., Bendix, J., and Samaniego, E. (2016). Rainfall and Cloud Dynamics in the Andes: A Southern Ecuador Case Study. *Advances in Meteorology*, 2016.

Célleri, R., Willems, P., Buytaert, W., and Feyen, J. (2007). Space-time rainfall variability in the Paute basin, Ecuadorian Andes. *Hydrological Processes*, 21(24):3316–3327.

Coltorti, M. and Ollier, C. D. (2000). Geomorphic and tectonic evolution of the Ecuadorian Andes. *Geomorphology*, 32(1-2):1–19.

Diederich, M., Ryzhkov, A., Simmer, C., Zhang, P., and Trömel, S. (2015). Use of Specific Attenuation for Rainfall Measurement at X-Band Radar Wavelengths. Part II : Rainfall Estimates and Comparison with Rain Gauges. *Journal of Hydrometeorology*, 16(April):503–516.

Feng, L., Xiao, H., Wen, G., Li, Z., Sun, Y., Tang, Q., and Liu, Y. (2016). Rain Attenuation Correction of Reflectivity for X-Band Dual Polarization Radar. *Atmosphere*, 7(164):1–17.

Frasier, S. J., Kabeche, F., Figueras i Ventura, J., Al-Sakka, H., Tabary, P., Beck, J., and Bousquet, O. (2013). In-Place Estimation of Wet Radome Attenuation at X Band. *Journal of Atmospheric and Oceanic Technology*, 30:917–928.

Fries, A., Rollenbeck, R., Bayer, F., Gonzalez, V., Oñate-Valivieso, F., Peters, T., and Bendix, J. (2014). Catchment precipitation processes in the San Francisco valley in southern Ecuador: combined approach using high-resolution radar images and in situ observations. *Meteorology and Atmospheric Physics*, 126(1-2):13–29.

Gabella, M. and Notarpietro, R. (2002). ERAD 2002 Ground clutter characterization and elimination in mountainous terrain. *Proceedings of ERAD*, pages 305–311.

Harrison, D. L., Driscoll, S. J., and Kitchen, M. (2000). Improving precipitation estimates from weather radar using quality control and correction techniques. *Meteorological Applications*, 6:135–144.

Hedir, M. and Haddad, B. (2016). Automatic system for radar echoes filtering based on textural features and artificial intelligence. *Meteorology and Atmospheric Physics*.

Heistermann, M., Collis, S., Dixon, M. J., Giangrande, S., Helmus, J. J., Kelley, B., Koistinen, J., Michelson, D. B., Peura, M., Pfaff, T., and Wolff, D. B. (2015). The emergence of open-source software for the weather radar community. *Bulletin of the American Meteorological Society*, 96(1):117–128.

Heistermann, M., Jacobi, S., and Pfaff, T. (2013). Technical Note: An open source library for processing weather radar data (wradlib). *Hydrology and Earth System Sciences*, 17(2):863–871.

Jacobi, S. and Heistermann, M. (2016). Benchmarking attenuation correction procedures for six years of single-polarized C-band weather radar observations in South-West Germany. *Geomatics, Natural Hazards and Risk*, 5705(July 2017).

Jing, W., Zhang, P., Jiang, H., and Zhao, X. (2017). Reconstructing Satellite-Based Monthly Precipitation over Northeast China Using Machine Learning Algorithms. *Remote Sensing*, 9:1–17.

Kirstetter, P.-E., Gourley, J. J., Hong, Y., Zhang, J., Moazamigoodarzi, S., Langston, C., and Arthur, A. (2015). Probabilistic precipitation rate estimates with ground-based radar networks. *Water Resources Research*, 51:1422–1442.

Krämer, S. and Verworn, H. R. (2009). Improved radar data processing algorithms for quantitative rainfall estimation in real time. *Water Science and Technology*, 60(1):175–184.

Kusiak, A., Wei, X., Verma, A. P., Member, S., and Roz, E. (2013). Modeling and Prediction of Rainfall Using

Radar Reflectivity Data : A Data-Mining Approach. *IEEE Transactions on Geoscience and Remote Sensing*, 51(4):2337–2342.

Lo Conti, F., Francipane, A., Pumo, D., and Noto, L. V. (2015). Exploring single polarization X-band weather radar potentials for local meteorological and hydrological applications. *Journal of Hydrology*, 531:508–522.

Marra, F. and Morin, E. (2015). Use of radar QPE for the derivation of Intensity-Duration-Frequency curves in a range of climatic regimes. *Journal of Hydrology*, 531:427–440.

Marshall, J. S. and Palmer, W. M. K. (1948). The Distribution of Raindrops With Size. *Journal of Meteorology*, 5:165–166.

McLaughlin, D., Pepyne, D., Chandrasekar, V., Philips, B., Kurose, J., Zink, M., Droegemeier, K., Cruz-Pol, S., Junyent, F., Brotzge, J., Westbrook, D., Bharadwaj, N., Wang, Y., Lyons, E., Hondl, K., Liu, Y., Knapp, E., Xue, M., Hopf, A., Kloesel, K., DeFonzo, A., Kollias, P., Brewster, K., Contreras, R., Dolan, B., Djaferis, T., Insanic, E., Fraiser, S., and Carr, F. (2009). Short-wavelength technology and the potential for distributed networks of small radar systems. *Bulletin of the American Meteorological Society*, 90(12):1797–1817.

McRoberts, D. B. and Nielsen-Gammon, J. W. (2017). Detecting Beam Blockage in Radar-Based Precipitation Estimates. *Journal of Atmospheric and Oceanic Technology*, 34:1407–1422.

Meyer, H., Kühnlein, M., Reudenbach, C., Nauss, T., Meyer, H., Kühnlein, M., Reudenbach, C., Meyer, H., Kühnlein, M., Reudenbach, C., and Nauss, T. (2017). 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 Sensing Letters*, 8(7):647–656.

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., Tolstoy, L., Kucera, P. A., Petersen, W. A., Grazioli, J., and Pazmany, A. L. (2016). Deployment and Performance Analyses of High-Resolution Iowa XPOL Radar System during the NASA IFloodS Campaign. *Journal of Hydrometeorology*, 17(Febraury):455–479.

Morin, E. and Gabella, M. (2007). Radar-based quantitative precipitation estimation over Mediterranean and dry climate regimes. *Journal of Geophysical Research*, 112(October):1–13.

Muñoz, P., Célleri, R., and Feyen, J. (2016). Effect of the Resolution of Tipping-Bucket Rain Gauge and Calculation Method on Rainfall Intensities in an Andean Mountain Gradient. *Water*, 8(11):534.

Nash, J. E. and Sutcliffe, J. (1970). River flow forecasting through conceptual models Part I - A discussion of principles. *Journal of Hydrology*, 10:282–290.

Oñate-Valdivieso, F., Fries, A., Mendoza, K., Gonzalez-Jaramillo, V., Pucha-Cofrep, F., Rollenbeck, R., and Bendix, J. (2018). Temporal and spatial analysis of precipitation patterns in an Andean region of southern Ecuador using LAWR weather radar. *Meteorology and Atmospheric Physics*, 130(4):473–484.

Orellana-Alvear, J., Célleri, R., Rollenbeck, R., and Bendix, J. (2017). Analysis of Rain Types and Their Z - R Relationships at Different Locations in the High Andes of Southern Ecuador. *Journal of Applied Meteorology and Climatology*, 56:3065–3080.

Orlandini, S. and Morlini, I. (2000). Artificial neural networks estimation of rainfall intensity from radar observations. *Journal of Geophysical Research*, 105(D20):849–861.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michael, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, É. (2011). Scikit-learn : Machine Learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

Rollenbeck, R. and Bendix, J. (2011). Rainfall distribution in the Andes of southern Ecuador derived from blending weather radar data and meteorological field observations. *Atmospheric Research*, 99:277–289.

Teschl, R., Randeu, W. L., and Teschl, F. (2007). Improving weather radar estimates of rainfall using feed-forward neural networks. *Neural Networks*, 20:519–527.

Thurai, M., Mishra, K. V., Bringi, V. N., and Krajewski, W. F. (2017). Initial Results of a New Composite-Weighted Algorithm for Dual-Polarized X-Band Rainfall Estimation. *Journal of Hydrometeorology*, 18(April):1081–1100.

van de Beek, C. Z., Leijnse, H., Hazenberg, P., and Uijlenhoet, R. (2016). Close-range radar rainfall estimation and error analysis. *Atmospheric Measurement Techniques*, 9:3837–3850.

Villarini, G. and Krajewski, W. F. (2010). Review of the different sources of uncertainty in single polarization radar-based estimates of rainfall. *Surveys in Geophysics*, 31(1):107–129.

Yang, T.-h., Feng, L., and Chang, L.-Y. (2016). Improving radar estimates of rainfall using an input subset of artificial neural networks. *Journal of Applied Remote Sensing*, 10(2):1–12.

Yang, X., Kuang, Q., Zhang, W., and Zhang, G. (2017). A terrain-based weighted random forests method for radar quantitative. *Meteorological Applications*, 414(May):404–414.

Yu, P.-s., Yang, T.-c., Chen, S.-y., Kuo, C.-m., and Tseng, H.-w. (2017). Comparison of random forests and support vector machine for real-time radar-derived rainfall forecasting. *Journal of Hydrology*, 552:92–104.

5 Assessment of native radar reflectivity and radar rainfall estimates for discharge forecasting in mountain catchments with a random forest model

This chapter is published in *Remote Sensing (MDPI)*, 12(12), 1986, 2020
Received: 09 May 2020
Accepted: 18 June 2020
First published online: 20 June 2020
DOI: https://doi.org/10.3390/rs12121986

Assessment of Native Radar Reflectivity and Radar Rainfall Estimates for Discharge Forecasting in Mountain Catchments with a Random Forest Model

Johanna Orellana-Alvear[1,2], Rolando Célleri[2,3], Rütger Rollenbeck[1], Paul Muñoz[2], Pablo Contreras[2] 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

Abstract: Discharge forecasting is a key component for early warning systems and extremely useful for decision makers. Forecasting models require accurate rainfall estimations of high spatial resolution and other geomorphological characteristics of the catchment, which are rarely available in remote mountain regions such as the Andean highlands. While radar data is available in some mountain areas, the absence of a well distributed rain gauge network makes it hard to obtain accurate rainfall maps. Thus, this study explored a Random Forest model and its ability to leverage native radar data (i.e., reflectivity) by providing a simplified but efficient discharge forecasting model for a representative mountain catchment in the southern Andes of Ecuador. This model

was compared with another that used as input derived radar rainfall (i.e., rainfall depth), obtained after the transformation from reflectivity to rainfall rate by using a local Z-R relation and a rain gauge-based bias adjustment. In addition, the influence of a soil moisture proxy was evaluated. Radar and runoff data from April 2015 to June 2017 were used. Results showed that (i) model performance was similar by using either native or derived radar data as inputs ($0.66 < NSE < 0.75$; $0.72 < KGE < 0.78$). Thus, exhaustive pre-processing for obtaining radar rainfall estimates can be avoided for discharge forecasting. (ii) Soil moisture representation as input of the model did not significantly improve model performance (i.e., NSE increased from 0.66 to 0.68). Finally, this native radar data-based model constitutes a promising alternative for discharge forecasting in remote mountain regions where ground monitoring is scarce and hardly available.

