UCUENCA

Universidad de Cuenca

Facultad de Ingeniería

Doctorado en Recursos Hídricos

Investigating the relationships between precipitable water vapor estimations and heavy rainfall over the Eastern Pacific Ocean and Ecuadorian regions

Trabajo de titulación previo a la obtención del título de Doctora (PhD) en Recursos Hídricos

Autora:

María Sheila Fabiola Serrano Vincenti

Director:

Marcos Joshua Villacís Erazo

ORCID: 0000-0002-4496-7323

Codirector 1

Thomas Condom

ORCID: 0000-0002-4408-8580

Codirector 2

Lenin Vladimir Campozano Parra

ORCID: 0000-0002-7205-5786

Cuenca, Ecuador

2023-03-20

Investigating the relationships between precipitable water vapor estimations and heavy rainfall over the Eastern Pacific Ocean and Ecuadorian regions

María Sheila Fabiola Serrano Vincenti

ESCUELA
POLITÉCNICA
NACIONAL

UNIVERSIDAD DE CUENCA

UTPL
UNIVERSIDAD TÉCNICA PARTICULAR DE LOJA

Resumen

La lluvia es un fenómeno atmosférico difícil de predecir. Más allá de su importancia en las actividades humanas; existen dificultades teóricas y técnicas que justifican el estudio de la lluvia y la lluvia intensa. Los Modelos Numéricos atmosféricos de Predicción como el Weather Research & Forecasting Model (WRF), son las herramientas que se utilizan para predecir y estudiar su comportamiento, aunque presentan limitaciones al trabajar con lluvias intensas y topografías complejas y empinadas. Recientes investigaciones proponen a la estimación vapor de agua troposférico (Precipitable Water Vapor PWV), como una herramienta que puede ayudar a la predicción y entendimiento de los mecanismos que desencadenan lluvia intensa. Productos satelitales y su derivación indirecta a través del retraso de señales de Sistemas de Posicionamiento Global GNSS, son las principales fuentes actuales de PWV. El presente trabajo estudia la relación entre la lluvia intensa y el PWV satelital sobre el océano, la relación de PWV-GNSS sobre la Costa, Sierra y Oriente del Ecuador; así como con los datos modelados en WRF sobre zonas andinas ecuatoriales. Como principales resultados, se tiene un modelo empírico entre el PWV satelital y los valores máximos de lluvia sobre el océano; además, se identifican períodos de carga y descarga del PWV-GNSS relacionados con el ciclo diurno de la lluvia sobre tierra, y relaciones con los eventos intensos de lluvia; y por último, se encuentran las principales discrepancias entre los datos observados PWV-GNSS y lluvia con datos modelados de WRF sobre zonas de los Andes Ecuatoriales.

Palabras clave: vapor de agua precipitable, GNSS, lluvia intensa, WRF, análisis armónico

María Sheila Fabiola Serrano Vincenti

Abstract

Among the weather phenomena, rainfall is difficult to forecast, despite the theoretical and technical challenges inherently related to its prediction, its impact in economic and everyday activities, clearly justify its study. Numerical Weather Prediction Models are widely used to predict rainfall, such as the Weather Research & Forecasting Model (WRF), However, they underperform when is set to predict intense events and when working with complex and steep topographies. Recent studies have proposed the estimation of Precipitable Water Vapor PWV, as a tool that can help predict and understand the mechanisms that trigger intense rainfall. PWV is mainly sourced from satellite products and from indirectly measurements which derive it through the delay of the Global Navigation Positioning System (GNSS) signals quite accurately. Thus, the present work studies the relationship between intense rain and satellite sourced PWV over the ocean, the relationship of PWV-GNSS over the Coast, Sierra and Amazon of Ecuador, and the comparison of the PWV-GNSS with the data modeled in WRF. As main results, we point an empirical model between the satellite PWV and the maximum values of rainfall over the ocean. In addition, PWV-GNSS loading and unloading periods related to the diurnal cycle of rainfall over the land, and relationships with intense rain events were identified; and finally, the main discrepancies between the observed PWV-GNSS data and rainfall with WRF modeled data over areas of the Equatorial Andes.

Keywords: precipitable water vapor, GNSS, intense rain, WRF, harmonic analysis

María Sheila Fabiola Serrano Vincenti

Index of contents

Index of figures

María Sheila Fabiola Serrano Vincenti

Index of tables

"A crown is nothing more than a hat that lets the rain in".

Frederick the Great

Acknowledgments

María Sheila Fabiola Serrano Vincenti

To my beloved Amelie and Nicole

Chapter 1

1 Introduction

1.1 Importance of the study

Rain is a key ingredient for life and a few economical and human activities are not intimately linked with its presence -or absence- (Lenderink & Fowler, 2017). Given its importance, numeric weather prediction models NWP were created at the end of the past century, in order to be prepared for drought or abundant rain seasons(Davis et al., 2003). These models are a group of fluid dynamics non-linear equations processed with high computational power in a system with constant data assimilation fed from ground stations and satellite data around the globe (Warner, 2010). Today, these models are the greatest joint attempt to understand the physics of the atmosphere and of weather and climate events on the planet.

However, there are events that are beyond the prediction of these models, and intense rain is one of them; Herman & Schumacher (2016) indicate that there are important limitations when studying intense rain, which occurs on small time scales (less than two hours) and in short portions of space (less than 10 km), and that they also become especially inaccurate when working on steep and complex orography. Added to this fact are the additional difficulties faced by developing countries, where the in-situ meteorological data network is insufficient to validate -and improve- the NWP models, and where support for base research is limited.

It can be mentioned that at the beginning of this century, research has focused on the prospective use of Precipitable Water Vapor PWV and its relationship with rain, as one of the best strategies to improve rain prediction. This magnitude corresponds to the amount of water vapor that exists in a column from the surface to the top of the troposphere and is also known as Tropospheric Water Vapor TPW or Integrated Water Vapor IWV (Bonafoni et al., 2019). And although the importance of its relationship with rain is evident, it was a difficult magnitude to obtain outside of radiosondes. It was not until 1997 that the Tropical Rainfall Measuring Mission TRMM provided this variable for the tropics (-30°, 30°) through satellite means, although was only for regions over the ocean, since the terrestrial albedo interfered with the PWV spectrum that the satellite assimilated (Grecu & Anagnostou, 2001).

Since then, and with new satellite products, research propose physical relationships between PWV and rainfall (Neelin et al., 2009; Peters & Christensen, 2002; Peters & Neelin, 2006),

María Sheila Fabiola Serrano Vincenti

which show the prospective benefit of PWV for rainfall understanding, especially for large intensities, and whose potential has not yet been fully exploited. Concurrently, from the geosciences, it was found that the signal delays of the GNSS systems produced by increase in the humidity of the PWV atmosphere, can be quantified and therefore be able to measure it quite accurately over the land (Bevis et al., 1992; Businger et al., 1996).

These novel techniques, which came to be called GNSS Meteorology, require a sufficiently dense network of high-quality or geodetic sensors over the study area, whose number depends on the complexity of the orography, in addition to considerable High Performance Computing (HPC) (Bonafoni et al., 2019). The meteorological products derived from GNSS systems were used in two ways: the first was the assimilation of this data directly in the NWP models, a fact that considerably increased their prediction ability (Risanto et al., 2021; Shoji et al., 2009); and through specific case studies, where possible physical relationships were found, such as the presence of PWV peaks or thresholds that preceded intense rain events, especially in low altitude zones as Portugal (Benevides et al., 2015, 2019), China (H. Li et al., 2020; L. Li et al., 2022) and the Amazon (Adams et al., 2013a; Sapucci et al., 2019).

In Ecuador, the study of the GNSS-derived PWV has been previously carried out on a monthly and daily scale (Ayala et al., 2018; Mackern, M. V., Mateo, M. L., Camisay, M. F., & Morichetti, 2020; Porras et al., 2016; Romero et al., 2019), and even to hourly resolutions, in several days, but in a specific location (Albán, 2019). On the other hand, various studies have been conducted on the implementation and validation of the NWP models in their estimation of rainfall in the steep orographies of the Equatorial Tropical Andes (Erazo et al., 2018; Heredia et al., 2019; Junquas et al., 2022), though their connection with the PWV has not yet been integrated, especially in the case of intense rain.

1.2 Objectives

Under this scenario, the following objectives are proposed based on the available data, with the best resolution and in the most representative way on the regions of Ecuador to find relationships between PWV and intense rain.

i. Develop a high temporal and spatial resolution model for maximum rainfall rates related to PWV over the ocean using satellite data.

ii. Looking for the relationship between intense rainfall and PWV derived from the GNSS delays over different zones over Ecuador´s mainland.

iii. Find through the comparison of the series between the WRF model and in-situ observations of PWV-GNSS, possible errors in the PWV modeling, which could be addressed in future work to improve the estimation of rain and intense rain simulated by the model.

1.3 Outline of the doctoral thesis and study area

This dissertation consists of five chapters that address PWV relations with rain and intense rain. Three manuscripts are the core of this thesis; two of them are already published in scientific journals (Chapter 2 and Chapter 3) and the third is currently submitted to peer-review (Chapter 4). Chapter 5 is a summary of them, including recommendations and future work.

Chapter 2 makes a first approach between PWV and rain with the satellite data provided by the Tropical Rainfall Measuring Mission TRMM, which only had over-the-ocean data. For the best possible resolution and fidelity of the data, products of second generation were employed as the 2A12 product with a resolution of 5.1 × 5.1 km every 1.67 sec, with four years of data, from January 2009 to December 2012. The selected area of study was Niño 2 (0 –5°S; 90–81°W) in front of Ecuador Coast, over the Pacific Ocean. Such resolution referred only to satellite-observed data which was not filled by models, such as third-generation products like the 3B43; this resulted in satellite swept regions that traversed the study area twice a day, which resulted in discontinuous time series despite the fact that PWV and rain data were available for each studied pixel. Thus, it was possible to propose an exponential empirical relationship between the maximum rainfall values associated with a given amount of PWV using model selection tools. Likewise, as a result, the higher potential in locating rain events from PWV with respect to the Sea Surface Temperature SST was obtained.

Chapter 3, seeks to explore the relationships between PWV and rain over the land of equatorial region. For this purpose the PWV is derived from the GNSS delays produced by the refraction of the GPS signal from the atmosphere, data was provided by the Geophysical Institute of the National Polytechnic School IGE-EPN , together with the assistance and computational resources of the l'Institut des Sciences de la Terre in Grenoble (ISTerre). In this analysis was possible to obtain a continuous time series of PWV on land for the year 2014 with an hourly resolution. The limitation of this study consisted in finding meteorological stations with data logged at hourly resolutions close enough to the GNSS stations. The outcome was to have five regions for analysis: on the coast (Esmeraldas), two in the Andean valleys (Ibarra and Quito), in the high mountains (Antisana) and in the Amazon (Tena). The

study found, through harmonic analysis tools, periods of charge and discharge of the PWV in the atmosphere that were not in phase with the periods corresponding to the diurnal cycle of rain. Likewise, when analyzing specific intense rain events, it was found the existence of PWV peaks that preceded intense rain events by 11 hours, except in the Amazon station, where the recorded peak was lower. However, these PWV peaks did not always occur before an intense rain event; or if they did occur, they did not always herald a rain event.

In Chapter 4, WRF model data previously validated for the Andes high mountain and Andes valleys regions (Heredia et al., 2019) with hourly and spatial resolution of 1km x 1km was analyzed, reason why Quito and Antisana stations were studied in this case. When comparing the observed and modeled data of PWV and rain using harmonic analysis, it was found that the model did not see the relationships of loading and discharge of PWV in antiphase to the diurnal cycle of rain observed in the previous Chapter; the model determined that these cycles were in phase, that is, that the PWV increased together with the rain. When analyzing specific intense rain events, the model showed a similar behavior, with lags of less than 3 hours. Likewise, it was concluded that the better the estimate of PWV, the better the estimate of rainfall.

Finally in Chapter 5, the advancements and limitations of each result are contrasted, pointing to the research that can still be carried out as plausible future work.

Chapter 2

An Empirical Model for Rainfall Maximums Conditioned to Tropospheric Water Vapor Over the Eastern Pacific Ocean

One of the most difficult weather variables to predict is rain, particularly intense rain. The main limitation is the complexity of the fluid dynamic equations used by predictive models with increasing uncertainties over time, especially in the description of brief, local, and high intensity precipitation events. Although computational, instrumental and theoretical improvements have been developed for models, it is still a challenge to estimate high intensity rainfall events, especially in terms of determining the maximum rainfall rates and the location of the event. Within this context, this research presents a statistical and relationship analysis of rainfall intensity rates, Precipitable Water Vapor (PWV), and sea surface temperature (SST) over the ocean. An empirical model to estimate the maximum rainfall rates conditioned to PWV values is developed. The performance of the maximum rainfall rate model is spatially evaluated for a case study. High-resolution TRMM 2A12 satellite data with a resolution of 5.1 × 5.1 km and 1.67 s was used from January 2009 to December 2012, over the Eastern Pacific Niño area in the tropical Pacific Ocean (0–5°S; 90–81°W), comprising 326,092 rain pixels. After applying the model selection methodology, i.e., the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC), an empirical exponential model between the maximum possible rain rates conditioned to PWV was found with $R^2 = 0.96$, indicating that the amount of PWV determines the maximum amount of rain that the atmosphere can precipitate exponentially. Spatially, this model unequivocally locates the rain event; however, the rainfall intensity is underestimated in the convective nucleus of the cloud. Thus, these results provide an additional constraint for maximum rain intensity values that should be adopted in dynamic models, improving the quantification of heavy rainfall event intensities and the correct location of these events.

Related publication:

María Sheila Fabiola Serrano Vincenti

Serrano-Vincenti, S., Condom, T., Campozano, L., Guamán, J., & Villacís, M. (2020). An Empirical Model for Rainfall Maximums Conditioned to Tropospheric Water Vapor Over the Eastern Pacific Ocean. Frontiers in Earth Science, 8(July), 1–12. https://doi.org/10.3389/feart.2020.00198

2.1 Introduction

Disasters produced by extreme, episodic and abrupt rain events are still the deadliest natural hazard globally, with the most destructive and long-term effects on the economy, infrastructure, ecosystems, food security and people (Jonkman, 2005; Pielke et al., 2013; Ilbay-yupa & Barragán, Ricardo Zubieta, Lavado-Casimiro, 2019; Y. Li et al., 2019;). This vulnerability not only reflects the importance of adequate population spatial planning but the low accuracy of Numerical Weather Prediction (NWP) models in terms of forecasting high rainfall intensity rates for global events, especially in the most vulnerable zones (Schumacher, 2016).

Empirical modeling was proposed as a complementary strategy to improve rainfall modeling and prediction capabilities in NWP models (Jakob Themeßl et al., 2011). It relates to observed meteorological variables in new relationships, improving rain estimations or even providing indications of new physical implications of rainfall. Empirical relationships could be used to estimate rainfall over orographically stepped zones in order to address shortcomings with NWP models (Haiden & Kahlig, 1994; Mathias Vuille, Raymond S. Bradley, 2000a; Buytaert et al., 2006; Villacís et al., 2008; Heredia et al., 2018) by improving the location of rain events (Herman & Schumacher, 2016) and improving the quantification of high-intensity rain events.

Over the ocean, temperature and atmospheric humidity have empirical relationships with rainfall Haga clic o pulse aquí para escribir texto.(Bretherton et al., 2004; Holloway & Neelin, 2009; Nathan et al., 2016; Ken Takahashi & Dewitte, 2016; Ahmed & Neelin, 2018). The positive relationship between sea surface temperature (SST) and rain intensities has been widely reported (Jauregui & Takahashi, 2018a; Johnson & Xie, 2010; Manabe, S., Han, D., Holloway, 1974) mainly due to the increase in parcel instability and the convective available potential energy (CAPE), which makes it possible to overcome the convective inhibition (Betts & Ridgway, 1989). Additionally, using monthly data, the observations show a rainfall peak between SST values of 26°C and 28°C, this temperature range was identified as a convection trigger by several authors, and after this SST maximum, the rain intensities decrease (Gray, 1998; Jauregui & Takahashi, 2018a; Johnson & Xie, 2010; Vincent et al., 2012).

With regards to the relationships with humidity, and using a different time resolution, Bretherton et al. (Bretherton et al., 2004) present an exponential dependence between the daily mean precipitation and the column-relative humidity (obtained by dividing the total PWV by its corresponding saturation value). PWV represents the depth of water in mm in a column of the atmosphere in the case that the water in that column was precipitated as rain. This study was conducted over four oceans and four years, and using satellite data with a resolution of 2.5° x 2.5° (279.3 km x 279.3 km).

In a similar study carried out with a temporal resolution of 3 hours and a spatial resolution of 0.25° x 0.25° (27.9 km x 27.9 km), Peters and Neelin (Peters & Neelin, 2006) reported a power law relationship between the mean rainfall and PWV, exceeding the critical value of PWV. In other research with a similar resolution, rainfall and PWV values were reported (Kuo et al., 2017; Schroeder et al., 2016) when studying tropical convection over the ocean at different pressure levels (Neelin et al., 2009a; Martinez-Villalobos & Neelin, 2018), and even over land (Bernstein & Neelin, 2016; Leon et al., 2016; Sapucci et al., 2019).

A high temporal and spatial resolution analysis of rainfall could provide more realistic results. If a coarse temporal and spatial resolution is used instead, this could lead to the pooling of small events and the underestimation of large events by averaging them with shorter or null events (Lovejoy & Mandelbrot, 1985; Newman, 2005; Peters et al., 2002; Dickman, 2003). Additionally, an adequate spatial resolution may give a more exact and precise location of the rainy event and its predictors.

One particular difficulty for estimating rainfall rates lies in the statistical features related to maximum rainfall events (Arakawa, 2006). Rainfall occurrence is characterized by numerous small intensity events and only a few heavy intensity events; however, these large events are strongly representative because, depending on the study area, they can discharge 70-90% of the total amount of rain registered in one year (Pendergrass, 2018). These large events are not well reproduced by the NWP model and are one of the causes of high vulnerability to extreme rainfall events. Then, the maximum rainfall is of special interest since it indicates the maximum amount that the atmospheric system can precipitate, producing the largest and most destructive rainy events.

Therefore, the objective of this study is to develop a high temporal and spatial resolution analysis of maximum rainfall rates related to PWV and SST. Thus, microwave satellite data from TRMM (Tropical Rainfall Measure Mission) Microwave Imager (TMI) Level 2 Hydrometeor Profile Product (TRMM Product 2A12), with a resolution of 5.1 x 5.1 km and 1.67 seconds, is used in this research. This data proves to be very useful and provides reliable data

over the ocean even during storms and hurricanes. Moreover, its advantages include reliability, broad area coverage and accessibility where *in situ* data are scarce (Khairoutdinov & Randall, 2006; Kummerow et al., 1996; Wang et al., 2009; Wilcox & Donner, 2007). The area chosen in our study is in the eastern part of the tropical Pacific Ocean (i.e. the Eastern Pacific Niño, or Niño E region), reported as a heavy rainfall area in which events could population (Ken Takahashi & Dewitte, 2016).

The aims of this paper are threefold: (i) to perform a statistical evaluation of the rainfall, PWV and SST, (ii) to determine the point estimate relationships between intense rainfall and its predictors by choosing the most suitable mathematical model able to estimate maximum rainfall by season, and (iii) to spatially analyze the performance of the empirical model with respect to the spatial location of the rainfall event as a study case.

2.2 Study area and data

Figure 1 presents the Niño E area, an area spanning 560,242 km^2 (0-5°S and 90-81°W). This area is important due to the presence of heavy rains and the influence of the ENSO (El Niño South Oscillation) phenomenon. It is in the vicinity of the northern South American coast, affecting Ecuadorian and Peruvian territories in particular (K. Takahashi et al., 2011; Ken Takahashi & Dewitte, 2016).
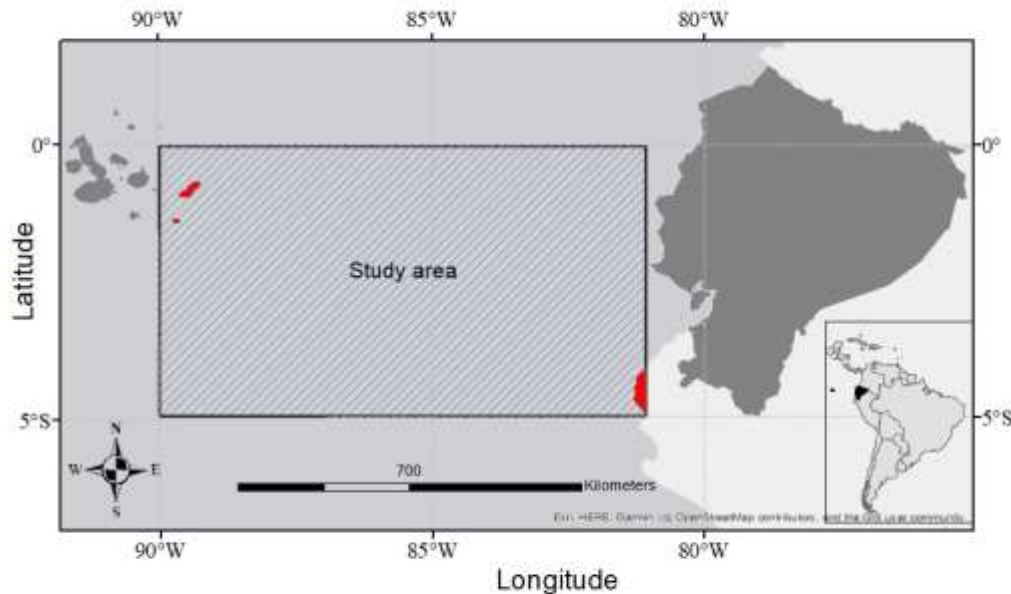
Figure 1 Study area over the eastern equatorial Pacific, called Niño E (Takahashi and Dewitte, 2016). The zones over land (red) are excluded from the study area.

This data comes from a TRMM Microwave Imager (TMI) satellite, which is a passive microwave radiometer and a dual-polarized (vertical and horizontal) multichannel sensor with the following five frequencies: 10.65 GHz, 19.35 GHz, 37 GHz, and 85.5 GHz. The 21.3 GHz frequency has vertical polarization only. The spatial coverage ranges from 40°N to 40°S and the satellite was launched on November 27, 1997 at an altitude of 380 km and an inclination of 35° (NASDA, 2001). Specifically, the 2A12 level has 16 orbits per day with 2,991 scans, each with 280 high-resolution pixels, and a data offering of 5.1 x 5.1 km each and 1.9 seconds, which is the scan rotation period (Wentz et al., 2001). The satellite swath lasts around 3 min overflying the study area, twice a day, resulting in a non-continuous but randomized time series, hereafter referred to as the Rainfall Intensity Data (RID). This information is available in HDF4 and .nc format at:

http://disc.sci.gsfc.nasa.gov/precipitation/documentation/TRMM_README/TRMM_2A12_readme.shtm.

María Sheila Fabiola Serrano Vincenti

The available TMI-TRMM instantaneous RID rates are calculated using the Goddard profiling algorithm (GPROF) (Kummerow et al., 2001), where the response functions of different channels detect different depths within the rain column, defining different brightness temperature vectors **Tb,** each one related to a vertical distribution of the hydrometeors **R**. However, the desired variable is the contrary, i.e. the vertical distribution of the hydrometeors **R**, for a given **Tb.** Then, in terms of probability, it is estimated by Bayes Equation 1:

$$\Pr(\mathbf{R}|\boldsymbol{Tb}) = \Pr(\boldsymbol{R}) \Pr(\boldsymbol{Tb}|\boldsymbol{R}) \qquad (1)$$

where $\Pr(\boldsymbol{R})$ represents a rain probability observed profile that is determined by CRM (cloud-resolving models). Once the $\Pr(\boldsymbol{R})$ has been reached, the intensity of the rain is derived; this estimation is especially accurate over the ocean (Furuzawa and Nakamura 2005). In order to focus on the maximum rain rates in this research, the chosen values are higher than 0.1 mm/h (Table 1).

Over the oceans, the GPROF as well as TMI 2A12 product, compare well with atoll rain gauge data. The TMI 2A12 rain rate is biased negatively by 9%, and with a correlation of 0.86. The correlation increases to 0.91 and is positively biased by 6% if two or more atolls are used in the validation (Kummerow et al., 2001). The expected sampling errors for the rain estimation are calculated by algorithms and depend on the rain intensity, i.e. the more intense the rainfall the more accurate the TMI estimation, following Eq.2:

$$\frac{\sigma}{R}(\%) = 0.26R^{-0.27} \qquad (2)$$

where R, is rainfall intensity in mm.h^{-1} and σ is its standard deviation. Therefore, TMI 2A12 data are better for determining intense rainfall rates (Olson et al., 1999; Bell et al., 2001).