Keywords native radar data; discharge forecasting; machine learning; mountain region; Andes; X-band; radar reflectivity; radar rainfall

---

5.1 Introduction

Discharge forecasting is of main importance for water management and decision-making support all around the globe. Drought and flood events can adversely affect the normal operation of water supply, irrigation, and hydropower systems, and produce socio-economic and ecological damages. Therefore, discharge modeling and forecasting are crucial and have been largely studied in the literature employing different approaches based on physical processes and data-driven techniques (Fatichi et al., 2016; Mosavi et al., 2018; Paniconi and Putti, 2015; Valizadeh et al., 2017; Yaseen et al., 2015). Distributed and semi-distributed models are by far the most adopted and well-known rainfall-runoff models used for discharge forecasting. Several studies (Fatichi et al., 2016; Paniconi and Putti, 2015) have proven their efficiency and good performance at different catchment scales and often at daily or monthly lead times. Nonetheless, these models often come with high-computational costs derived from the increase of the spatial resolution of relevant variables used in the model as in Heuvelink et al. (2020). Furthermore, a high density and even distribution of rain gauge networks are a critical condition for their successful implementation (Paz et al., 2020). Unfortunately, remote

and mountainous regions are usually scarcely monitored which restrict the use of distributed models for discharge forecasting. This is due to the necessity of spatially detailed description of several hydro-geomorphological variables on the study area, which are frequently limited or non-existing. This is especially true in regions such as the Andean tropical mountains. Here, Sucozhañay and Célleri (2018) assessed the impact of using interpolated maps generated with different density of rain gauges by means of the HBV (Hydrologiska Byrans Vattenbalansavdelning)-light semidistributed model in a small Andean catchment. The authors found that the location and number of rain gauge highly influenced the model uncertainty.

As an alternative for capturing a detailed spatial distribution of rainfall, data provided by satellite and meteorological radar has emerged as a solution (Berne and Krajewski, 2013; Editorial Board, 2015; Li et al., 2016). These data sources can effectively be blended with Numerical Weather Predictions (NWP) as in Yoon (2019) or assimilation methods as in Khaki et al. (2019). Moreover, the use of radar rainfall estimates (McKee and Binns, 2016) for streamflow forecasting has been extensively documented (Abon et al., 2016; Chen et al., 2016; He et al., 2013; Hsu et al., 2019; Keblouti et al., 2013) with satisfactory outcomes. The added value of high-resolution radar data in comparison to the use of interpolated rain gauge maps has been demonstrated when the precipitation is highly variable in space (Emmanuel et al., 2015; Lobligeois et al., 2014) and in situations of a decreasing rain gauge density (He et al., 2013). Recently, Mejía-Veintimilla et al. (2019) used bias-adjusted radar rainfall maps over both a distributed and a semi-distributed runoff model for discharge forecasting in a very small catchment (5 km^2) in the Andean region. Despite its good results ($0.6 < R^2 < 0.92$), an exhaustive pre-processing of radar imagery — in addition to ancillary data (e.g., rain gauges, vegetation, and soil maps) — was mandatory. Unfortunately, the availability of hydro-geomorphological data — well distributed in space — is extremely unusual in Andean catchments.

On the other hand, new attempts of using machine learning models for discharge forecasting have been pursued in the last decade with encouraging results. Some authors (Mosavi et al., 2018; Valizadeh et al., 2017; Yaseen et al., 2015) provide reviews of the application of machine learning-based models for this purpose. Several streamflow forecasting studies (Dinu et al., 2017a,b; Falck et al., 2018; Ogale and Srivastava, 2019; Ragettli et al., 2017) use data-driven models employing radar-derived rainfall as input; this requires a preliminary step for transforming reflectivity (native ground radar variable) into rainfall rate (McKee and Binns, 2016). Either way, derivation of radar rainfall estimates remains an intensive task that needs to be tackled before using ground radar data for discharge forecasting.

From the multiple types of machine learning techniques, the Random Forest (RF) algorithm has received particular attention of hydrologists (Tyralis and Papacharalampous, 2017) due to its less complex implementation for machine learning non-specialists compare to other methods. For instance, RF does not need a pre-processing of the inputs; its number of hyper-parameters to tune is far less than the widely used Artificial Neural Networks (ANN); the interpretation of the algorithm as a sequence of binary decisions is more intuitive than other black-box models that result on complex mathematical equations; just to name a few. An RF model has been recently tested by Muñoz et al. (2018) for runoff forecasting in an Andean mountain catchment and performed well with point-scale data. Nonetheless, the recent availability of X-band radar imagery on the site due to the implementation of a radar network in southern Ecuador (Bendix et al., 2017), RadarNet-Sur, opens a new data source for the model. Some studies (Célleri et al., 2007; Guallpa et al., 2019; Rollenbeck and Bendix, 2011) already revealed the high spatial and temporal variability of precipitation in the Andean mountains of southern Ecuador. Moreover, rainfall derivation of the highest radar in the network was performed successfully (Orellana-Alvear et al., 2017, 2019) by using a RF model and reflectivity related inputs. Here, Orellana-Alvear et al. (2019) proposed a new alternative for radar rainfall derivation that reduced the complexity of using the traditional step-wise correction approach which is very often tailor-made depending on the specific site (e.g., Lo Conti et al., 2015; van de Beek et al., 2016). However, due to the complex orography in the Andes and the limitations of single polarized X-band technology, the derivation of accurate rainfall rates still remain quite challenging. Altogether, this brings out the constraints and intensive efforts needed for generating adequate radar rainfall estimations for discharge forecasting models in mountain regions.

Due to the nature of data-driven models, one could expect their ability to map discharge forecast estimations from the reflectivity variable itself — without further pre-processing. For instance, Chaipimonplin et al. (2011) used raw radar reflectivity as input for building an Artificial Neural Network for streamflow forecasting up to 24 h in advance. While the study accomplished good results, no other attempts have been reported which may be a result of the complexity of training a neural network. Thus, to the knowledge of the authors native radar data (reflectivity) has still not been explored as input data in any RF model application for runoff forecasting. This could have enormous implications to overcome the need for extensive and complex processes to obtain radar rainfall estimations, which are normally applied before they can be used in streamflow forecasting.

Consequently, this study aims to compare the forecasting performance of RF models trained with either quasi-native radar data (i.e., corrected reflectivity values) or radar-derived rainfall. In addition, the influence of a proxy variable that resembles the soil moisture in the catchment will be

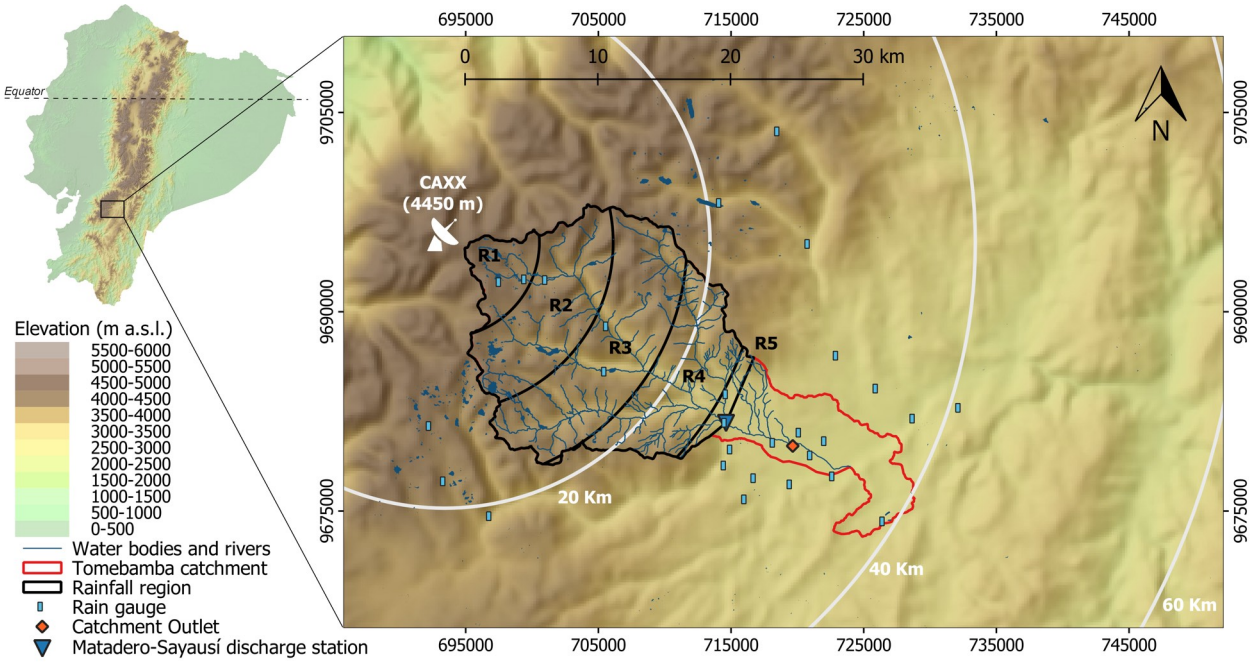

Figure 5.1: The Tomebamba catchment and delimitation of five rainfall regions at the Matadero-Sayausí discharge station.

evaluated through several metrics of goodness of fit.

5.2 Materials and Methods

5.2.1 Study site

The study was conducted in the Tomebamba catchment, located at the southern Andes of Ecuador. The catchment has an approximate extension of 300 km^2 at the Matadero-Sayausí hydrological station as seen in Figure 5.1 and it is representative of medium scale mountain catchments. The elevation ranges from 2800 to 4100 m a.s.l. The city of Cuenca lies at its outlet and, thus, it is exposed to a high risk of floods. Mountain catchments in the tropical Andes provide vital watershed services to downstream communities (Hamel et al., 2018). In this case the most important are potable and irrigation water supply. Additional details of the catchment climatology can be found in Muñoz et al. (2018). Figure 5.1 illustrates the study site and the instruments used in the current investigation.

5.2.2 Instruments and data

5.2.2.1 Radar

Data from a rainfall radar located at 4450 m a.s.l. on the Paragüillas peak in the north border of the Cajas National Park in southern Ecuador was used in the current study. The radar has a maximum range of 100 km and provides 2D polar images of raw reflectivity records as described in Orellana-Alvear et al. (2019). The bin resolution is 2 degrees' azimuth and 100 m range, therefore a matrix of 180x1000 is obtained every 5-min. Data are available from the period April 2015 to June 2017.

Radar data was used with the purpose of obtaining two different spatio-temporal series: one of radar reflectivity and another of radar rainfall retrievals, both at hourly scale. These data were used to derive the inputs of the models. The time series of reflectivity records was obtained from the application of a quality control procedure. For this, specific corrections related to both the clutter and attenuation effects over the native reflectivity records were performed. The reflectivity correction was achieved through the methodology used by Orellana-Alvear et al. (2019) that deals with static and dynamic clutter. This allowed to remove unrealistic reflectivity measurements (e.g., spikes and at certain point attenuation issues) from the radar images. Henceforth for simplicity we will refer to the resulting time series as native radar data.

The series of radar rainfall depth was obtained by using the step-wise correction model used in Orellana-Alvear et al. (2019) that applies clutter and attenuation correction. In addition, a bias adjustment was applied by using the rain gauge data defined below in section 5.2.2.2. For this, the multiplicative bias correction model described in Goudenhoofdt and Delobbe (2009) was used. Finally, data was transformed from rainfall rate to rainfall depth. This entire process resembles the usual efforts of getting the best quantitative estimations of rainfall maps to be used in distributed models. Finally, an accumulation to hourly scale was performed for both spatio-temporal series, reflectivity, and rainfall depth from the 5-min resolution records.

5.2.2.2 Rain Gauges

A rain gauge network, comprised by 28 rain gauges of 0.1 mm resolution each, was used in this study. The rain gauges are unevenly distributed over the study site and its surroundings as illustrated in Figure 5.1. A time series of hourly rainfall depth was obtained for every rain gauge. Data from the period April 2015 to June 2017 were used to adjust the radar rainfall retrievals. Data were

quality checked and operational monitoring was frequently performed during the study period.

5.2.2.3 Discharge

Discharge is measured in the Tomebamba river at the Matadero-Sayausí station located at 2693 m.a.s.l. (see Figure 5.1). Mean hourly discharge, calculated for the period 1997-2017, is 7 m^3 s^{-1} at this hydrological station which was permanently monitored to ensure data quality. Data from the period April 2015 to June 2017 at hourly time step was used as target variable for the models. All extreme discharge values found in the historical data were kept for the analysis. Available data was divided in training and testing periods for the modeling process. Thus, data used in the training phase corresponded to April 2015 to June 2016, while the last year — from July 2016 to June 2017 — was used for the test phase. The latter constitutes an independent dataset that was not used for training the model. Consequently, it ensures the model's performance evaluation on unseen data.

5.2.3 Methods

5.2.3.1 Random Forest model for Discharge Forecasting

Random Forest (RF) is a decision tree-based model from the machine learning family. It is an ensemble method, which means that several trees (a forest) are built and the predicted estimation is the combination of the results of all tree models in the forest. The RF algorithm for regression derives n datasets of random samples with replacement (bootstrapping) from the original dataset. Then, these new datasets are used to build n trees, which ensures that a different subset is used for the construction of each decision tree of the model. A fraction of the data (out-of-bag, OOB), usually the 33%, is used for an internal validation process, which acts as independent data from the training process allowing to obtain unbiased estimations of the regression error. For building each regression tree, a random subset of N predictors (features) is used to create the binary rule at each node of the tree. The selection of the feature used for the binary rule at each step is based on the sum of square residuals. The tree is expanded until a certain depth has been reached (depth of the tree). Then, observations of the OOB subset are evaluated in each constructed tree; and the average of all estimations from the trees is the final estimation of the specific RF model. Finally, the performance of the RF model at the training stage is evaluated with a OOB score (i.e, metric of goodness of fit, here the coefficient of correlation R^2) by comparing the estimations of the RF model and the corresponding observations of the OOB dataset. Therefore, n trees (i.e, number of trees), N

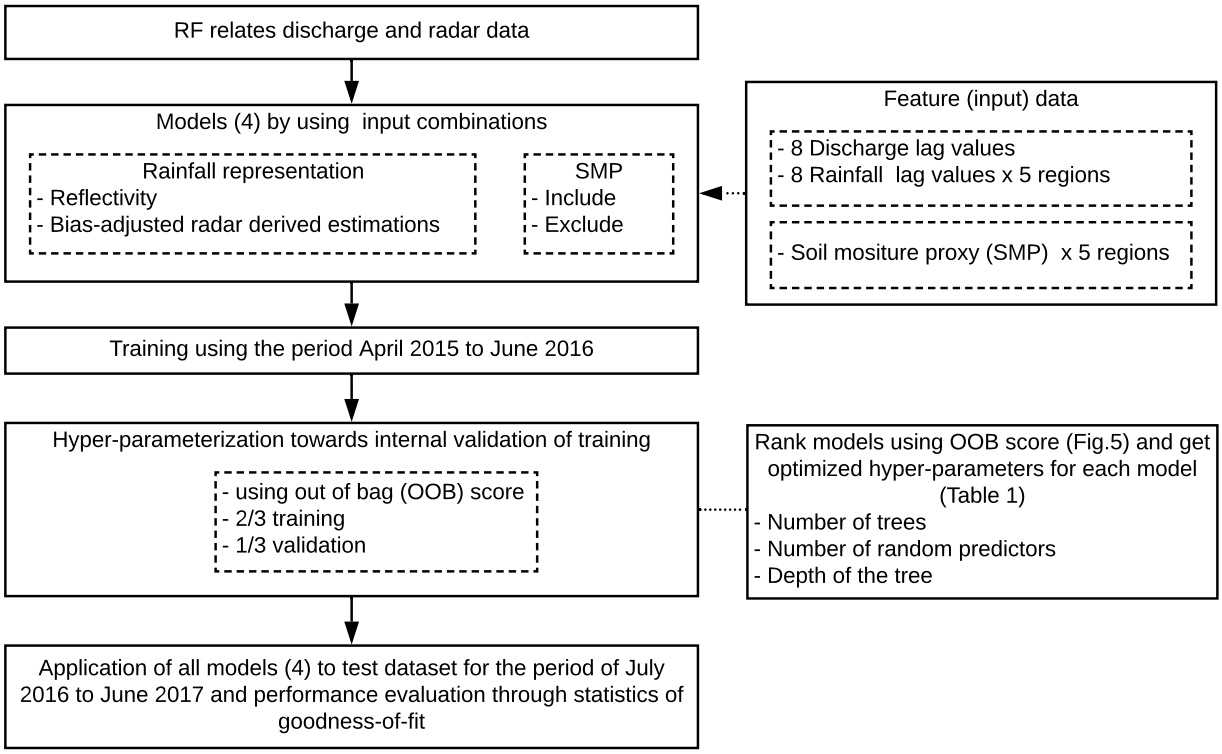

Figure 5.2: Workflow of the implementation of the RF models in the current study.

features (i.e, number of random predictors used at each node to construct each tree) and the depth of the tree (i.e, limit of nodes in each tree) act as hyper-parameters that need to be optimized on the model, which is accomplished by a grid search approach that uses the OOB score for ranking all possible RF models.

For a more detailed explanation of how RF works the reader is referred to Breiman (2001) in general, and to Orellana-Alvear et al. (2019) in an application to this study region in a related work. The RF algorithm was used to develop discharge forecasting models with a lead time of four hours. A flowchart of the models' implementation is provided in Figure 5.2. The following sections provide details about input data and configurations of these models.

5.2.3.2 Input data

To reduce the complexity of the spatial representation of the radar data, we decided to delineate several regions where each one can be considered as a virtual rain gauge. Then, every time series is derived from the spatial mean of the region at each hourly time step. The definition of the regions

is given by considering two factors: (i) The spatial distribution of the virtual rain gauges should resemble the most common distribution of rain gauges within mountain catchments (i.e., along the altitudinal gradient) and (ii) we assume that each virtual rain gauge corresponds to a rainfall region which could be influenced due to the altitude. Thus, we divided the Tomebamba catchment in five regions from the headwater to the outlet by drawing concentric circles of increasing radius of 5565 m. We cropped the last region (R5) up to the Matadero Sayausí discharge station. This small area coincides with the urbanized area of the catchment (see Figure 5.1). By using this regionalization, rainfall regions from R1-R5 have mean elevations of 4027, 3867, 3736, 3463, and 3033 m respectively.

The simplification of the radar image into rainfall regions is required for reducing the number of inputs to the model; the use of all radar pixels as individual inputs would produce high dimensionality issues that will increase the computational costs. We are aware that this regionalization scheme may be improved; however, this is out of the scope of this study whose main objective is to evaluate the potential of leveraging native reflectivity records from meteorological radars on data-driven discharge forecasting modeling.

Inputs for the model are commonly determined by analyzing the target variable (i.e., discharge) and its relation with lagged values from related time series (e.g., discharge and precipitation; Tyralis and Papacharalampous, 2017). This process allows retaining the most relevant observations in the model. For this, different analyses of correlation were carried out between the target variable and the derived-rainfall radar time series.

First, an autocorrelation and partial correlation analyses, as in Muñoz et al. (2018), were applied from the autocorrelation function (ACF) and the partial autocorrelation function (PACF) for identifying the number of previous discharge observations that mostly influenced the discharge at the time of interest. Derivation of ACF and PACF has been already documented in studies related to data-driven models for discharge forecasting (Sudheer et al., 2002). Both analyses serve as complementary tools to define the number of lags (hours) to be included in the model. The ACF allows identifying autoregressive patterns in the time series, while the PACF is aimed at identifying the extent of the lag influence. Second, a cross-correlation analysis between the discharge and, either the reflectivity or rainfall depth, allowed finding the number of influential lags from the virtual rain gauge time series. Finally, the use of an additional input that represents the soil moisture condition into a data-driven hydrological model has been previously documented (Ba et al., 2018; Jadidoleslam et al., 2019; Javelle et al., 2010) and proven to be efficient. Thus, in order to account for the soil moisture condition of the catchment, a 3-day precipitation accumulation was considered

to be proxy. Therefore, a 3-day proxy per region was calculated through the virtual rain gauge time series. It means that the 3-day precipitation accumulation is derived by either the sum of radar reflectivity values or the radar rainfall estimates depending on the data type of the time series. We selected a 3-day window due to rainfall data availability. Since this accumulation is calculated from the observed radar imagery which has not been filled in any way (e.g., interpolation of missing radar images as a result of technical maintenance), we need a continuous time frame for such derivation that allows obtaining an acceptable number of instances for training the model.

Finally, it is worth mentioning that usually several years of rainfall-runoff records are needed for streamflow forecasting modeling when using physical-based models. However, in this study we are able to exploit the limited data of a very short time (2.5 years) because there are not aquifers in our study catchment which means discharge is mainly controlled by rainfall. This allowed us to obtain a good number of events for data-driven discharge forecasting for most of the flow duration curve.

5.2.3.3 Input Data Configuration and Model Optimization

Four models were defined based on different input data scenarios. Input data configuration differs in two conditions: (i) data type of the time series (e.g., reflectivity (dBZ) or derived-rainfall radar (mm)) and (ii) inclusion or exclusion of the 3-day proxy variable. Therefore, four possible models were evaluated in the current study through the combination of the conditions mentioned above.

It should be pointed out that the number of lags resulting from the runoff and virtual rain gauge time series remained equal for all input data configurations. Thus, only the data representation (dBZ or mm) and the addition or omission of the 3-day proxy modify the inputs to the models.

The hyper-parameters of each model (i.e., number of trees, number of features and tree depth) were optimized using a grid search approach during the model training process. Thus an independent set of optimized hyper-parameters was found for each model configuration. This ensured that each model accomplished the best possible use of its own input data.