The amount of PWV detected by TMI that is related to the liquid equivalent of the total water vapor column in mm has its absorption line in the 21.3 GHz band, i.e. it is an accurate measurement due to this high signal-to-noise ratio (Furuzawa and Nakamura 2005); the root mean square error (RMSE) is 8.1 mm with a bias of -0.76 mm**,** after radiosonde validation (Sajith et al., 2007). However this data is only reliable over the ocean, since it overlaps with the terrestrial albedo over land. It is important to mention that CRM ancillary data report few PWV values higher than 70 mm, and this difference could lead to numerical instabilities over this value (Kummerow et al. 2001). Finally, the SST product is estimated directly by using lower bands such as 19.35 GHz and 10.65 GHz, and is given in K (NASDA, 2001); thus, the accuracy of the estimation is 0.95 K (Guan & Kawamura, 2004).

2.3 Materials and Methods

A statistical analysis of the studied variables (RID, PWV and SST) is presented in section 3.1 to find the model for the maximum rainfall values. In section 3.2, the relationships of the maximum rainfall values conditioned to PWV (Peters & Neelin, 2006; Neelin et al., 2009a;) and SST (Jauregui & Takahashi, 2018b) were analyzed. Then, the choice of the best model and the confidence intervals were found according to the model selection methodology. Finally, in section 3.3, the predictive performance and location of the model in a case study were analyzed.

2.3.1 P, PWV and SST statistical metrics

First, the data were downloaded and decrypted using Python as detailed in (Serrano S. 2016). Big data post-processing and data selection were performed in R using High-Performance Computing (HPC) MODEMAT at the Escuela Politécnica Nacional laboratory. The chosen data (rain > 0.1 mm/h) were 326,092 rain pixels observed over the study area.

Monthly boxplot analyses of the data were performed to determine the seasonality of RID, PWV, and SST. The characteristics of the rainy and dry seasons were identified (Figure 2). Then, a statistical evaluation of RID, PWV and SST was presented by using frequency histograms (Figure 3). Table 1 shows the main descriptive estimators to determine the normality of the data and its statistical implications.

2.3.2 Relationships between intense rainfall rates, PWV and SST

The data were divided into identical 200 bins to determine the dependence of the rainfall rates on PWV and SST. The data were first aggregated into 0.25 mm PWV bins, and then into 0.07°C SST bins. The selected number of bins does not affect the behavior of the studied meteorological variables. The maximum RID values related to each bin were collected and identified with their corresponding SST and PWV values. Based on the scatter plots of the data and the literature, different models were evaluated ranging from exponential (Bretherton, Peters, and Back 2004), power laws (Peters and Neelin 2006) and stretched exponential (Martinez-Villalobos and Neelin 2018), among others.

Distribution fitting was performed using mixed non-linear regression methods in the nls R package (Baty et al. 2015). These methods optimize a modification of the least-squares criterion, which can be applied to non-Gaussian variables. The parameters of the model are estimated iteratively by using starting values. The fit of the non-linear regression is evaluated using graphical tools and the standard confidence intervals are derived, assuming normality in the standard deviations, by obtaining the typical metrics of the residuals. This procedure was repeated for each model: i.e. exponential, power-law, stretched exponential and polynomial relationships.

The best model was chosen using the standard errors (calculated using the Hessian matrix estimate in the maximum likelihood estimation), as well as the AIC and BIC. The last two criteria estimate the information that a model loses when fitting the data, balancing the underfitting by goodness-of-fit and by using likelihood with the overfitting, due to the unnecessary addition of parameters. Therefore, the model that lost the least amount of information was selected as the best model (Clauset, Shalizi, and Newman 2009).

Finally, the chosen model was proven in the dry (ASO) and wet season (FMA) separately, repeating the procedure described above, in order to determine the influence of these seasons on the RID data.

2.3.3   Spatial analysis

A study case (April 5, 2012) is presented to determine the performance of the selected model with regards to rain intensity estimation and spatial location. This event corresponded to the largest rain event registered in the observed satellite swath (lasting 3 min, twice a day, over the study area from 2009 to 2012). A pixel-by-pixel comparison of the RID with PWV and SST was carried out using R raster libraries.

The maximum rainfall pixels were calculated using the model selected in Section 2.3, based on each PWV value, in order to obtain the maximum modeled rainfall. Finally, the differences between the modeled and observed RID values are used to evaluate the performance of the spatial model. This type of spatial analysis will allow to evaluate the performance of the model in the different regions of the cloud, as well as to verify the ability of PWV and SST to locate the maximum rainy events, with low-cost calculations.

2.4 Results and discussion

2.4.1 Monthly analysis

In the RID, PWV and SST boxplots presented by month (Figure 2), most of the events have small intensities in both the wet and dry season due to the high resolution of the data. Unimodal seasonality is evident with a marked rainy season from February to April, and a dry season from August to October. This behavior is well known (Manabe, S., Han, D., Holloway 1974), and is mainly due to the influence of the Intertropical Convergence Zone (ITCZ), which is at the southernmost location in March, resulting in low-pressure systems and a marked increase in SST (Jauregui and Takahashi 2018) favoring convective rainfall in the first part of the year. To the contrary, the ITCZ is at its northern maximum in July-August, decreasing low pressure and precipitation as well (Campozano et al., 2016, 2018). During the dry season, the events are not larger than 1 mm/h, while in the wet season they can reach magnitudes up to 40 times larger, presenting a large seasonal contrast. Similar behavior –but not as marked as that observed for rainfall– is reported for PWV and SST, evidencing their direct relationship with precipitation. More evaporation occurs with warmer SST values, and then more warm water vapor over the surface could ascend, increasing the PWV (Wu, Stevens, and Arakawa 2009). This response is related to the Clausius-Clapeyron relation: the warmer the air, the more moisture it can contain; and later it condensates as precipitation (Roderick, Wasko, and Sharma 2019).

Table 1 RID, PWV, and SST main statistical metrics. Note the marked non-gaussian behavior of RID distribution marked in bold.

	SKEWNESS	KURTOSIS	MIN	Q1	MEDIAN	Q3	MAX
RID [MM/H]	**4.95**	**41.48**	**0.1**	**0.16**	**0.3**	**0.68**	**39.32**
PWV [MM]	-0.99	1.43	19.22	56.48	61.32	65.20	75
SST [°C]	-1.45	3.09	15.8	26	27	27.7	30

2.4.2 Statistical evaluation of RID, PWV and SST

In Figure 3, the frequency histograms for PWV and SST present a Gaussian left-skewed behavior. However, the RID distribution presented on a semi-logarithmic scale shows right skewness and high kurtosis (See Table 1). It is evident that RID follows a non-Gaussian heavy-tailed distribution with Q1, Q2, and Q3 less than 0.68 mm/h, whereas the maximum value is 39.32 mm/h.

Due to its Gaussianity, the best parameters for describing the behavior of PWV and SST were the mean and variance. However, the strong heavy-tailed behavior of RID shows the importance of large values in these data. Maximum events shape rain distributions and need to be analyzed carefully (Clauset, Shalizi, and Newman 2009).

María Sheila Fabiola Serrano Vincenti
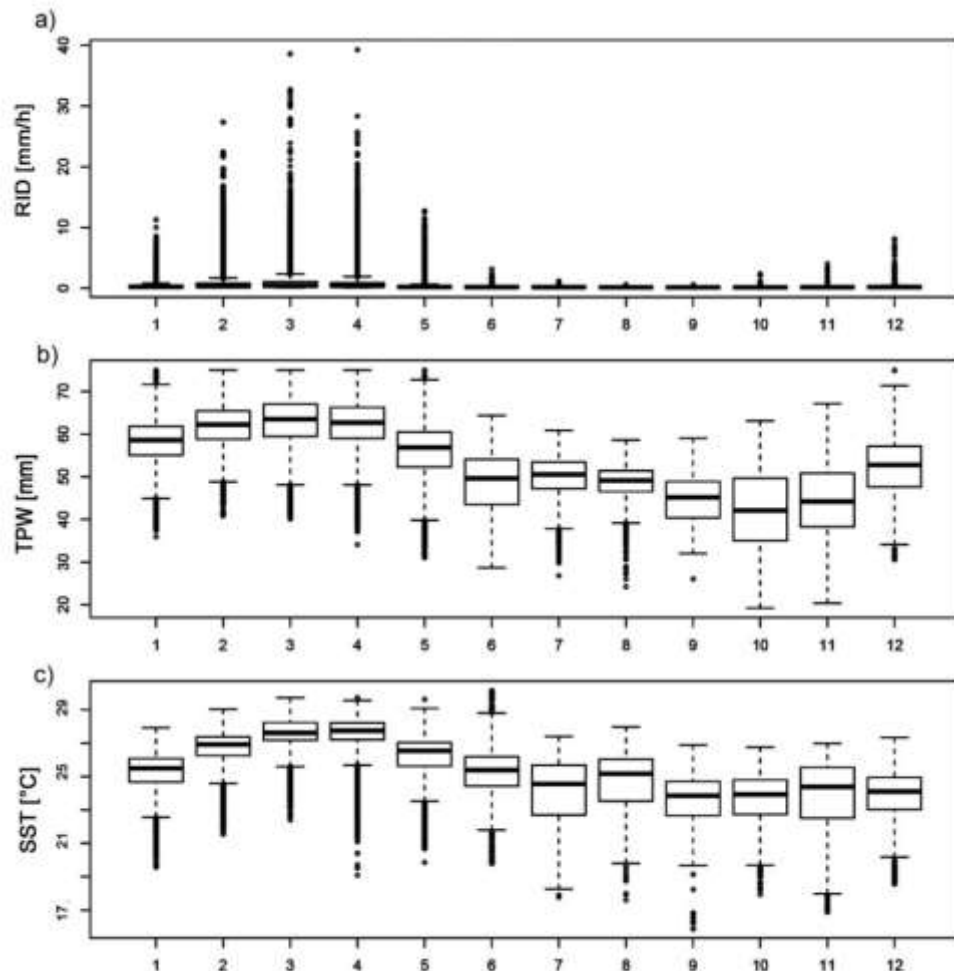
Figure 2 Monthly boxplot analysis for a) RID, b) PWV and c) SST. The rainy season extends from February to April and the dry season extends from August to October.
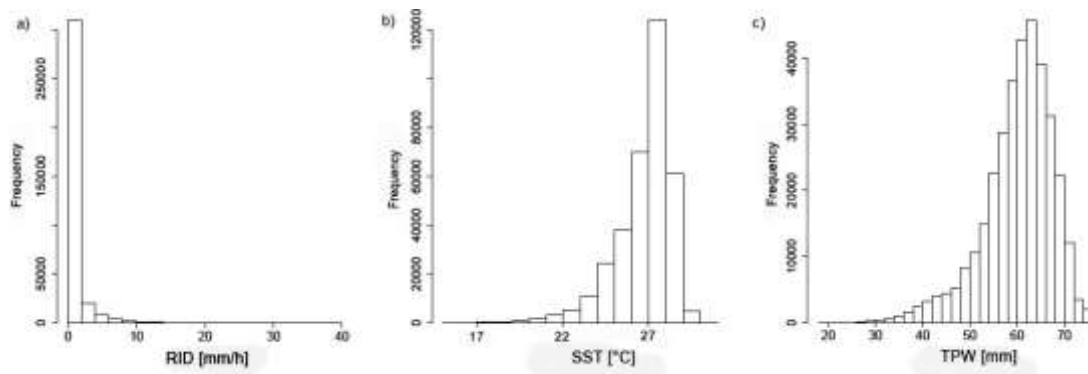
Figure 3 High-resolution data frequency histograms of studied variables: a) RID b) SST and c) PWV from 2009-2012.

2.4.3 Functional relationships between rainfall, PWV and SST.

The maximum RID values associated with each PWV and SST bin are presented in Figure 4. These 200 values are hereafter referred to as MRI_{PWV} for the Maximum Rainfall Intensity conditioned to PWV, and their corresponding PWV and SST values are denoted as PWV_{Max} and SST_{MaxPWV}. Similarly, MRI_{SST} is the Maximum Rainfall Intensity conditioned to SST, and the corresponding PWV and SST are PWV_{MaxSST} and SST_{Max}, respectively.