5.2.3.4 Performance evaluation

An independent evaluation was performed after the models were trained and optimized. For this, a discharge prediction of each model was obtained at each time step (i.e., four hours) of the one-year data used for the test phase. Afterwards, the forecasted discharges were compared with the corresponding observations through several metrics of goodness of fit. Thus, Root Mean Squared Error (RMSE), Percentage Bias (PBIAS) and Mean Absolute Relative Error (MARE) were

calculated. In addition, model performance statistics, which are commonly used in hydrological studies, were derived. These include the Nash-Sutcliffe efficiency (NSE), the Kling-Gupta efficiency (KGE) (Gupta et al., 2009) and its modified version (KGE') (Kling et al., 2012) as described in Equations 5.1,5.2,5.3

$$NSE = 1 - \frac{\sum_{i=1}^{n}(Q_{sim_i} - Q_{obs_i})^2}{\sum_{i=1}^{n}(Q_{obs_i} - \overline{Q}_{obs})^2} \tag{5.1}$$

where n is the length of the evaluated time series, Qsim$_i$ and Qobs$_i$ is the simulated and observed discharge at time i respectively, and $\overline{Q}$obs is the mean of the observed values.

$$KGE = 1 - \sqrt{(r-1)^2 + (\beta-1)^2 + (\alpha-1)^2} \tag{5.2}$$

$$KGE' = 1 - \sqrt{(r-1)^2 + (\beta-1)^2 + (\gamma-1)^2} \tag{5.3}$$

where r is the Pearson correlation coefficient, β is the ratio between the means of the simulated values and the observed ones, α is the ratio between the standard deviations of the simulations and observations. Similarly, in Eq.5.3 γ is the ratio between the coefficients of variation (CV) of the simulated values to the observed ones.Thus the decomposition of the KGE equations into correlation (r), bias (β) and variability (α / γ) facilitates to identify the relative importance of each independent metric over the derived KGE value.

5.3 Results and Discussion

In the following, the analysis of the models' construction as well as their optimization process is documented. Moreover, the discharge forecasting models are evaluated and compared at first in terms of the radar data type, reflectivity, or rainfall depth, used as input for each model and secondly, regarding the use of the soil moisture proxy (i.e., 3-day precipitation accumulation) as additional input into the models.

5.3.1 Feature Selection and Model Optimization

The number of precipitation and discharge lags for the inputs of the model were determined as eight lags (hours) for both discharge and precipitation variables. As seen in Figure 5.3, the 95%

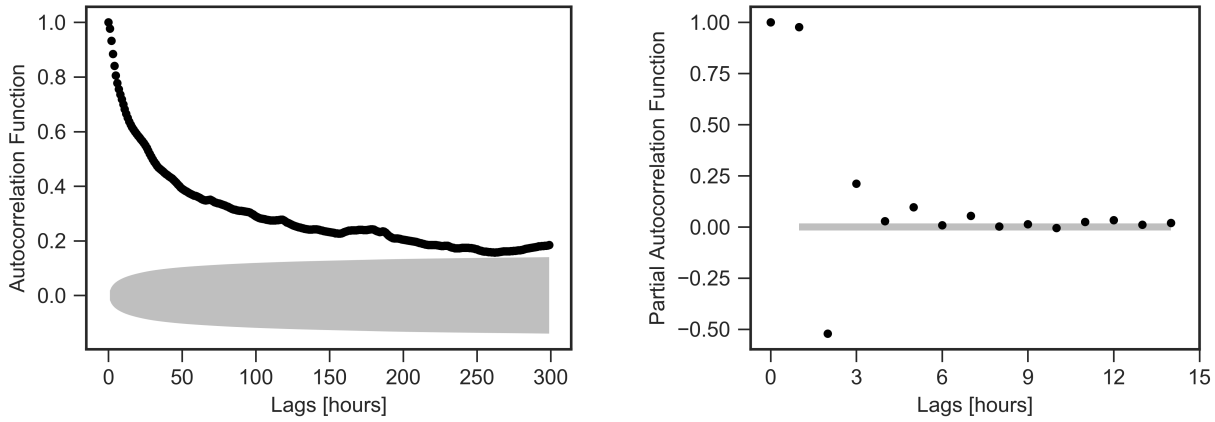

Figure 5.3: Autocorrelation function (ACF) and Partial autocorrelation function (PACF) of the Matadero-Sayausí discharge series. Gray hatch indicates the 95% confidence band.

confidence interval (i.e., values out of the gray area are considered very probably a correlation and not a statistical chance) from the ACF in the left hand side reveals a correlation up to around 250 lags showing a dominant autoregressive pattern. In addition, the results of PACF illustrated in the right hand side of Figure 5.3 reveal no significant correlation from lag 8. These results are in agreement with Muñoz et al. (2018) that used a slightly larger dataset of this discharge time series.

Furthermore, Figure 5.4 illustrates the Pearson cross-correlation analysis between the discharge time series and the one derived from the native radar data as a virtual rain gauge. Similar results were found when using the adjusted radar data. It can be seen that the highest correlation is found between lags 4 and 8 depending on the virtual rain gauge (i.e., rainfall region). Muñoz et al. (2018) already discussed the relation of lag 4 to the mean concentration time of the catchment. However, in contrast to that study, we decided to keep eight lags from the precipitation variable because of the slight variability among rainfall regions. R1 has a lower correlation which could be related to three aspects: (i) the presence of few tributaries in this region; (ii) its smaller area that drains to a location where there is an important presence of lakes that heavily affect water flow and transit times and; (iii) rainfall in the headwaters falls mainly as drizzle. Nonetheless, we maintained this layer because we aimed to build the models as simple and homogenous as possible, so we can mainly focus on the input data-type influence (i.e., native or adjusted radar data). In summary, the inputs for each model correspond to the values of the eight discharge lags (hours) and eight derived precipitation (or reflectivity) lags. The latter applies for every region (R1-R5). Thus we defined [8+(8x5)] inputs for modeling, which could increase in 5 due to the soil moisture proxy (i.e., one soil moisture proxy

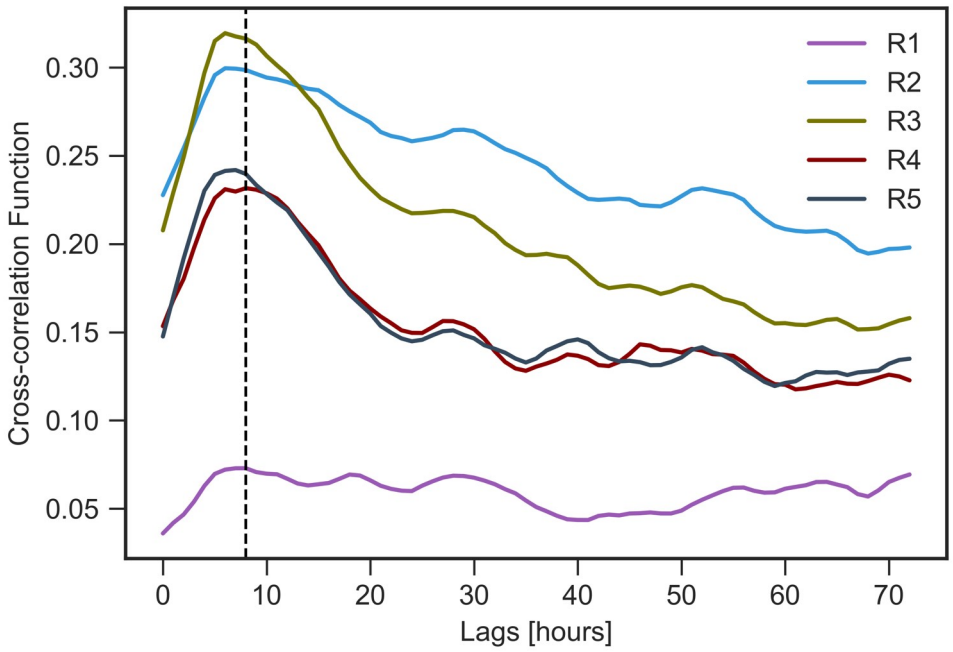

Figure 5.4: Pearson cross-correlation comparison between the different precipitation time series (derived from the rainfall regions R1-R5 representing the virtual rain gauge stations) and the Matadero-Sayausí discharge station. The dashed line denotes lag 8.

per region).

Figure 5.5 shows the convergence of the best hyper-parameters for all forecasting model configurations at 4 h in advance. The distribution of the OOB score for each number of trees follows from the results of the combination with different number of features and depths of the tree. It can be seen that both models using adjusted radar data have smaller variability than the models using native radar data. OOB score slightly improves when using adjusted radar data ($\sim$0.04). Optimized hyper-parameters corresponding to the number of trees, number of features and depth of the tree, as well as OOB scores for the training phase are shown in Table 5.1.

5.3.2 Performance Evaluation of Discharge Models with Test Data

In the following, the analyses are performed based on the evaluation of all four models by using the test dataset. The correlation between forecasted and observed values is illustrated in Figure 5.6 while the performance statistics of the models are summarized in Table 5.2.

Figure 5.6 points out to slight differences between the models. The 95% confidence interval of

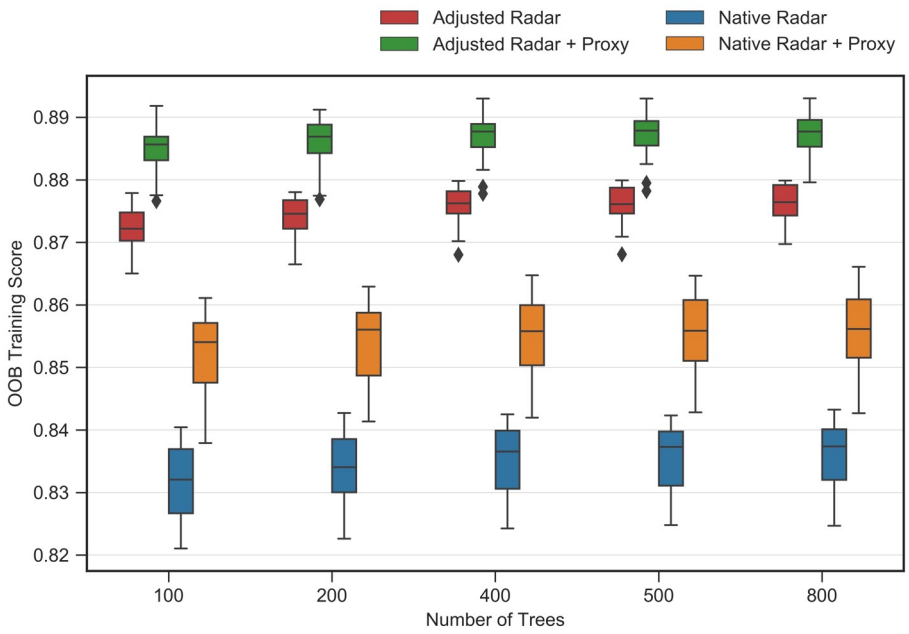

Figure 5.5: Evolution of the OOB score for different configuration models.

the regression for the native radar data-based models is wider than their counterparts. Moreover, it can be observed that the scatter at all models starts to decouple at 25 m^3 s^{-1}, with two branches, a lower and a higher. Finally, the lower retains all measurements higher than 50 m^3 s^{-1} which means that all models underestimate the discharge above this threshold. These observations of high discharge values will be discussed in detail later on.

Adjusted radar data-based models are in general slightly better than the native radar data-based models as shown in Table 5.2. Nonetheless, the component r from the KGE index reveals that differences in the models' performance are mainly related to the linear correlation between forecasted and observed values. Thus, for the adjusted radar data-based models r reaches 0.87, while for native radar data-based models $0.81<r<0.83$. The PBIAS oscillates around 10% for all models, while the MARE values increase up to 30% because of the absolute value of the error. Unexpectedly, the performance (NSE, KGE and KGE') of the adjusted radar data-based model that uses the moisture proxy as additional input is equal or lower than the one that do not use the moisture proxy. In contrast, the native radar data-based model has, in general, a consistent improvement at all evaluation metrics when the proxy is included. This points out to the inherent error related to the rainfall adjustment process from the reflectivity values toward the bias correction

Table 5.1: Optimized hyper-parameters for the discharge forecasting models and their OOB score at the training phase.

Model	N trees	N features	Depth of tree	OOB Score
Adjusted	400	18	40	0.88
Adjusted + Proxy	400	30	40	0.89
Native	400	18	30	0.83
Native + Proxy	400	36	30	0.85

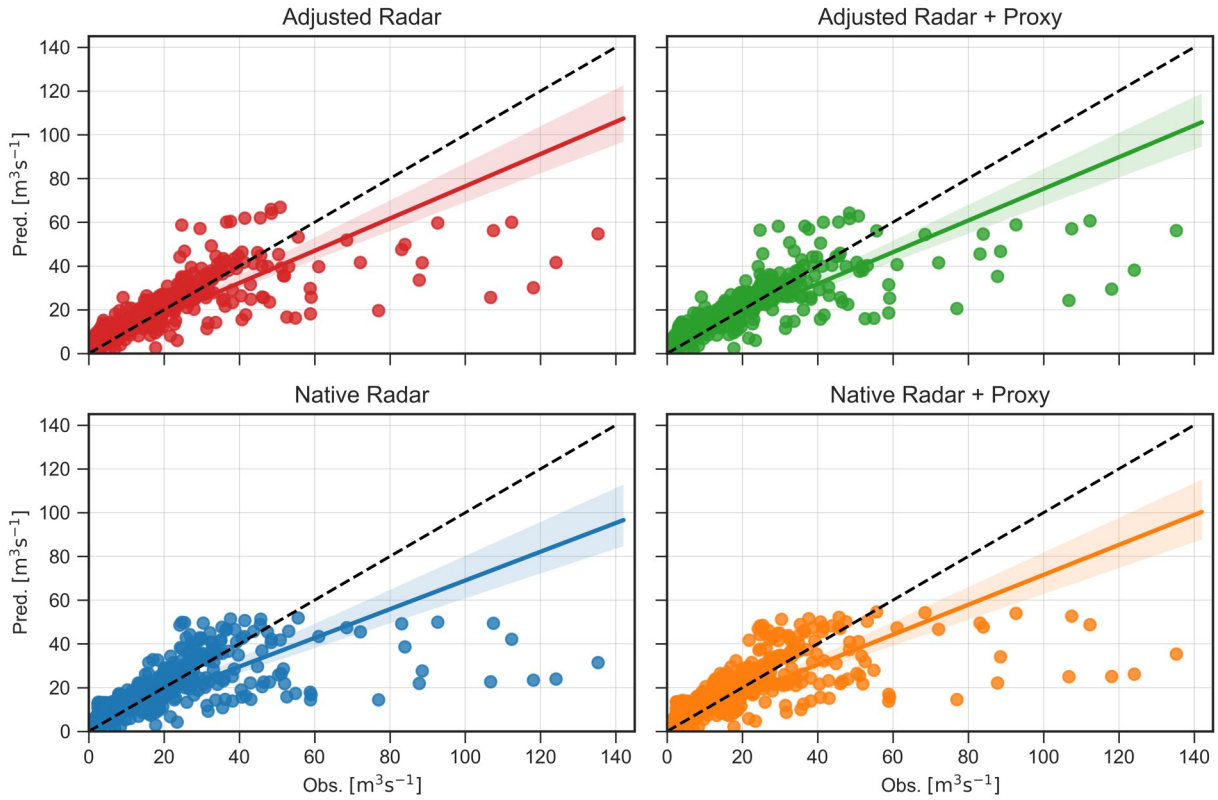

Figure 5.6: Correlation between observed and forecasted discharge of the four different model configurations. The bisector line is showed in hatched black. The continuous lines denote the linear regressions and the shadow areas represent the 95% confidence interval band of each regression respectively.

based on the rain gauge data, which is present in the adjusted radar data-based models. It may be the case that while certain spatial points are properly adjusted regarding the rainfall quantities, this process also adds errors (i.e., inadequate bias correction) to other regions of the radar imagery.

Table 5.2: Performance of all discharge forecasting models for the test period. *Metrics for the data subset where observations are higher than 50 m^3 s^{-1} are not shown due to the low number of samples.

Model	Data*	RMSE	PBIAS	MARE	NSE	Original KGE				Modified KGE	
						KGE	r	β	α	KGE'	γ
Adjusted	All	5.38	10.02	0.25	0.75	0.78	0.87	1.10	0.85	0.72	0.77
Adjusted + Proxy	All	5.33	9.87	0.25	0.75	0.77	0.87	1.10	0.83	0.71	0.76
Native	All	6.23	9.62	0.30	0.66	0.72	0.81	1.10	0.81	0.66	0.74
Native + Proxy	All	6.00	10.38	0.29	0.68	0.73	0.83	1.10	0.82	0.68	0.75
Adjusted	<50 m^3s^{-1}	3.08	16.26	0.22	0.84	0.81	0.94	1.16	1.08	0.81	0.93
Adjusted + Proxy	<50 m^3s^{-1}	3.04	17.27	0.24	0.85	0.81	0.94	1.17	1.04	0.79	0.89
Native	<50 m^3s^{-1}	3.47	17.06	0.26	0.80	0.80	0.92	1.17	1.05	0.79	0.90
Native + Proxy	<50 m^3s^{-1}	3.53	20.51	0.29	0.80	0.77	0.92	1.21	1.05	0.75	0.87

This in turns obscures the influence of the proxy within the adjusted radar data-based models. It is because the soil moisture proxy is indeed dependent of the adjusted rainfall and while the proxy may improve the discharge forecasting for some rainfall events, it may also have the contrary effect.

At first sight, the performance statistics of the models depicted in Table 5.2 seems slightly lower than other related studies in Andean regions such as Mejía-Veintimilla et al. (2019) that accomplished a 0.77<NSE<0.80 for three different runoff events by using radar data. However, by exploring the discharge observations higher than 50 m^3 s^{-1} depicted in Figure 5.6, it turns out that they correspond only to 1% of the test dataset. From this sample, all observations higher than 60 m^3 s^{-1} (67% of the data) belong exclusively to two very strong rainfall events that occurred in 2017. When these very extreme events are omitted, the performance statistics improve substantially. For instance, for the Adjusted + Proxy model, NSE increases from 0.75 to 0.85, KGE increases from 0.77 to 0.82 and KGE' improves from 0.71 to 0.80. Similarly, for the Native + Proxy model, the same metrics improved from 0.68 to 0.81, 0.73 to 0.80 and 0.68 to 0.79 respectively. Interestingly, for both models the γ component of KDE' improves to $\sim$0.90 which means that the CV between observations and predictions is comparable. Despite the fact that our results can not be directly compared to those of Mejía-Veintimilla et al. (2019) that used distributed models, it is worth mentioning that the random forest models in the current study have been evaluated toward the entire discharge time series and show satisfactory results (KDE$\sim$0.80, NSE>0.80) for discharges lower than 50 m^3 s^{-1}. This points out to the leverage of the radar data (even as native variable) and the usefulness of the simplified structure of the random forest models. Altogether, the differences in the performance of the random forest models are even smaller when considering discharge values lower

than 50 m^3 s^{-1}. Therefore, it confirms that the use of native radar data as inputs for the models are able to produce as good results as those based on adjusted radar data.

Independently of the hydrological model, very high runoff events are always challenging to simulate (Abon et al., 2016; Ovando et al., 2016). In our study site, discharge of 50 m^3 s^{-1} is exceeded less than 5.4% of the time by considering a long-term series of more than 23 years at Matadero-Sayausí station. Nonetheless, in order to learn from these particular events and provide a potential explanation of the diminished performance of the models, we investigated from the radar imagery the event of 2017.04.13 that produced discharge values higher than 60 m^3 s^{-1} (see Figure 5.7). It is evident that a strong radar signal attenuation may compromise the reflectivity records for several hours of the event (i.e., heavy rainfall cores closer to the radar site). On the other hand, the rainfall adjustment based on rain gauge data at ground may not properly describe the spatial distribution and size of the strongest rain cells. Also, at certain time steps the rain cells are occurring very close to the outlet of the catchment, producing a flash discharge response, shorter that the 4 h lead time of our models, which heavily compromise their forecasting capability. In these cases, a precipitation nowcasting should be beneficial to the model in order to anticipate heavy rainfall events as in Heuvelink et al. (2020). In addition, the lower performance of all models for these extreme events is probably due to the small number of training samples of high discharge values (>50 m^3 s^{-1}). As these events are infrequent, their particular characteristics regarding the velocity of the storm and spatial occurrence are not properly learned by the model. In addition, Guallpa et al. (2019) showed that fast storms in the southern Andes of Ecuador need to be recorded by ground radar at higher temporal resolution (e.g., 1-min), otherwise the reliability of the reflectivity records and in consequence radar-derived rainfall was compromised. Moreover, as the RF discharge estimation is the result of the average of all predictions from the trees, it tends to underestimate high discharge values. Finally, it is worth mentioning that despite a less accurate forecast was accomplished for discharge peaks higher than 50 m^3 s^{-1}, local emergency services can still be benefited from the discharge forecasting developed in this study. It is because a flooding alert can already be emitted once the forecast has exceeded a certain threshold without the need of a very accurate quantitative estimation.