The associated PWV_{Max} vs. SST_{MaxPWV} values for the 200 MRI_{PWV} bins are presented in Figure 4a. It shows a positive but scattered dependence: as the SST increases, the PWV content increases as well (Gamache and Houze 1983). An increase in the slope is presented above 24°C for SST_{MaxPWV}. Similar behavior is shown in Figure 4b where the maximum rainfall MRI_{PWV} values and their corresponding SST_{MaxPWV} values are presented. The same threshold of 24°C marks a phase of higher and more dispersed rain intensity. Below this value, MRI_{PWV} does not exceed 0.5 mm/h.

The MRI_{PWV} bins and their corresponding PWV values are presented in Figure 4c. It shows that maximum rainfall increases rapidly with PWV. The rainfall increase related to PWV is smooth and continuous, which points to a rapid growth function. From the findings in the literature, similar studies with 3-hourly mean rainfall values indicate that these suitable models could be: exponential (Bretherton et al., 2004), power law (Peters and Neelin 2006; Neelin,

Peters, and Hales 2009) or stretched exponential models (Martinez-Villalobos and Neelin 2018), among others like potential functions.

The model selection is presented in Table 2. Based on the convergence of the parameters, fitting residuals, and the AIC and BIC information criteria, it is determined that the best model is the exponential model, followed by the power law model.

Table 2 Multiple R-squared error (R2), Residual standard error (RMSE), Akaike (AIC), and Bayesian (BIC) information criteria for model selection between MRIPWV and PWV, the best model is exponential (bolded).

Model	R^2	RMSE mm/h	AIC	BIC	Parameters	
$MRI_{TPW} = ae^{b*TPW}$ **Exponential**	**0.96**	**0.34**	**126.2**	**135.7**	$a = 0.01 \pm 0.0012$	$b = 0.109 \pm 0.0063$
$MRI_{TPW} = aTPW^b$ Power Law	0.89	0.52	290.3	299.9	$a = 0.01 \pm 0.0012$	$b = 4.3885 \pm 0.74$
$MIR_{TPW} = MIR_{TPW}^{-b} * e^a$ Streched Exp.	0.79	2.98	337.8	387.9	$a = 8.69 \pm 1.3$	$b = 4.3885 \pm 0.4$

Then, the rainfall maximum MRI conditioned to PWV follows an exponential model (Equation 3):

$$MRI_{TPW} = ae^{b*TPW} \qquad (3)$$

It shows that the model performs well for the maximum rainfall values, and only the highest values are underestimated. This behavior could be due to two reasons: first, the proximity to the 70 mm threshold, where the scarcity of the CRM data needed for the rain satellite calculation could result in artificial numerical instability effects as mentioned in Section 2.3 (Kummerow et al. 2001). The second reason could be due to the highest rain values, which could correspond to a reinforced convection typical of the center of the storm, as can be seen in Section 2.4.3. The predictability of convection events is limited due to turbulence complexity (Nielsen and Schumacher 2016). Therefore, it is important for future to determine if this value of 70 mm represents an artificial threshold or a physical one.

Scattered relationships are presented for the MRI data associated with the SST values (right side of Figure 4). First, in Figure 4a, the PWV vs. SST plot for the maximum values related to SST are presented. The data presented a mild threshold between 26°C and 28°C, similarly to

María Sheila Fabiola Serrano Vincenti

Figure 4b, where the reported threshold of 26-28°C for the rainfall peaks is presented (Jauregui and Takahashi 2018; Johnson and Xie 2010; Gray 1998). In this study the peak coincides although the time scale used in the literature is monthly.

Due to the great thermal inertia of the water, and consequently of SST, the increase over 26°C is due to the climatological presence of ITCZ, which results in general warming and, therefore, the arrival of the rainy season (Figure 2c). In other words, SST serves to characterize rain behavior on a climatic scale, rather than to characterize specific events of rainy precipitation. In Section 2.4.4, the low localization capacity of SST can be seen in a case study (Figure 6b). Therefore, SST is not a suitable variable in terms of proposing a functional model for rainfall. Finally, Figure 4f presents the rainfall $MRIss_T$ points and their corresponding PWV_{MaxPWV} values. Over 60 mm of PWV indicate a high rainfall intensity, but in a very scattered manner.

Figure 5 presents the above analysis for the wet and dry season separately. Including the PWV and SST values registered in each season. The events that appear most in the model correspond to opposite seasons, i.e. the driest (ASO) and wettest (FMA) seasons (Figure 5a). This result is remarkable because the exponential model represents the absolute opposite seasons instead of the intermediate –and most frequent– seasons. Most of the maximum rain rates associated with the PWV values belong to the wet season, as expected, however the smallest maxima have small PWV values (< 34 mm), which are only present in the dry season (Figure 5b). And its intensity does not exceed 1 mm/h over the entire range. This fact proves that the model can be applied over the entire range of the rainfall and PWV values.
Seasonality is also marked when analyzing SST (Figure 5c), it is seen that in the wet season, the majority of events are between 26 ° C and 28 ° C. As mentioned before, it is a region with high rainfall rates.
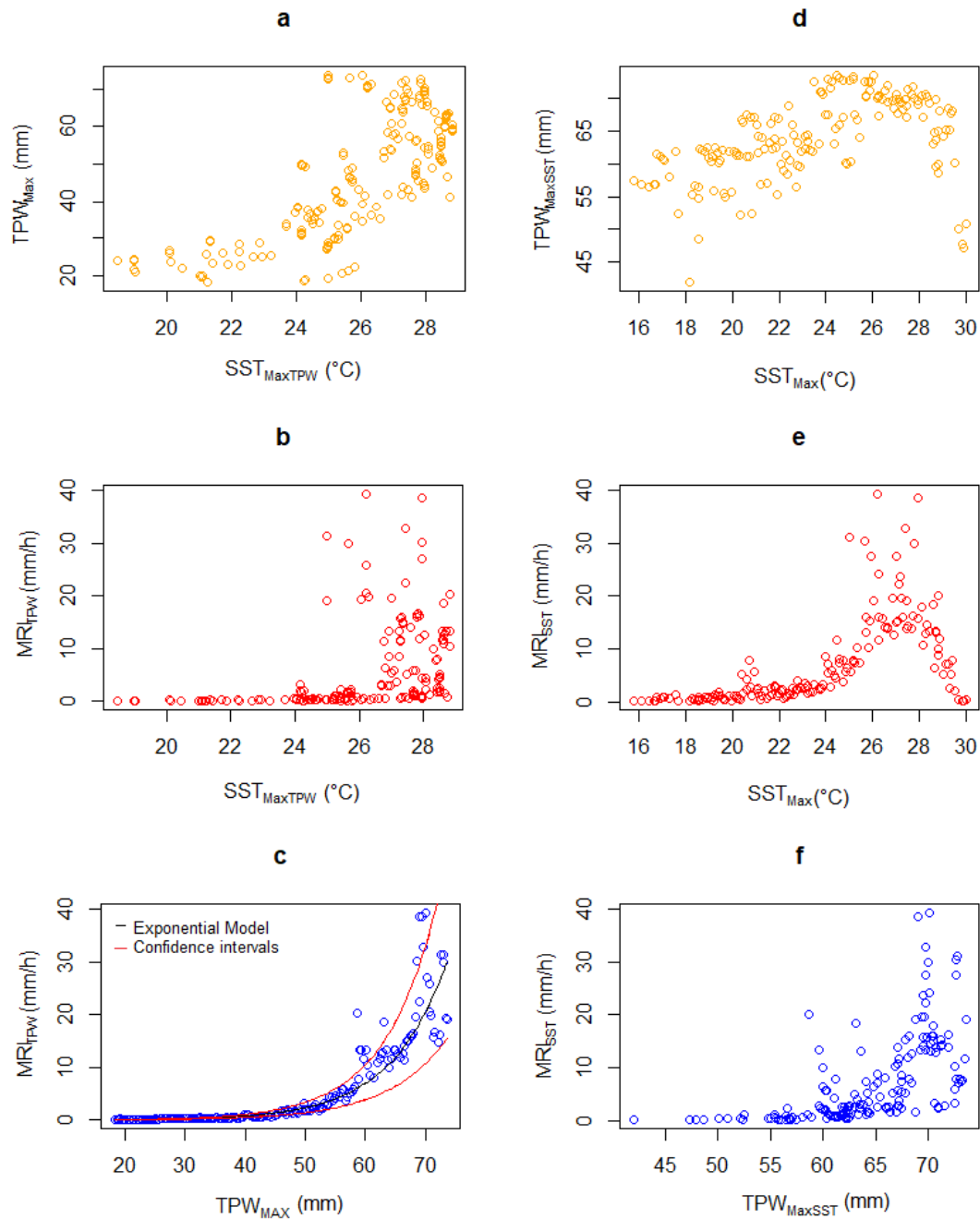
Figure 4 MRI points associated with PWV (left side): a) PWVMax vs. SSTMaxPWV; b) MRIPWV points and their corresponding SST values (SSTMaxPWV); c) MRIPWV associated with the PWVMAX bins, including the adjusted exponential model selected (continuous black line) and the 95 percent confidence

intervals (red lines). Right side: MRI points associated with SST: d) PWV_{MaxSST} vs. SST_{Max}; e) MRI_{SST} points and their corresponding SST values (SST_{Max}) ; f) MRI_{SST} associated with PWV_{MaxSST}.
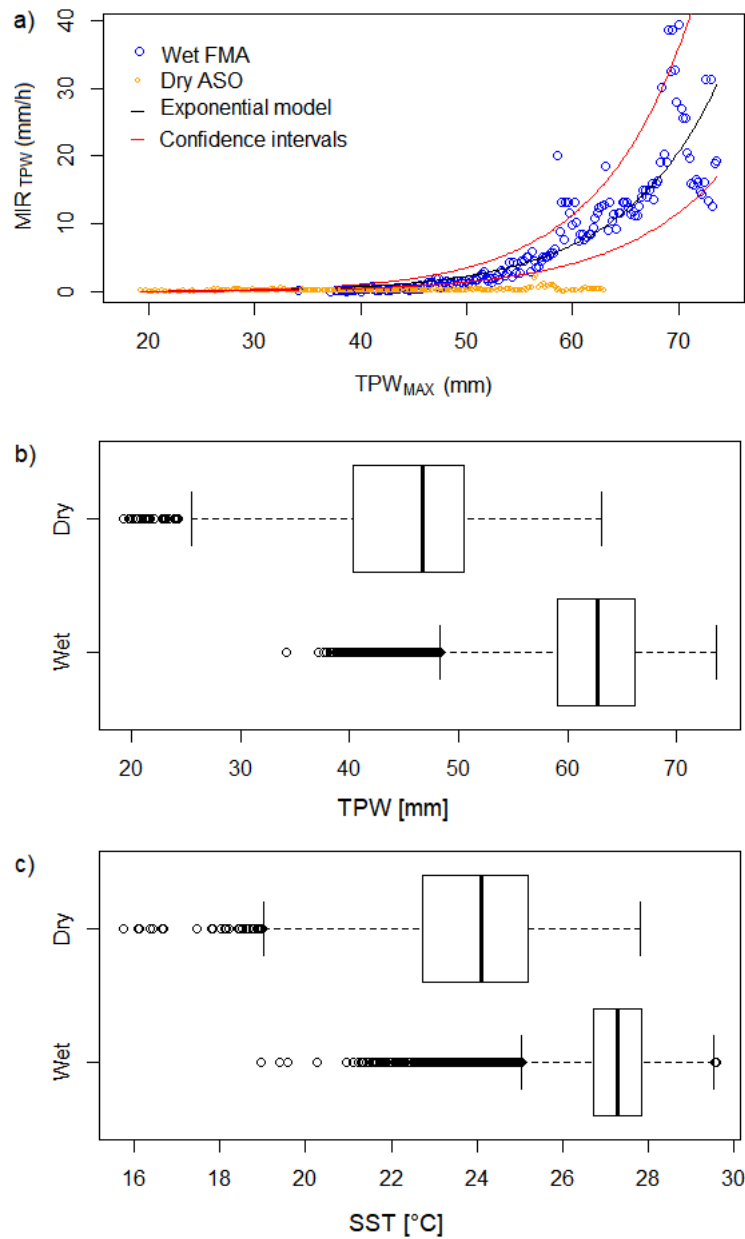
Figure 5 a) MRI values for the wet and dry seasons, separately, associated with the PWV values. The adjusted exponential model and its confidence intervals are presented; b) PWV values of the Dry and Wet season; c) SST values registered in the Dry and Wet season.

2.4.4 Spatial relationships

Figure 6 shows the spatial distribution of the studied variables in the study case of April 5, 2012, from 8:50 am to 8:53 am. In the spatial analysis, a comparison between RID, SST, PWV, the model estimation and the difference between the observed and modeled data are presented.

The first observations indicate that PWV (Figure 6a) is able to locate the rain event, but not SST (Figure 6b) which has a more uniform behavior that is primarily influenced by cold coastal upwelling than by the rain event. Furthermore, it is noticed that SST experiences a decrease in the area beneath the rain event. Three main processes could control this decrease in SST under the rain event: oceanic vertical mixing, air-sea latent heat exchange and advection; all of which are mainly influenced by the convective winds produced by the storm (Vincent et al. 2012). As mentioned previously in Section 4.3, SST is influenced more by seasonal variability than by specific rain events.

María Sheila Fabiola Serrano Vincenti
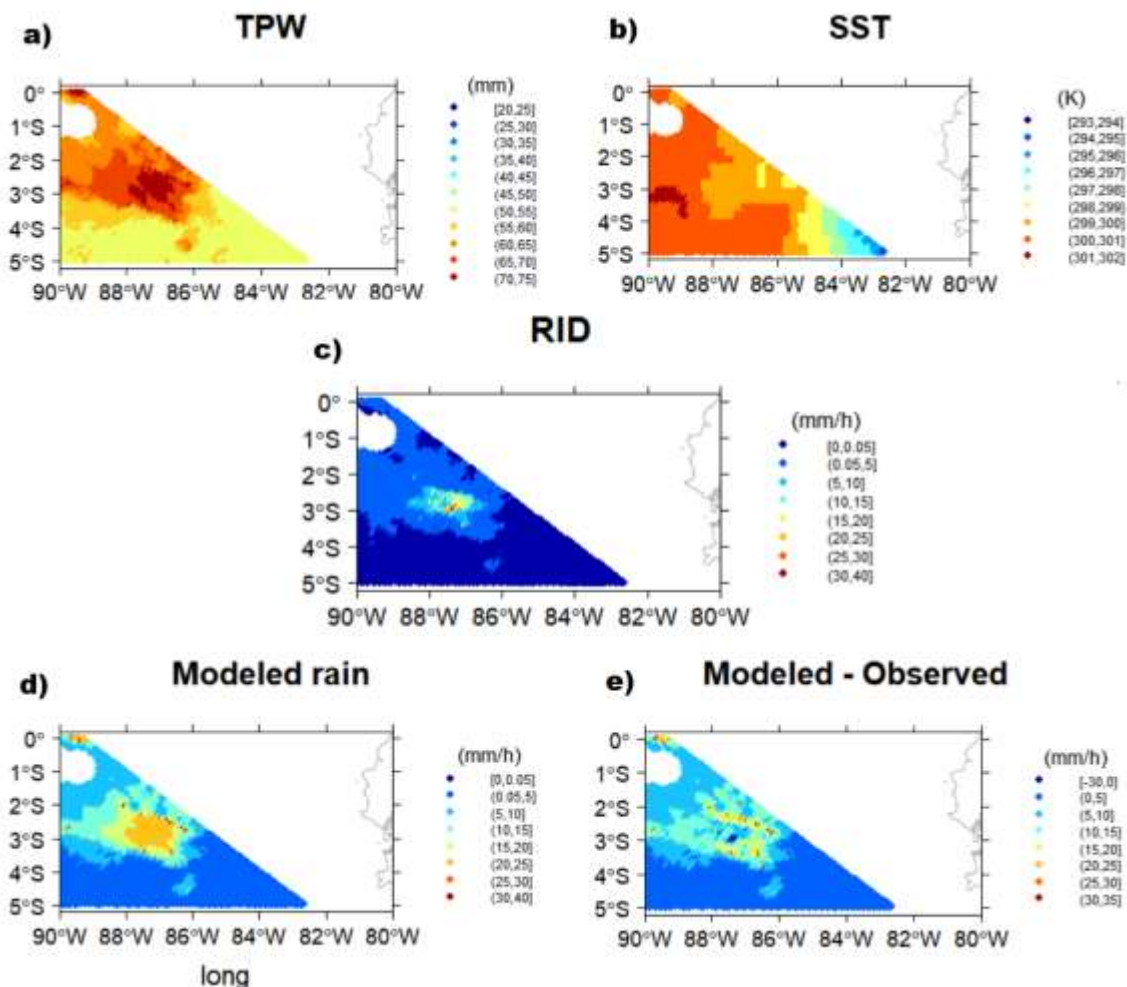
Figure 6 Graphical representation of the satellite scan over the study area on April 5, 2012, from 8:50 am to 8:53 am. The variable analyzed for this rainfall event are a) PWV, b) SST, c) RID, d) maximum rainfall modeled by the exponential model and e) difference between the modeled and observed rainfall.

Figure 6c shows the RID data, Figure 6d presents the maximum modeled rain using the exponential model (Equation 2), and Figure 6e depicts the modeled rain minus the observed RID. Two aspects are evident from this figure: first, the model is able to determine the proper location of the rain event and second, this exponential model overestimates rain events in the surrounding advective zone because the model describes the maximum rain events related to PWV; however, there could be smaller rain events which correspond to the same amount of PWV. The development of a rain process does not depend solely on the amount of PWV. Other dynamic and physical variables must be considered, such as winds, convection and variations in the CAPE (Convective Available Potential Energy). If all of the weather conditions are well combined, heavy rain will develop, but its intensity will be limited exponentially by the amount of PWV in the area.

All of the MRI values associated with PWV > 55 mm were analyzed. Each MRI occurred in 23 rainfall separated events for a total of 29,353 RID (several MRI_{PWV} events correspond to the same rainfall event). In order to quantify the performance of the model, RMSE and the difference between the modeled and observed rainfall are calculated and presented as a histogram and the associated boxplot (Figure 7).
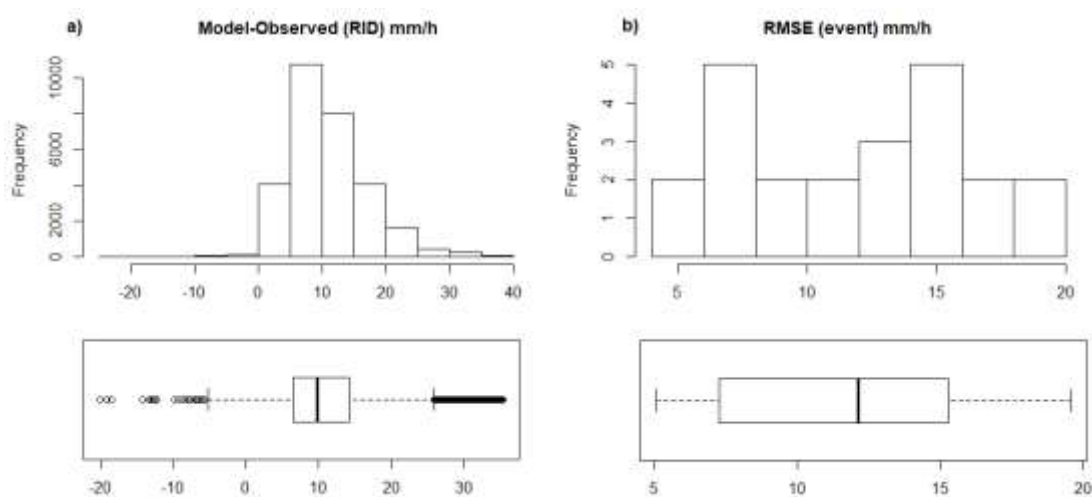
Figure 7 a) Histogram and boxplot of the differences between the maximum modeled rainfall and observed rainfall using 29,353 RID data. The positive values represent an overestimation (correct

model). The negative data represent an underestimation (incorrect model) by 0.4%. b) Histogram and boxplot of RMSE for the corresponding 23 rainfall events.

The differences between the modeled and observed rainfall in Figure 7a shows only 129 negative values, i.e. 0.4%, and therefore 99.6% of the values correspond to an overestimation, which indicates that the model calculating the maximum rainfall values is performing correctly in case studies, where the meteorological conditions for the maximum rainfall development will not always occur. The mean value of these differences represented by the RMSE and calculated for the 23 rainfall events is 11.8 mm/h, the minimum is 5.1 mm/h and the maximum is 19.6 mm/h (Figure 7b).

Therefore, the most important issue related to this model is that the amount of PWV determines the maximum (but not the minimum) rainfall that the atmosphere can precipitate exponentially. To the contrary, the center of the storm is characterized by deep convection (Figure 8), and the model underestimates the rain intensity at around 20 mm/h, but over a very small area. Extreme convective rain is present into the nucleus of the storm, principally due to a mass convergence from the areas surrounding the convective area (Gray, 1998). Although, another factor could be the numerical instabilities around 70 mm of PWV (Kummerow et al. 2001) as mentioned in Section 2.2.2.
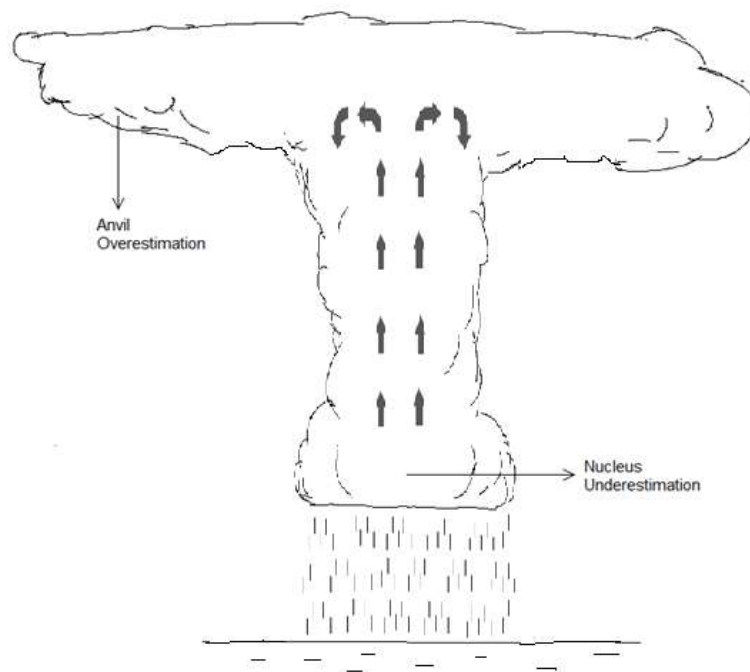
Figure 8 Storm convective cloud diagram, identifying overestimated (nucleus) and underestimated (anvil) areas using the exponential model.

2.5 Conclusions and discussion

In summary, these results empirically show that the maximum amount of rain that the atmosphere can precipitate depends exponentially on the amount of tropospheric water vapor under the specific geographical conditions of this study. This finding has low computational cost implications for improving heavy rain estimations in terms of the location of the event as well as proposing an upper limit for its intensity.

It is useful to evaluate the risk of a possible extreme event of rain related to a very stable and well-predicted meteorological variable such as PWV. In addition, PWV correctly locates the rain event. However, because it is a model for maximum rainfall, it overestimates the rain in the areas surrounding the cloud, corresponding to the anvil advective area. In contrast, an underestimation was reported in its nucleus, supporting literature findings and the fact that strong convection has its own difficulties in modeling as well as in its prediction due to turbulence complexity (Bellenger et al. 2015).

Finally, it is important to consider the two main limitations in this study. The first is the non-continuity of the collected data, both temporally and spatially, which prevents us from finding cause-effect relationships between PWV and rainfall. The second limitation is the cutoff over 70 mm of PWV; this is due to the lack of base data used to estimate precipitation via satellite by using Equation 1, which could reproduce non-physical effects over 70 mm. These two limitations can be overcome with future studies on the current topic, using other non-satellite techniques, and obtaining data that allow continuous series in space and time, in order to determine the cause-effect dynamics between PWV and intense rain, as well as over the ocean and land.

There are several techniques to measure PWV with high confidence over land, such as differential absorption LIDAR (Spuler et al., 2015) and atmospheric emitted radiance, among others. However, the most recommended is to estimate PWV via tropospheric delays derived from the Global Positioning System (GPS), which is proportional to the atmospheric PWV content, and can be measured with high accuracy (± 2 mm) in all meteorological conditions, at low cost and with high resolution (15 min) (Businger et al. 1996; Walpersdorf et al. 2007; Jade and Vijayan 2008; Sapucci et al. 2019;). The rainfall quantification for a specific studied location would be improved by contrasting these GPS-PWV values with rainfall *in situ* measurements and by finding empirical relationships between them (Adams et al., 2013a; Kuo et al., 2017; Sapucci et al., 2019).