5.3.3 Data Type Influence

Figure 5.8 shows that the performance of the models that used native or adjusted radar data is very similar. Although the use of adjusted radar data slightly improves the runoff forecast, main

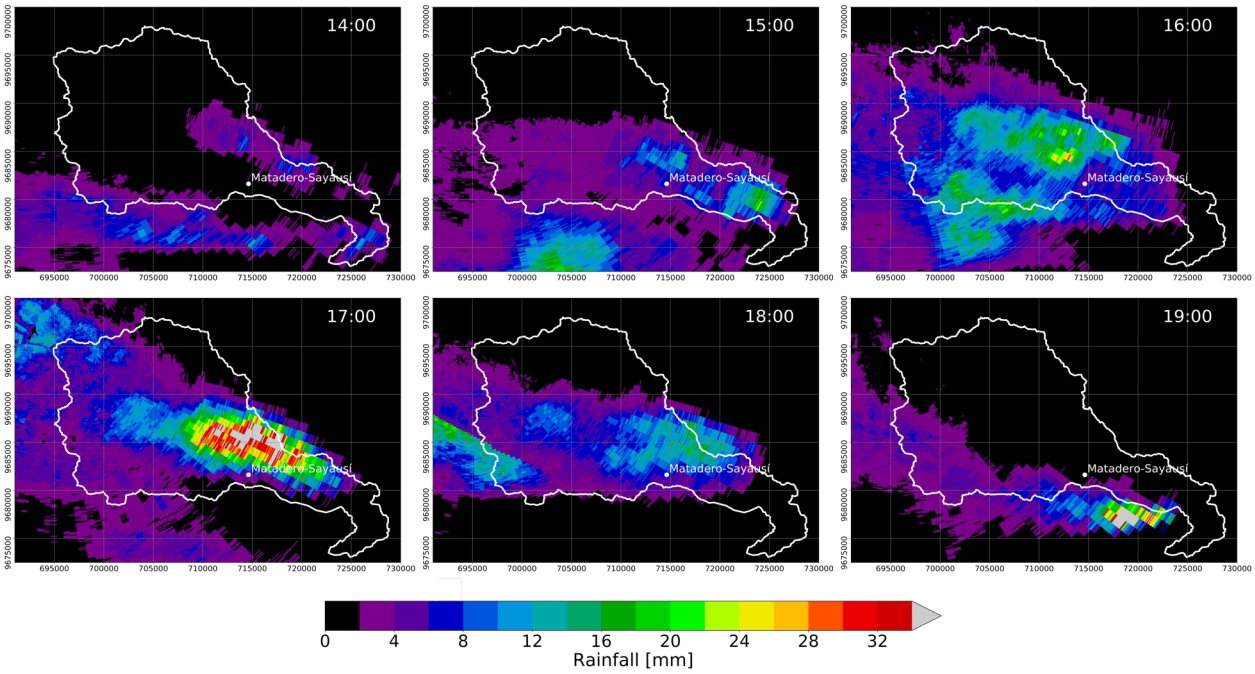

Figure 5.7: Hourly rainfall images from adjusted radar estimates corresponding to the rainfall event of 2017.04.13 (local time).

differences occur in high discharges (> 50 m^3 s^{-1}). Below this threshold, the native radar data-based model denotes a lower scatter around the regression line whereas the adjusted radar data-based model tends to overestimate the discharge forecast. It could be a result of the added noise while applying the step-wise correction process to derive the adjusted rainfall radar data. On the other hand, better forecasts of higher discharge values may be benefited for the adjustment (bias removal) of rain rates by using the rain gauge records. For instance, by considering only the runoff observations higher than 50 m^3 s^{-1}, the RMSE (PBIAS) from the Adjusted Radar + Proxy model are 44.46 ($\sim$-47%), while in contrast they are 49.95 ($\sim$-54%) for the Native Radar + Proxy model.

Altogether, the results bear out that the influence of data type might be overlooked (0.72<KGE<0.78; 0.66<KGE'<0.72). This is an interesting finding since the pre-processing for the generation of input data for each model type (native- or adjusted-radar) is quite different in terms of auxiliary data and correction process chain. The use of native radar data would be preferred due to the very few correction steps needed, which are mainly focused on fixing unrealistic observations. More importantly, the use of native data overcomes the problem of not having adequate rain gauge networks for radar image adjustment. This is of utmost importance because mountain remote areas, as the high Andes, are usually of very complex terrain and access.

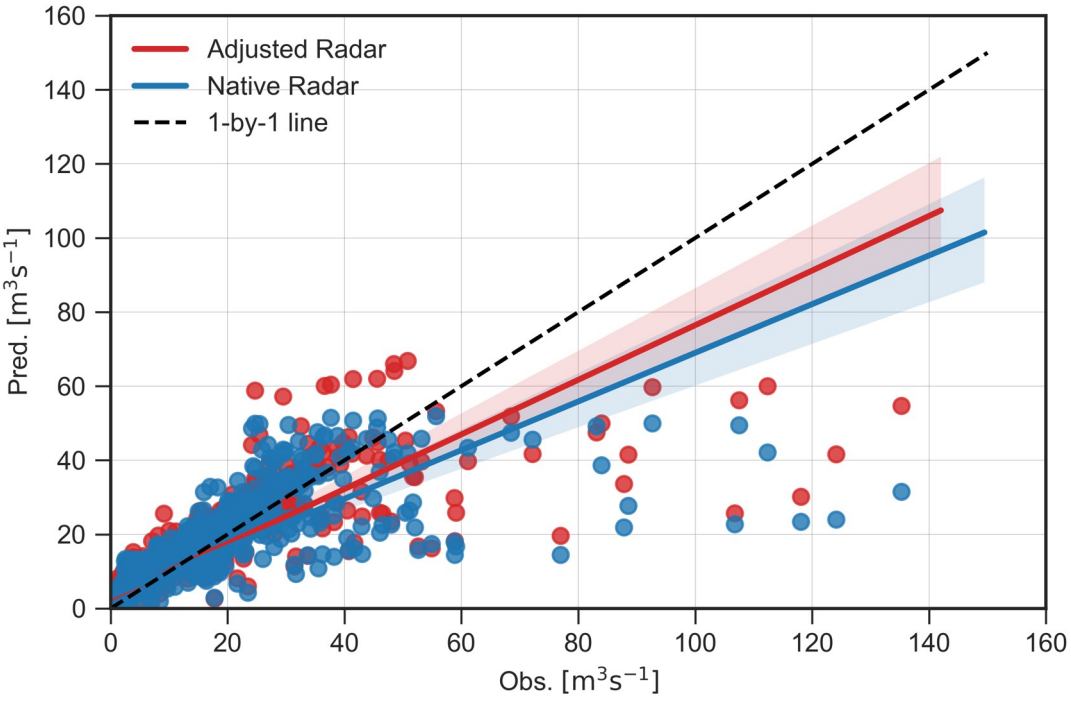

Figure 5.8: Influence of the use of data type (adjusted or native — reflectivity — radar data) in the discharge forecasting models. The bisector line is showed in dotted black; the continuous lines denotes the linear regressions and the shadow areas represent the 95% confidence interval band of each regression respectively.

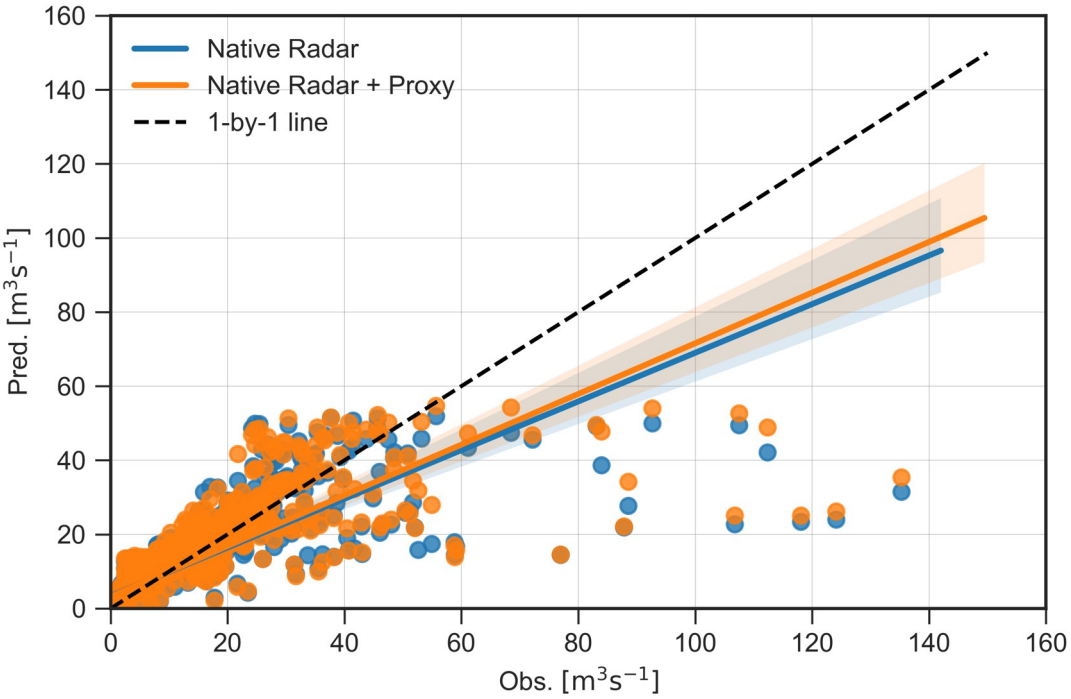

Figure 5.9: Influence of the use of a soil moisture proxy variable in the discharge forecasting models. The bisector line is showed in dotted black; the continuous lines denotes the linear regressions and the shadow areas represent the 95% confidence interval band of each regression respectively.

5.3.4 Proxy of Soil Moisture Influence

A slight improvement in model performance by using the soil moisture proxy is observed in Figure 5.9. Nonetheless, this enhancement seems narrow due to the regionalization of the radar imagery. Other strategies by using soil moisture data derived from satellite such us in Jadidoleslam et al. (2019) or derived from rainfall and evaporation as in Javelle et al. (2010) are not feasible for this catchment due to its relatively small size and need of additional variables. Thus, different strategies of radar data regionalization are encouraged for future research since finding the optimal spatial representation is not the aim of the current study. Moreover, the limitation of the temporal length of the precipitation window (3-day precipitation accumulation) due to the number of samples could also influence the results. An adequate evaluation of an optimal window is also suggested for further studies.

5.4 Conclusions

A discharge forecast model by using a random forest algorithm by means of native radar data (i.e., reflectivity instead of derived rain rate) was developed and evaluated. In addition, a comparison with an equivalent model that used adjusted radar data (i.e., bias-adjusted radar derived rainfall) was performed. The use of an antecedent soil moisture proxy as input for the models was also evaluated. From the results, the following conclusions can be drawn:

(i) Similar goodness of fit was accomplished when using native radar data as well as adjusted radar data as input to the forecasting models. It demonstrates that the use of native radar data as input can properly map the expected discharge quantitatively. This is of great importance and interest since the need for exhaustive data pre-processing for converting the reflectivity values into rainfall depth can be omitted for discharge forecasting modeling applications using data-driven techniques.

(ii) Satisfactory results were obtained from the native radar data-based model (NSE=0.81; KGE=0.80) when evaluated in the discharge time series for values lower than 50 $m^3\ s^{-1}$. The decay in the model performance when considering higher discharge values was identified mainly as a result of two strong rainfall events. Several reasons could influence in the inadequate response of the model: a) strong attenuation within the catchment, which lead to too low mean rainfall values for distant affected zones which coincides with the region close to the catchment outlet, b) the reduced number or training samples for discharge values higher than 50 $m^3\ s^{-1}$ which complicates the RF model during its averaging of the predictions from the trees and c) a suboptimal delineation of the rainfall regions for the virtual rain gauges' derivation.

(iii) The inclusion of the derived soil moisture proxy used in this study (i.e., 3-day accumulation of precipitation) did not improve the model performance significantly, showing only a small increment in NSE at the expense of including an additional predictor. Given that the antecedent soil moisture condition affects the rainfall to runoff formation, more research is needed to address how to include an adequate soil moisture representation in the model configuration.

(iv) The RF model developed in the current study is particularly suitable for its use in (remote) mountain regions where catchments usually remain scarcely monitored. Therefore, models that perform well in absence of (dense) rain gauge networks overcome a strong limitation in the application of simulation tools for disaster prevention through early warning systems. This can be extrapolated to other sites around the globe where where limitations in access restrict the possibility of dense rain gauge monitoring.