The insights gained from this research may be useful to future studies over land that could be used to prove the relationships found between PWV and intense rain (Eq. 3), and its implications over different climates and orographic conditions. Even so, having reliable PWV data on land could, on its own, improve the forecast of regional mesoscale NWP, where PWV could be adjusted *in situ*, creating an assimilation process that will improve the forecast modeling in all of the meteorological variables with a low computational cost.

Chapter 3

Harmonic Analysis of the Relationship between GNSS Precipitable Water Vapor and Heavy Rainfall over the Northwest Equatorial Coast, Andes, and Amazon Regions

This study finds the relationship between increases in precipitable water vapor (PWV), and intense rainfall events in four different climatological regions of South America's equatorial northwest: the coast, Andes valley, high mountains, and Amazon. First, the PWV was derived from tropospheric zenith delay measured by Global Navigation Satellite System (GNSS) instrumentation located near meteorological stations within the regions of interest using hourly data from the year 2014. A harmonic analysis approach through continuous wavelet cross-spectrum and coherence, as well as discrete wavelets, was used to determine a measure of the lags found between PWV and specific heavy rain events and then compared with satellite IR images and meteorological anomalies. The link between PWV peaks and rainfall was the most evident on the coast, and less discernible in the other stations possibly due to local dynamic factors. The results showed a lag of 11 h between the preceding PWV increase and an intense rainfall event. This was apparent in all of the stations, except in Amazon where it was 6 h, with the highest precision at the coast and with the largest dispersion in the high mountains. The interpretation of this lag for each region is also discussed.

Related publication:

3.1 Introduction

The latest technological advances for meteorological applications based on GNSS (Global Navigation Satellite Systems) have received increased attention from the scientific community due to their availability in all types of atmospheric conditions, providing high temporal resolution, reliability, accuracy of the data, being a low-cost solution for the analysis, modeling, and forecasting of extreme rainfall events (Bonafoni et al., 2019; Businger et al., 1996).

The GNSS signals characterize the atmosphere as they pass through it, related to the conditions of the pressure, temperature, and humidity of the troposphere. The troposphere generates variable delays (ZTD for Zenit Total Delay) that can be quantified in the GNSS analysis(Bevis et al., 1992, 1994). This is how the term "GPS Meteorology" is used, which is based on two techniques: the first uses ground-based geodetic GNSS stations (GB-GNSS in this work) that indicate the amount of tropospheric water vapor (PWV for Precipitable Water Vapor, TPW for Total Precipitable Water or IWV for Integrated Water Vapor) precisely over the GNSS station (Yeh et al., 2016). If the GB-GNSS station network is sufficiently dense, it is possible to know the vertical structure of the troposphere through tomographic reconstructions (Brenot et al., 2014; Sguerso et al., 2016; Walpersdorf et al., 2007). The second is a remote sensing technique called RO-GNSS (RO for Radio-occultation), which measures from a Low Earth Orbit (LEO) satellite the refraction effect of the signal of a GNSS satellite rising or settling behind the Earth. This measurement provides vertical profiles of the structure of the Earth's atmosphere. Both techniques have been used in the study of intense rains, convective storms, tropical cyclones, atmospheric rivers, heat waves, and droughts (Bonafoni & Biondi, 2016; Padullés et al., 2022).

The potential of these technologies has led to the generation of several international projects in the United States, Europe, and Asia that seek to improve the constellation of associated satellites, such as the Global Positioning System (GPS, USA), GLObal NAvigation Satellite System (GLONASS, Russia), Galileo (European Space Agency, Paris, France), and BeiDou (China), as well as improving the GB-GNSS network and associated real-time tropospheric devices. As a result, significant improvements have been reported in the forecasting of Numerical Weather Prediction (NWP) models, especially in severe weather conditions, by assimilating observed PWV data with high accuracy and high temporal resolution (Shoji et al., 2009; Risanto et al., 2021).

In this way, not only the modeling of intense rainfall events has been improved by PWV assimilation, but new relationships have also been found between PWV and rainfall

María Sheila Fabiola Serrano Vincenti

characteristics. Several studies mention the occurrence of PWV thresholds (H. Li et al., 2020; L. Li et al., 2022)and/or peaks before heavy rain events (Labbouz et al., 2015). Extreme rain is commonly defined as rain which exceeds high percentiles, 95th, 97th, or 99th, of the distribution of accumulated frequencies of daily precipitation, although this definition depends a lot on the climatic characteristics of the place (Pendergrass, 2018). Thus, investigations with hourly resolution data study windows of 6 h for the location of PWV peaks before intense rain events on the eastern coasts of China (at around 170 m above the sea level) (Yao et al., 2017; Zhao et al., 2020) a 6 h window was also found in Lisbon, Portugal (~125 m.a.s.l.) (Benevides et al., 2015, 2019), and 3 h windows in various locations in Italy (Bonafoni & Biondi, 2016). Likewise, using finer resolutions (1 min.) in Sao Paulo (Brazilian Amazon) (~85 m.a.s.l.), PWV peaks were found 32–36 min before heavy rain events (Sapucci et al., 2019), while lags of 4 and 8 h, respectively, between PWV peaks and shallow-to-deep convection transitions were reported (Adams et al., 2013b). Finally, researchers in the Argentina Valley of Mendoza (~769 m.a.s.l.), detected PWV increases before heavy storms but not a conclusive lag (Calori et al., 2016).

However, the evolution of the PWV time series is different for each reported case, both before and after the rainy event and no generalizations can be made about the time between the peak of PWV and the start of heavy rain (Bonafoni et al., 2019). In fact, all studies have reported significant increases in PWV that have not preceded any rain event, showing that the increase in PWV is just one more of the conditions for the generation of intense rain (Benevides et al., 2015). Therefore, it is necessary to analyze additional events in regions with contrasting climates and on complex topographies, in order to take into account the effect that altitude and different synoptic conditions may have on the behavior of the PWV; this us to help draw general conclusions regarding the establishment of critical PWV values or other dynamic indicators.

In this context, the present study has been conducted in Ecuador. Despite its small area, it has a diverse topography with three climatically different regions: the coast, the Andes, and the Amazon region. Additionally, it has a moderately dense network of GB-GNSS stations distributed at different altitudes, making it an ideal place to improve the understanding of the relationship of PWV with heavy rainfall events.

Note that previous studies conducted with GNSS-based meteorological applications in this area have been scarce and involve only daily or monthly resolutions (Ayala et al., 2018; Romero et al., 2019). The present research, in contrast, uses hourly resolutions to contribute to the understanding of the sub-diurnal dynamics of the aforementioned regions. It combines

for this purpose a time-frequency harmonic analysis with statistical methods in the form of wavelet coherence and wavelet cross-spectrum analyses, while taking into consideration the dynamic and climatic characteristics of each region.

3.2. Geographical Location of the Measurement Network and Methods

3.2.1. GNSS-ZTD Data

Data used for the present analysis has been obtained from the National Geodetic Network RENGEO of the Geophysics Institute of the National Polytechnic School (Instituto Geofísico de la Escuela Politécnica Nacional IG-EPN) https://www.igepn.edu.ec/red-nacional-de-geodesia, which is continually updated in the database of UNAVCO https://www.unavco.org/data/gps-gnss/gps-gnss.html, accessed on 21 march 2019.

First, in this database, there were 17 stations that corresponded to the geographical regions of interest and had complete data allowing ZTD estimation with hourly resolutions. The selection was narrowed down to five sites, which were the ones close to meteorological stations with available hourly data for the year 2014. These GNSS-selected stations are located in the coastal region (SNLR), Andes region (BELI, IBEC, ASEC), and the Amazon region (TEN1) (See Figure 9). The location and technical specifications of these stations are found in Table 3 and Table 4, respectively.
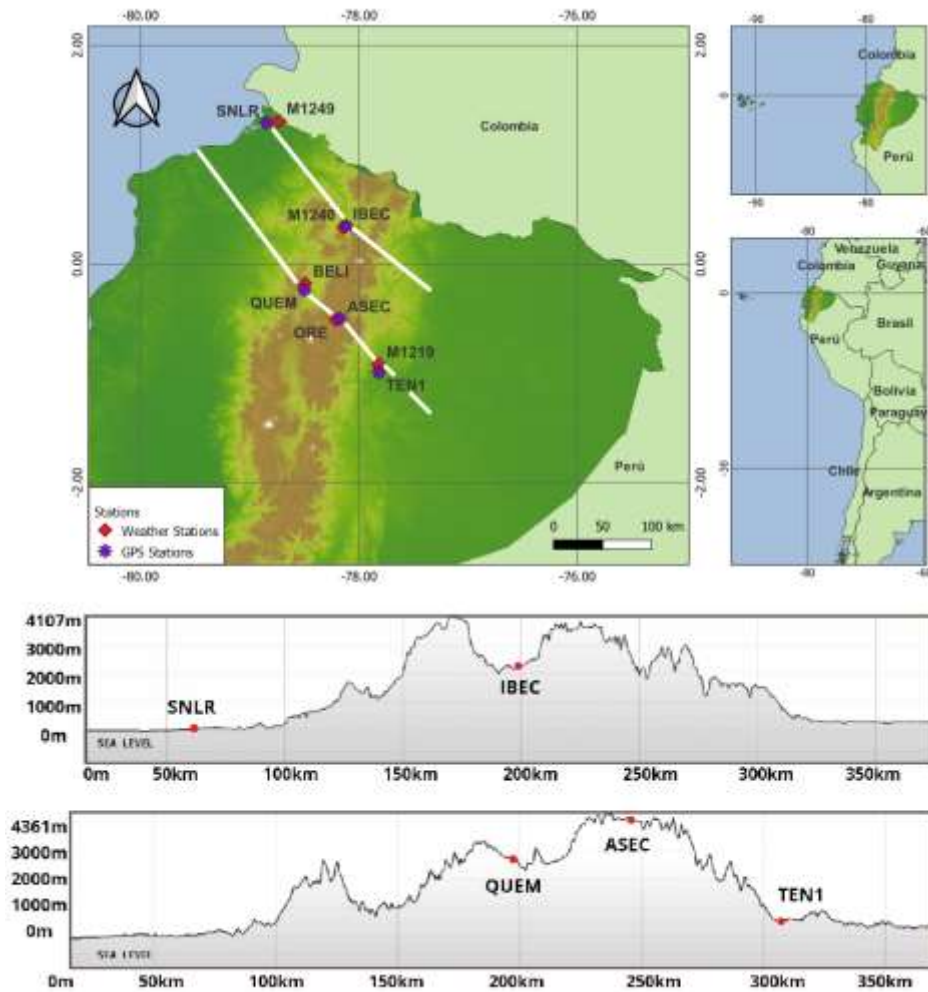
Figure 9 (Top) GPS stations where the tropospheric delays were calculated in purple and their nearby meteorological stations in red. (Bottom) GPS stations alongside the altitude profile of the segments marked in the map in white (See Table 3 for details).

Table 3 Weather station coordinates, altitude, and distance to the nearest GPS station in kilometers.

Weather Station (Code)	Region	LON (°)	LAT (°)	Altitude m.a.s.l	Variables	Time Resolution (Time Span)	Nearest GPS Distance (km)
Antisana (ORE)	High Mountain	− 78.2112	− 0.5092	4059	Rain (mm/h) Temperature (°C) Rel. Humidity (%) Wind speed (m/s)	30 min (2005 - 2018)	ASEC (4.75)
Tena (M1219)	Amazon	− 77.8190	− 0.9168	553		Hourly (2014 - 2020)	TEN1 (8.18)
Esmeraldas (M1249)	Coast	− 78.7316	1.30583	45	Rain (mm/h) Temperature (°C) Rel. Humidity (%) Wind speed (m/s) Pressure (hPa)	Hourly (2014 - 2018)	SNLR (12.97)
Quito (BELI)	Andes valley 1	− 78.49	− 0.18	2835		Hourly (2013 - 2020)	BELI (6.42)
Ibarra (M1240)	Andes valley 2	− 78.1397	0.33388	2247		Hourly (2014 - 2020)	IBEC (3.4)

Table 4 Technical specifications of the weather and GNSS stations instrumentation.