In summary, this is a pioneer study that leverages the native radar variable (i.e., reflectivity) from an X-band single polarized radar located in the southern Andes of Ecuador for discharge forecasting by using a RF model. The usefulness of reflectivity records has been confirmed and highlights the benefits of leaving out the complexity of the radar rainfall process derivation. This means that the spatial-distributed rainfall pattern captured from ground radar can be used and applied without further pre-processing. It has enormous implications for remote and mountain regions in the world where additional monitoring can be expensive or even unaffordable and logistically challenging due to access issues. Moreover, the limitations of the RF model in situations with high attenuation (i.e., heavy rainfall events) were shown. This is comparable with the impact that the inadequate spatial distribution of precipitation — as a result of an uneven rain gauge network — has in physical-based models. This points out the relevance of the accuracy of the input data over the model itself. Thus, this study is a first step moving forward to focus the efforts on combining data sources (radar reflectivity and useful bias-adjusted radar rainfall areas) as inputs for discharge forecasting models. Further work will also focus on the improvement of the regionalization of precipitation and derivation of soil moisture through different techniques.

Acknowledgments

We acknowledge the Ministry of Environment of Ecuador (MAE) for providing research permissions. We are grateful to the technical staff that contributed to the meteorological monitoring and particularly setting-up and operational monitoring of the CAXX radar equipment: Ing. Mario Guallpa and Andreas Fries.

References

Abon, C. C., Kneis, D., Crisologo, I., Bronstert, A., Constantino David, C. P., and Heistermann, M. (2016). Evaluating the potential of radar-based rainfall estimates for streamflow and flood simulations in the Philippines. *Geomatics, Natural Hazards and Risk*, 7(4):1390–1405.

Ba, H., Guo, S., Wang, Y., Hong, X., and Zhong, Y. (2018). Improving ANN model performance in runoff forecasting by adding soil moisture input and using data preprocessing techniques Huanhuan. *Hydrology Research*, 49(3):744–760.

Bendix, J., Fries, A., Zárate, J., Trachte, K., Rollenbeck, R., Pucha-Cofrep, F., Paladines, R., Palacios, I., Orellana, J., Oñate-Valdivieso, F., Naranjo, C., Mendoza, L., Mejia, D., Guallpa, M., Gordillo, F., Gonzalez-Jaramillo, V., Dobbermann, M., Célleri, R., Carrillo, C., Araque, A., and Achilles, S. (2017). Radarnet-Sur First weather Radar Network in tropical high mountains. *Bulletin of the American Meteorological Society*, 98(6):1235–1254.

Berne, A. and Krajewski, W. F. (2013). Advances in Water Resources Radar for hydrology : Unfulfilled promise or unrecognized potential ? *Advances in Water Resources*, 51:357–366.

Breiman, L. (2001). Random forests. *Machine learning*, 45(1):5–32.

Célleri, R., Willems, P., Buytaert, W., and Feyen, J. (2007). Space-time rainfall variability in the Paute basin, Ecuadorian Andes. *Hydrological Processes*, 21(24):3316–3327.

Chaipimonplin, T., See, L., and Kneale, P. (2011). Improving neural network for flood forecasting using radar data on the Upper Ping River. In *19th International Congress on Modelling and Simulation*.

Chen, X., Zhang, L., Gippel, C. J., Shan, L., Chen, S., and Yang, W. (2016). Uncertainty of Flood Forecasting Based on Radar Rainfall Data Assimilation. *Advances in Meteorology*, 2016.

Dinu, C., Drobot, R., Pricop, C., and Blidaru, T. V. (2017a). Flash-Flood Modelling with Artificial Neural Networks using Radar Rainfall Estimates. *Mathematical Modelling in Civil Engineering*, 13(3):10–20.

Dinu, C., Drobot, R., Pricop, C., and Blidaru, T. V. (2017b). Genetic Programming Technique applied for Flaash-Flood Modelling using Radar Rainfall Estimates. *Mathematical Modelling in Civil Engineering*, 13(4):27–38.

Editorial Board (2015). Hydrologic applications of weather radar. *Journal of Hydrology*, 531:231–233.

Emmanuel, I., Andrieu, H., Leblois, E., Janey, N., and Payrastre, O. (2015). Influence of rainfall spatial variability on rainfall-runoff modelling : Benefit of a simulation approach ? *Journal of Hydrology*.

Falck, A. S., Maggioni, V., Tomasella, J., Diniz, F. L., Mei, Y., Beneti, C. A., Herdies, D. L., Neundorf, R., Caram, R. O., and Rodriguez, D. A. (2018). Improving the use of ground-based radar rainfall data for monitoring and predicting floods in the Iguaçu river basin. *Journal of Hydrology*, 567(October):626–636.

Fatichi, S., Vivoni, E. R., Ogden, F. L., Ivanov, V. Y., Mirus, B., Gochis, D., Downer, C. W., Camporese, M., Davison, J. H., Ebel, B., Jones, N., Kim, J., Mascaro, G., Niswonger, R., Restrepo, P., Rigon, R., Shen, C., Sulis, M., and Tarboton, D. (2016). An overview of current applications, challenges, and future trends in distributed process-based models in hydrology. *Journal of Hydrology*, 537:45–60.

Goudenhoofdt, E. and Delobbe, L. (2009). Evaluation of radar-gauge merging methods for quantitative precipitation estimates. *Hydrology and Earth System Sciences*, 13(2):195–203.

Guallpa, M., Orellana-Alvear, J., and Bendix, J. (2019). Tropical andes radar precipitation estimates need high temporal and moderate spatial resolution. *Water (Switzerland)*, 11(5):1–22.

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, 377(1-2):80–91.

Hamel, P., Riveros-Iregui, D., Ballari, D., Browning, T., Célleri, R., Chandler, D., Chun, K. P., Destouni, G., Jacobs, S., Jasechko, S., Johnson, M., Krishnaswamy, J., Poca, M., Pompeu, P. V., and Rocha, H. (2018). Watershed services in the humid tropics: Opportunities from recent advances in ecohydrology. *Ecohydrology*, 11(3).

He, X., Sonnenborg, T. O., Refsgaard, J. C., Vejen, F., and Jensen, K. H. (2013). Evaluation of the value of radar QPE data and rain gauge data for hydrological modeling. *Water Resources Research*, 49(September):5989–6005.

Heuvelink, D., Berenguer, M., Brauer, C. C., and Uijlenhoet, R. (2020). Hydrological application of radar rainfall nowcasting in the Netherlands. *Environment International*, 136:105431.

Hsu, S. Y., Chen, T. B., Du, W. C., Wu, J. H., and Chen, S. C. (2019). IntegrateWeather radar and monitoring devices for urban flooding surveillance. *Sensors (Switzerland)*, 19(4):1–15.

Jadidoleslam, N., Mantilla, R., Krajewski, W. F., and Goska, R. (2019). Investigating the role of antecedent SMAP satellite soil moisture , radar rainfall and MODIS vegetation on runo ff production in an agricultural region. *Journal of Hydrology*, 579(June):124210.

Javelle, P., Fouchier, C., Arnaud, P., and Lavabre, J. (2010). Flash flood warning at ungauged locations using radar rainfall and antecedent soil moisture estimations. *Journal of Hydrology*, 394(1-2):267–274.

Keblouti, M., Ouerdachi, L., and Berhail, S. (2013). The use of weather radar for rainfall-runoff modeling , case of Seybouse watershed (Algeria). *Arabian Journal of Geosciences*.

Khaki, M., Hoteit, I., Kuhn, M., Forootan, E., and Awange, J. (2019). Assessing data assimilation frameworks for using multi-mission satellite products in a hydrological context. *Science of the Total Environment*, 647:1031–1043.

Kling, H., Fuchs, M., and Paulin, M. (2012). Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios. *Journal of Hydrology*, 424-425:264–277.

Li, Y., Grimaldi, S., Walker, J. P., and Pauwels, V. R. (2016). Application of remote sensing data to constrain operational rainfall-driven flood forecasting: A review. *Remote Sensing*, 8(6).

Lo Conti, F., Francipane, A., Pumo, D., and Noto, L. V. (2015). Exploring single polarization X-band weather radar potentials for local meteorological and hydrological applications. *Journal of Hydrology*, 531:508–522.

Lobligeois, F., Andréassian, V., Perrin, C., Tabary, P., and Loumagne, C. (2014). When does higher spatial resolution rainfall information improve streamflow simulation? An evaluation using 3620 flood events. *Hydrology and Earth System Sciences*, 18(2):575–594.

McKee, J. L. and Binns, A. D. (2016). A review of gauge-radar merging methods for quantitative precipitation

estimation in hydrology. *Canadian Water Resources Journal*, 41(1-2):186–203.

Mejía-Veintimilla, D., Ochoa-Cueva, P., Samaniego-Rojas, N., Félix, R., Arteaga, J., Crespo, P., Oñate-Valdivieso, F., and Fries, A. (2019). River discharge simulation in the high andes of southern ecuador using high-resolution radar observations and meteorological station data. *Remote Sensing*, 11(23).

Mosavi, A., Ozturk, P., and Chau, K. W. (2018). Flood prediction using machine learning models: Literature review. *Water (Switzerland)*, 10(11):1–40.

Muñoz, P., Orellana-Alvear, J., Willems, P., and Célleri, R. (2018). Flash-Flood Forecasting in an Andean Mountain Catchment - Development of a Step-Wise Methodology Based on the Random Forest Algorithm. *Water (Switzerland)*, 10(1519):1–18.

Ogale, S. and Srivastava, S. (2019). Modelling and short term forecasting of flash floods in an urban environment. *25th National Conference on Communications, NCC 2019*, pages 1–6.

Orellana-Alvear, J., Célleri, R., Rollenbeck, R., and Bendix, J. (2017). Analysis of Rain Types and Their Z - R Relationships at Different Locations in the High Andes of Southern Ecuador. *Journal of Applied Meteorology and Climatology*, 56:3065–3080.

Orellana-Alvear, J., Célleri, R., Rollenbeck, R., and Bendix, J. (2019). Optimization of X-Band Radar Rainfall Retrieval in the Southern Andes of Ecuador Using a Random Forest Model. *Remote Sensing*, 11(1632):1–20.

Ovando, A., Tomasella, J., Rodriguez, D. A., Martinez, J. M., Siqueira-Junior, J. L., Pinto, G. L., Passy, P., Vauchel, P., Noriega, L., and von Randow, C. (2016). Extreme flood events in the Bolivian Amazon wetlands. *Journal of Hydrology: Regional Studies*, 5:293–308.

Paniconi, C. and Putti, M. (2015). Physically based modeling in catchment hydrology at 50: Survey and outlook. *Water Resources Research*, 51:2498–2514.

Paz, I., Tchiguirinskaia, I., and Schertzer, D. (2020). Rain gauge networks' limitations and the implications to hydrological modelling highlighted with a X-band radar. *Journal of Hydrology*, 583(January):124615.

Ragettli, S., Zhou, J., Wang, H., Liu, C., and Guo, L. (2017). Modeling flash floods in ungauged mountain catchments of China : A decision tree learning approach for parameter regionalization. *Journal of Hydrology*, 555:330–346.

Rollenbeck, R. and Bendix, J. (2011). Rainfall distribution in the Andes of southern Ecuador derived from blending weather radar data and meteorological field observations. *Atmospheric Research*, 99:277–289.

Sucozhañay, A. and Célleri, R. (2018). Impact of Rain Gauges distribution on the runoff simulation of a small mountain catchment in Southern Ecuador. *Water (Switzerland)*, 10(9).

Sudheer, K. P., Gosain, A. K., and Ramasastri, K. S. (2002). A data-driven algorithm for constructing artificial neural network rainfall-runoff models. *Hydrological Processes*, 16(6):1325–1330.

Tyralis, H. and Papacharalampous, G. (2017). Variable selection in time series forecasting using random forests. *Algorithms*, 10(4).

Valizadeh, N., Mirzaei, M., Allawi, M. F., Afan, H. A., Mohd, N. S., Hussain, A., and El-shafie, A. (2017). Artificial

intelligence and geo-statistical models for stream-flow forecasting in ungauged stations : state of the art. *Natural Hazards*.

van de Beek, C. Z., Leijnse, H., Hazenberg, P., and Uijlenhoet, R. (2016). Close-range radar rainfall estimation and error analysis. *Atmospheric Measurement Techniques*, 9:3837–3850.

Yaseen, Z. M., El-shafie, A., Jaafar, O., Afan, H. A., and Sayl, K. N. (2015). Artificial intelligence based models for stream-flow forecasting: 2000-2015. *Journal of Hydrology*, 530:829–844.

Yoon, S.-S. (2019). Adaptive Blending Method of Radar-Based and Numerical Weather Prediction QPFs for Urban Flood Forecasting. *Remote Sensing*, 11(6):642.

6 Conclusions and Outlook

6.1 Conclusions

This research constitutes a pioneer study that exploited SP X-band radar data at the high Andes of southern Ecuador and its application on discharge forecasting towards a machine learning (ML) approach. The investigation mainly focused in the leverage of weather radar data in mountain regions where ancillary rainfall monitoring at ground is scarce and unevenly distributed.

This project responded to the fact that tropical mountain rainfall in the Andes has been under-explored despite its importance for understanding climatic phenomena such as El-Niño Southern Oscillation (ENSO) that constantly affects precipitation in the region. The huge knowledge deficit is due to a very scarce operational network of rain gauge stations which do not allow to analyze spatio-temporal pattern of rainfall in the complex Andean orography. This imposes many restrictions to decision-makers regarding in the risk management of natural hazard such as floods and landslides.

Therefore, the deployment of a radar network over the Southern Andes of Ecuador (A1) was pursued in order to provide a suitable area-wide monitoring source that accounts for the high spatial variability of rainfall data in the mountain region. Thus, the first component of WP1 (Chapter 2) documented the implementation of two cost-effective X-band radars as well as complementary instrumentation that significantly extends the rainfall monitoring initially established earlier with the former LAWR at Loja. The result is the first weather radar network in tropical high mountains (RadarNet-Sur) worldwide. A framework for data transmission was set up and application examples that highlighted the benefits of the fine spatial resolution of radar imagery were provided.

The highest radar of the network (CAXX, 4450 m a.s.l.), arguable the highest worldwide at that time, was used for this study and yield to the main objective of this thesis which is the exploitation of X-band radar imagery in the complex orography of the Andes for hydrological applications. As a first step by pursuing a physically-based radar QPE, derivation of Z-R relations was needed. Thus, the second component of WP1 addressed this process by testing the first hypothesis (H1).

H1 wanted to test if different Z-R relations within the radar coverage needed to be used for radar

QPE. As a result of the analyses (Chapter 3), an identification of rainfall events and its characteristics at different locations was accomplished by using high resolution rainfall data (disdrometers). This allowed for the first time to document the rainfall structure of tropical high Andes in terms of DSD. Classification of rainfall events was based on the Dm (mean volume diameter) as a compact representation of rainfall microphysics. The results showed that light rain category [0.1< D_m(mm)≤0.5] exhibits a higher droplet concentration with smaller Dm in comparison with heavier rainfall classes. These changes on DSD across the rain types were also reflected in their derived rain-type Z-R relationships where the parameters A and b of the Z-R relations increased according to the Dm categorization However, both parameters had different ranges at all stations. In addition, a single Z-R relation was derived for each location and its adjustment was evaluated in comparison with the use of rain-type Z-R relationships. It was found that rain-type Z-R relations provided a better adjustment for converting Z to R. Finally, a comparison performed by using theoretical distribution models showed that the gamma distribution better fitted the data at lower stations in the study site whereas the exponential Marshall-Palmer model was more suitable for the highest station. Altogether, this evidenced the high variability in the rainfall structure occurring at different locations and altitudes within the radar coverage. Consequently, rain-type Z-R relationships greatly differed among geographical locations and, therefore H1 was confirmed.

The second aim (A2) seek the application of the step-wise approach for radar QPE and its comparison with a RF model from the ML-based family that served the same purpose. As part of the step-wise approach, the locally Z-R relations previously derived were used. This aim was achieved through the WP2 (Chapter 4) in which the second hypothesis (H2) was tested.

H2 wanted to test if a ML-based model could optimize the physically-based radar QPE. The major challenge in this particular mountain region came from the scarce and uneven distribution of rain gauges among the study site. This restricted its effective use for radar rainfall bias adjustment which is commonly performed after the step-wise approach is applied. Results showed that the use of the locally derived Z-R relationship Z= $103R^{2.06}$ had a slightly better adjustment than the Marshall-Palmer equation Z=$200b^{1.2}$. However, these models that follow the step-wise approach (i.e., several steps for reflectivity corrections and upscaling before applying a Z-R relation), without the ability of a proper bias correction had an insufficient performance for further applications. In addition to the lack of bias adjustment, this was partially an effect of the attenuation of the radar signal with an increasing distance from the instrument. This results in lower reflectivity values and thus the derived radar-rainfall estimates were highly underestimated. In contrast, the RF model (ML-based) achieved a far better performance for different locations, distances and rainfall events

within the radar range. Here, a decision-tree based RF model was applied using features derived from a SP X-band radar data by means of single PPI scans for the first time. Thus, by considering unavailability of rain gauges in real-time in Andean mountain regions, the RF model provided an outstanding alternative for obtaining radar rainfall estimates. Hypothesis H2 could therefore be confirmed as well.

The third aim (A3) follows from the need for exploiting X-band radar data in the high Andes for decision-making support as well as the seek for alternatives to avoid the requirement for high resolution rain gauge data. Thus, the goal was to compare the discharge forecasting performance of two RF models that used either native radar data (i.e., reflectivity) or derived radar rainfall estimates (i.e., rainfall rates). WP3 (Chapter 5) allowed carrying out A3 and testing the third hypothesis (H3).

H3 wanted to test if radar reflectivity could be exploited as input feature for a ML-based discharge forecasting model. The catchment of interest was divided in five rainfall regions under the assumption of different rainfall regimes that were previously described along the altitudinal gradient. Rainfall and reflectivity time series were derived from the radar for each of the five regions. As an additional input, the use of 3-day precipitation accumulation as a soil moisture proxy was tested in the models. Several configurations regarding the data-type input (Z or R) and inclusion or exclusion of the soil moisture proxy were evaluated. The results revealed that a similar goodness of fit was achieved for models that used (i) native radar data as well as (ii) derived radar rainfall estimates. Hence, the RF model was able to properly convert radar reflectivity directly into rainfall rates without intermediate steps by taking advantage of the principle of data representation in ML-based models. This in turn allowed the use of native radar data as input for an RF discharge forecasting model that accomplished as good results as those obtained when commonly using radar rainfall estimates. This highlighted the advantages of a reduced pre-processing for generating input features derived directly from reflectivity values in contrast to an exhaustive process for performing radar QPE as a preliminary step for discharge forecasting. Consequently, H3 can be confirmed.

Finally, the relevant contributions and novelties of this work are:

- A framework of and operational X-band weather radar network was deployed in the southern Andes of Ecuador providing a key source of high spatio-temporal resolution data of tropical mountain rainfall.
- Tropical rainfall structure in terms of drop size distribution (i.e., microphysical characteristics) and related Z-R relationships at high altitudes over complex terrain have been documented

for the very first time.

- Radar reflectivity data was not used in its original form but used to construct features that allow a radar QPE model to learn about the rain occurrence along the radar beam, partially overcoming the common attenuation problem. This in turn allowed to expand the useful range of rainfall estimates (up to $\sim$50 km). This approach was unprecedented for single-polarization radar applications and highlighted promising results.

- The usefulness of native radar data (i.e., reflectivity variable) for discharge forecasting was validated in this study. This came as a novel alternative to overcome the limitations of sparse rain gauge networks in mountain regions needed for exhaustive radar rainfall process observation and monitoring.

6.2 Outlook

The main conclusions of this investigation have been summarized in the previous section. In the following, planned future work and an outlook onto further research is provided.

After obtaining the first insights about rainfall types and its structure in tropical mountain regions, additional analyses related to the study of extreme rainfall events is of utmost importance for a proper identification of their dynamics, propagation and microphysical evolution. Such studies will increase the understanding of the complex rainfall formation process in tropical high mountains, and its connection with other hydrometeorological variables such as soil moisture and air temperature.

The proposed RF model for radar QPE will serve as a starting point for its application in two twin X-band radars of the updated radar network (RadarNetPlus) located in Celica-Ecuador and Piura-Peru which have the same technology and data format. Moreover, additional inputs derived from satellite imagery that are related to rainfall process formation (e.g., cloud top, temperature, etc) can be evaluated for improving the performance of the radar QPE model. This will allow the generation of more accurate rainfall maps to the local governments and provide them with effective tools for decision-making related to risk management and agricultural planning, just to name a few. These rainfall maps would open an extraordinary data source for hydrological and climatological applications that can benefit from the high spatio-temporal resolution of tropical mountain rainfall. Additional work on the evaluation on the rainfall accumulation at different temporal scales and its uncertainty is suggested.

Since the usefulness of native radar data has been validated for discharge forecasting, further exploration of the regionalization and/or simplification of the radar data should be pursued in order to better represent the spatial rainfall pattern and capture the main features of the radar imagery. This can be tackle by using deep learning as a feature extraction method that allows obtaining the most relevant image characteristics. Moreover, further investigation needs to be addressed regarding the soil moisture proxy derivation. It should be mostly focus on finding an optimal time-window for the soil moisture proxy derivation and looking for alternative ancillary data that may be able to provide a representation of the catchment state.

Erklärung

Ich erkläre an Eides statt, dass ich meine Dissertation mit dem Titel

"Exploitation of X-band weather radar data in the Andes high mountains and its application in hydrology: a machine learning approach"

selbstständig ohne unerlaubte Hilfe angefertigt und mich dabei keinerlei anderen als der von mir ausdrücklich bezeichneten Quellen und Hilfen bedient habe.
Die Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Marburg a. d. Lahn, den 03.11.2020

_____

Johanna Marlene Orellana Alvear


THIES
CLIMA