IRD-INAMHI Antisana (High Mountain)		
Parameter	Sensor Type	Accuracy
Precipitation (kg m^{-2})	Geonor T-200B, Davis rain collector II	±0.1 × 10^{-3} ± 0.2 × 10^{-3}
Air temperature (°C)	Vaisala HPM45C, ventilated	±0.2
Relative humidity (%)	Vaisala HPM45C, ventilated	±2 on [0–90]; ±3 on [90–100]
Wind speed (m/s)	Young 05103	±0.3
INAMHI: ESMERALDAS (Coast), TENA (Amazon), IBARRA (Andes valley 2)		
Rain (mm/h)	TR525M Texas	±0.1 (mm)
Air temperature (°C)	HPM155D Vaisala	±(0.055 + 0.0057 × T) °C
Relative humidity (%)	HPM155D Vaisala	±1%RH on [40–95%]
Wind speed (m/s)	WMT 702D Vaisala	±0.8 (m/s)
Atmospheric Pressure (hPa)	Vaisala PTB110	±0.3 (hPa)
REMMAQ: BELISARIO-QUITO (Andes valley 1)		
Rain (mm/h)	AQMR25–Vaisala	±5%
Air temperature (°C)	AQMR25–Vaisala	±0.3 °C
Relative humidity (%)	AQMR25–Vaisala	±3%RH
Wind speed (m/s)	AQMR25–Vaisala	±0.3 m/s
Atmospheric Pressure (hPa)	AQMR25–Vaisala	±0.5 hPa
IG-GNSS-EPN		
GNSS	Trimble NetRS, NetR8 and NetR9	15 and 1 seg (volcanoes) 30, 1 and 0.2 seg (tectonic structures)

After the selection process, the ZTD time series were estimated by processing the GNSS data using the GAMIT/GLOBK software (Herring et al., 2018).

3.2.2. Meteorological Data

Weather data from the year 2014 were obtained from locations close to the selected GNSS stations. These weather stations were sparse and managed by the National Institute of Meteorology and Hydrology of Ecuador (Instituto Nacional de Meteorología e Hidrología—INAMHI), in a particular case in cooperation with the French Institute of Research for Development (Institut de recherche pour le développement, IRD) (Wagnon et al., 2009), and by the Quito Atmospheric Monitoring Metropolitan Network (Red Metropolitana de Monitoreo Atmosférico de Quito—REMMAQ). Figure 9 shows the location of these weather and GPS stations on the map and Table 3 contains their geographic coordinates with additional information such as altitude, monitored weather parameters, a time resolution of the data, and distance to the closest GPS station.

3.2.3. Calculating ZTD from GNSS Data

The GNSS data processing is carried out through the GAMIT/GLOBK R10.7 software using a doubly differenced approach to eliminate clock errors from satellites and ground receivers in addition to a linear combination L3 (also called LC) of the GNSS phase signals in the L1 and L2 band to mitigate the errors introduced by the ionosphere. Only GPS data have been employed, as the receivers had no multi-GNSS capacities at the time of this study. Due to the fact that GAMIT determines the position of the GPS-GB stations with great accuracy and precision in order to quantify tectonic displacements, several models are taken into account for its computation, to separate tectonic deformation from other geophysical features. Modeled are for example the deformation of the earth's crust due to the ocean tides, also known as ocean loading, atmospheric loading, and the influence of solar and lunar movements on Earth's atmosphere (atmospheric tides) (Herring et al., 2018).

GAMIT parameterizes the ZTD as a stochastic variation of the Saastamoinen model, with a stepwise linear interpolation. The variation is calculated over a first-order Gauss–Markov process, with a potential density of 2 cm^2/h, known as the "zenith restriction parameter". A sliding window strategy is used to avoid offsets at the limits between two successive sessions in order to prevent day boundary discontinuities of the satellite orbits, e.g., (Brenot et al., 2014). Two 24 h sessions are calculated, starting at 0 h UTC and 12 h UTC, respectively.

From both sessions, we keep the central 12 h to obtain a continuous well-constrained time series of ZTD for each station across the day boundary. The first and last 6 h of each analysis session are removed in order to take away the less well-constrained values at the session's start and end while adjusting the final IGS orbits to smooth the day boundary discontinuity (Brenot et al., 2006). The GAMIT analysis producing these hourly ZTD values for each station was conducted at the Institute of Earth's Sciences, ISTerre, of the University Grenoble Alpes, using the ISTerre computing cluster IST-OAR.

3.2.4. Quality Control and Data Synchronization

Once continuous ZTD data was obtained through the GAMIT sliding window analysis, repetitions of the values at session boundaries at 06:00 and 18:00 UTC, which account for 1% to 2% of the total data, were eliminated. In this way, it was possible to generate complete hourly ZTD series for the entire year 2014 at the Quito (Andes valley 1) and Antisana (High Mountain) stations. For the three remaining GB-GNSS stations, available data was found only from 21 July 2014 at 16:00 until the end of the year.

The quality of the hourly meteorological data from the in-situ stations required a thorough treatment, which was carried out on the available data series for each meteorological parameter rainfall: temperature, atmospheric pressure, wind speed, and relative humidity. The data series covered thirteen years for Antisana (High Mountain), seven years for Tena (Amazon), Esmeraldas (Coast) and Ibarra (Andes valley 2), and eight years for Quito (Andes valley 1). This amount of available data was used to determine the average values of the meteorological variables in order to determine the mean meteorology of the site presented in Section 3.3.1.

The quality control carried out for these meteorological data series was performed in accordance with the standards for hourly data treatment (Climate World & Programme, 1986). It was carried out first on the bias error, which corresponds to a discontinuity in the moving averages of the data, frequently due to the deviation of the sensors, which must be corrected by adding a constant value to the discontinuous period so that the moving averages are maintained (Lanzante, 1996).

Next, the extreme discontinuous values (outliers) were eliminated, defined as those values which were outside the interval $[\mu \pm 4\sigma]$ (Lanzante, 1996), with the exception of the value of rainfall, which, being characterized by non-normal distributions (Cho et al., 2004; Serrano-

Vincenti et al., 2020) could present higher rainfall intensities. In this case, values greater than [μ + 10σ] were eliminated (Lewis et al., 2018). Once the outliers were eliminated, no padding was performed in these meteorological data.

Only in the case of the Antisana station (high mountain), the surface atmospheric pressure was estimated since it was not available in the corresponding weather station. The estimation used the time series of the surface atmospheric pressure of the Quito station instead (Andes valley 1), given its similar mean hourly homogenous behavior, but subtracting the value of 156.1 hPa, corresponding to the difference of the average pressures calculated by height differences (Garratt, 1992) (Table 5).

Table 5 Accumulated rain, annual average, and standard deviation of the studied hourly meteorological variables for the year 2014. NA% is the percent of data loss per variable and * means that the data for these stations were only available from 21 July to 31 D

Variable	Coast (Esmeraldas) 45 m.a.s.l.	NA% *	Andes Valley1 (Quito) 2835 m.a.s.l.	NA%	Andes Valley2 (Ibarra) 2247 m.a.s.l.	NA%*	High Mountain (Antisana) 4059 m.a.s.l.	NA%	Amazon (Tena *) 553 m.a.s.l.	NA% *
Accum.rain (mm/year)	2640	5.6	1224.5	9.4	1098.76	5.6	415.7	0	3700	33.5
Mean PWV ± SD (mm)	58.2 ± 3.9	4.3	17.6 ± 1.7	5.9	24.07 ± 3.8	4.3	21.52 ± 2.2	4.5	46.3 ± 4.5	7.7
Mean hourly Temperature ± SD (°C)	25.3 ± 1.6	5.6	11.9 ± 3.3	9.4	16.32 ± 3.8	5.6	1.02 ± 1.68	5.8	23.1 ± 3.2	31.8
Mean Relative Humidity ± SD (%)	94.3 ± 3.8	6	78.1 ± 14.4	12.6	74.9 ± 17.4	6.1	81.9 ± 12.9	5.8	89.3 ± 11.4	31.8
Mean atmospheric pressure (hPa)	1001.4 ± 1.7	5.6	708.1 ± 0.56	9.4	780.6 ± 1.4	5.6	552 ± 1.24	9.4	951.5 ± 2.5	48.6
Mean wind speed ± SD (m/s)	1.5 ± 1	6.1	2.46 ± 0.63	17.3	1.6 ± 1.25	6.1	4.56 ± 4.15	9	0.8 ± 0.6	48.9

Finally, the hourly time series of the ZTD was changed from UTC to local time, and common work periods were selected in order to synchronize the meteorological data with the ZTD.

3.2.5. Estimation of the PWV from the ZTD

The Total Zenith Delay ZTD depends on the refractivity that the GNSS signal encounters when crossing the atmosphere. This delay has two components: a Zenith Hydrostatic Delay ZHD due to a hydrostatic or total troposphere, which generally explains 90% of ZTD (Fernández L.I., Meza A.M., 2009), and a wet component or Zenith Wet Delay ZWD due to the tropospheric water vapor, which is a strong variable and related to the PWV that explains the remaining 10% of the total delay. Equation (4) (Bevis et al., 1992, 1994):

$$ZTD = ZHD + ZWD \qquad (4)$$

ZHD in [m] is obtained from the surface atmospheric pressure P_s in [hPa] by

$$ZHD = (2.2768 \pm 0.0015).\,10^{-5}P_s/f(\theta, H)$$

and,

$$f(\theta, H) = 1 - 0.00266cos2\theta - 0.000279H$$

where θ is the latitude, and H is the ellipsoidal height in [km]. Calculating ZHD from surface pressure and the station position, ZWD can be isolated in the ZTD measurement (Equation 4). According to Bevis et al. (Bevis et al., 1994), ZWD has a relationship with the PWV, given by:

$$PWV = \kappa \cdot ZWD$$

The κ constant depends on the vertical integral of the vapor density and the temperature, and it has a 2% precision (Brenot et al., 2014). In similar climatology studies, a $\kappa_{E\&D}$ constant is used as presented by (Emardson & Derks, 1992), which has a precision of 1.15% and has the advantage that it only depends on the surface temperature T_s:

$$\kappa_{E\&D} = \frac{10^3}{6.324 - 0.0177\,(T_s - 289.76) + 0.000075\,(T_s - 289.76)^2}$$

Therefore, PWV can be estimated from GNSS ZTD measurements and meteorological ground pressure and temperature observations as in Equation (5):

María Sheila Fabiola Serrano Vincenti

$$PWV_{E\&D} = \kappa_{E\&D} \cdot ZWD \qquad (5)$$

3.2.6. Conditioning of the Resulting PWV and Rain Data for Harmonic Analysis

Further analysis of the PWV and rain data required some data conditioning to assure the standardization and correct interpretation of the results while working with these two different magnitude time series: rainfall and PWV, using harmonic analysis tools.

The time series was normalized, dividing the values of the series by the maximum value recorded in order to preserve their original envelope, especially in the case of rainfall where sudden increases and the different intensities of the rainfall events where noticeable.

The choice of wavelets as the main tool to analyze these observables made it necessary to inspect if any incomplete or corrupted data was going to affect the periodicities of the series. Thus, the incomplete data were padded with linear interpolation in the case of PWV and with zeros in the case of rainfall, which correspond to the mode of this data. This padding was made between the available samples at each end of every empty (missing) piece of data in order to preserve the smoothness of the series rather than creating harsh discontinuities and, at the same time, minimizing the creation of artificial events. Figure 10 shows the time series of rain and PWV after conditioning per station. Regions padded with data are those marked with vertical gray lines.
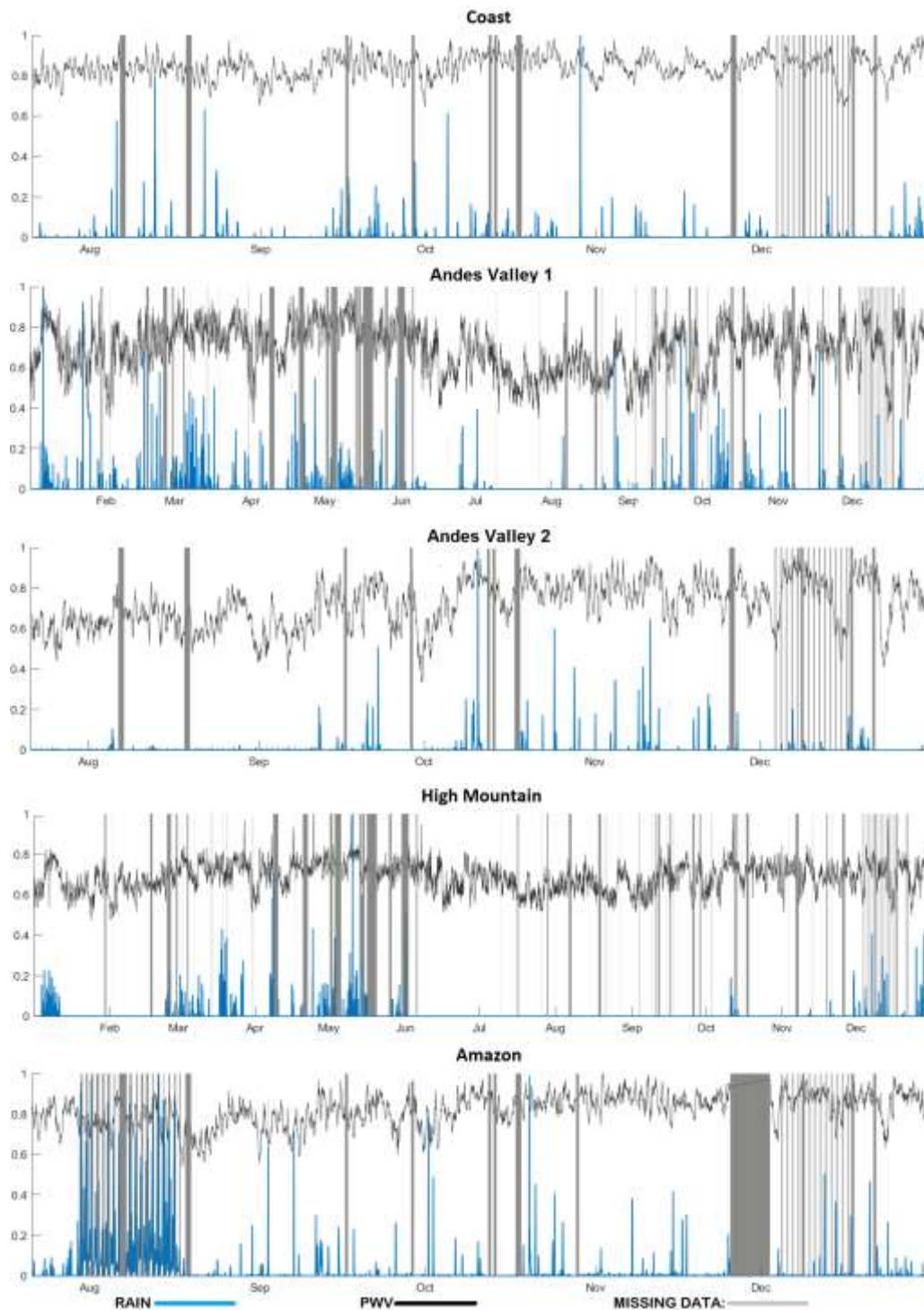
Figure 10 Normalized time series of rain and PWV after data conditioning in 2014. Missing data is shown in gray (See Table 5 for details).

3.2.7. Harmonic Analysis by Descriptive Statistics and Wavelets

The presented methodology uses the harmonic analysis of the time series via wavelets, a tool of great interest within applied mathematics that has allowed novel applications in several areas of knowledge, including those related to the study of natural phenomena (Torrence & Compo, 1998; Grinsted et al., 2004; Yang et al., 2006). The intrinsic relationship between time and frequency helps us to obtain information from rainfall and PWV data by visualizing the elementary atoms, or values of interest that are well concentrated in time and frequency, the variability of the signals, the analysis with various levels of detail of the signals by means of a change of scale (zooming) of time or frequency, and the characterization of signal structures such as singularities (Mallat, 1999).

Additionally, it is recommended to read (Chui, 1992; Daubechies, 1992) for a deeper understanding of this tool. Note that the wavelet and statistical analysis was performed in Matlab R2021a (Mathworks, 2021) using built-in and third-party libraries as described later.

3.2.7.1. Continuous Wavelet Analysis: Transform, Coherence, and Cross-Spectrum

One of the objectives of this work is to find a correlation in the time-frequency domain between rain and PWV, that is, to determine the periods in which these two meteorological variables interact in a coordinated manner and in a localized period of time. The most appropriate tools for the effect are the wavelet coherence and cross-spectrum (Grinsted et al., 2004), as they provide the phase information between these signals. However, the continuous wavelet transform CWT is required for the calculation of any of them, and therefore becomes a previous step which enables the visualization of the harmonic content of the signals, that is, frequencies of interest that could provide additional or complementary information.

The computational analysis, data conditioning, simulations, and the plot of the results were performed with the aid of the Wavelets toolbox, Signal Processing toolbox, and Descriptive Statistics toolbox libraries (MathWorks, 2021), as well as the toolbox of Cross-wavelet and Wavelet Coherence by Grinsted (Grinsted, 2014).

3.2.7.2 Continuous Wavelet Transform

As a first step, the continuous wavelet or CWT transform was implemented to identify the harmonic behavior, if any, of both rain and water vapor. In essence, this transform decomposes a signal on a particular continuous wavelet wave.

A wavelet is a function $\psi(t)$ of zero average,

$$\int_{-\infty}^{\infty} \psi(t)\, dt = 0 \qquad (6)$$

This function $\psi(t)$ is dilated with a scaling parameter "s" and translated in time with a parameter "u":

$$\psi_{u,s}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right) \qquad (7)$$

So, the wavelet transform W of a function $f(u,s)$, on the scale "s" and time position "u", is calculated by correlating f with the wavelet signal (Mallat, 1999).

$$Wf(u,s) = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{s}} \psi^*\left(\frac{t-u}{s}\right) dt \qquad (8)$$

As it is observed, $\psi(t)^*$, is the complex conjugate. $Wf(u,s)$ can therefore measure the variations of spectral components just as the windowed Fourier transform or spectrogram does (Benedetto, 1997), but with a different time and frequency resolution. This is how through this CWT, very well-localized time and frequency packages can be determined but with the limitations of Heisenberg's uncertainty principle (Heisenberg, 1927). For a mathematical explanation of this principle applied to time and frequency, see (Benedetto, 1997; Mallat, 1999).

These packets that concentrate the energy of both time and frequency make up what is graphically known as a scalogram, which allows an identification at different scales of how the function f varies. If the "s" scale decreases, time support also decreases but the frequency range increases towards high frequencies. Similarly, if the "s" scale is dilated, the frequency range is located in the low-frequency range, but the support of the transform over time is enlarged.

The scalogram applied to the rain and PWV series, which are discrete functions in time (with hourly intervals in our case), would indicate in each case if they have harmonics or modes of interest throughout the year 2014. That is, the busiest intervals of repetitive events, whether they would be daily, weekly, monthly, or seasonal.

For our study, the scalograms were obtained by means of the bump-type analytical wavelet (Mathworks, 2021), given the fact that it narrows the variance of the CWT in frequency and thus allowing us to visualize with more finesse the periods of repetition of rain or PWV during certain time intervals of the series.

3.2.7.3 Bivariate Analysis with the Continuous Wavelet Transform

According to (Rösch & Schmidbauer, 2018), if the CWT is applied to two time series, there is a geometric interpretation analogous to the covariance, which allows finding the common frequencies of these series, their location in time, and the common spectral power between these variables in two-dimensional graphs known as wavelet cross-correlation spectrum. It is important to mention that the covariance, unlike the correlation which is a value between −1 and +1, can take any value. Equation (9) determines the CWT of two time series x_t and y_t, after applying the individual wavelet transform on each of them, obtaining the functions $W_x(u, s)$. y $W_y(u, s)$ where u and s are the time-location and frequency scale parameters, respectively, as it was noted in Equation (8). Therefore, the bivariate wavelet is:

$$W_{xy}(u, s) = S\left(W_x^{\star}(u, s).W_y(u, s) \right) \qquad (9)$$

Note that the asterisk represents the complex conjugate of the wavelet transform of the first series and S is a smoothing operator in time and frequency (Mathworks, 2021). It can be clearly seen from this equation that the values to be operated on are complex. This is because the bivariate CWT uses an analytical (or complex) wavelet, thus obtaining the typical definition of the covariance of two complex random variables (Park, 2018).

Then, we define the cross-wavelet power spectrum XWT in the following fashion:

$$\text{XWT}(u, s) = \left| W_{xy}(u, s) \right| \qquad (10)$$

the instantaneous lead phase, or relative phase angle between x_t and y_t, in the time-frequency space, is determined by calculating the complex argument of the bivariate transform. See (Grinsted et al., 2004).

$$\theta(u, s) = \arg\left(W_{xy}(u, s) \right) \qquad (11)$$

The libraries allow this phase to be calculated within the cross-spectrum plot through arrows, with the interval [−π, π]. An absolute value less than π/2 indicates that the two series tend to

move in phase, having an ideal phase when the arrows point to the right horizontally. If the series tend to lag one another, the absolute value of the angle must be greater than π/2. If both signals oppose each other, or are in a complete out-of-phase, then the arrows are horizontal but left-oriented (Percival & Walden, 2000).

From Equations (10) and (11), it can be deducted that the XWT allows unveiling common high energy regions between the series, in our case rainfall and PWV, in addition to providing valuable information on the relationship of phase between these two. In other words, this is an important tool to be applied for the lead–lag analysis of the series of interest.

The libraries used the Morlet wavelet for the calculation of the XWT, since it provides an ideal balance in a location both in time and in frequency; this is because it has a complex exponential or carrier multiplied by a Gaussian or envelope window. This type of wavelet is also known as Gabor (Mallat, 1999).

Note that a bivariate wavelet analysis benefits remarkably when evaluating the continuous wavelet coherence WTC, since the XWT cross-power spectrum is insufficient to analyze the correlations between a pair of processes in the time-frequency space, especially when the series has flat characteristics due to data loss or padding, such as in the case of the PWV. Their respective counterpart, on the other hand, presents pronounced peaks, as is the case of the rain. While the XWT reveals only the areas with the highest common concentrations of energy, the coherence reveals the local correlation of the two series in time and frequency throughout their extension and their localized phase behavior.

The continuous wavelet coherence WTC is obtained by normalizing the XWT, and it is appropriately known as the magnitude square coherence, and can be written as follows:

$$WTC_{xy}(u,s) = \frac{\left| S\left(W_x^{\star}(u,s) \cdot W_y(u,s) \right) \right|^2}{S(|Wx(u,s)|^2) \cdot S(|Wy(u,s)|^2)} \quad (12)$$

The smoothing operator "S" is equally present in the WTC in the same way as in Equation (9). This operator as explained in (Torrence & Compo, 1998) performs a time smoothing followed by a frequency smoothing, although only time smoothing is possible according to (Cohen, 2011). Paying close attention to the WTC definition, it can be noticed that its structure is similar to the traditional method of obtaining the statistical correlation coefficient.

Therefore, the WTC is a complementary tool that serves to better analyze the behavior of rainfall and PWV, even in time and frequency segments that were not necessarily those

María Sheila Fabiola Serrano Vincenti

detected by the XWT, but that could render useful information on the dynamics of these meteorological variables.

3.2.7.4. Simple Statistical Correlation as a Reference Parameter

As a precaution, the proposed wavelet tools cannot be applied blindly, but rather the data sets must be compared or tested with other models or mechanisms applicable to the process (Grinsted, 2014). In our study, the control parameter for contrasting the results obtained by the XWT is the simple statistical cross-correlation (Stoica & Moses, 2005) between the discrete vectors of the rainfall and PWV time series, since it would yield the lead of one signal over the other in units of time, where the maximum correlation coefficient is calculated. No previous treatment to the application of the cross-correlation was included because the stationarity of the specific rainfall events series is preserved (Guevara Díaz, 2014).

3.2.7.5. Selection of Events of Interest by Their Statistical Significance Boundary

An event of interest is defined for this study as an intense rainfall occurrence, or in other words a sudden increase or peak in the hourly accumulated rain in the year 2014, which are selected for further analysis in order to determine its covariance with an associated PWV event preceding in time. Table 6 shows the max. rain occurrence, rain thershold and percentile of the most prominent events per station analyzed in this study.

Table 6 Characteristics of selected events of interest. * Points the series with data from 21 July to 31 December 2014. ** Indicates that the six chosen events correspond to the most intense as well.

Variable	Coast * 45 m.a.s.l.	Andes Valley 1 2835 m.a.s.l.	Andes Valley 2 * 2247 m.a.s.l.	High Mountain Antisana 4059 m.a.s.l.	Amazon Region * 553 m.a.s.l.
Events with XWT and CWT/Num. selected events	6/10 = 60% **	9/14 = 64.3%	8/15 = 53%	7/9 = 77%	6/9 = 66.7%
Maximum hourly accumulated rainfall event (mm/hour)	27.5	19.8	16.8	9.5	49.1

Rain threshold for the selected events (mm/hour)	12.9	8.4	3.6	3.9	35.1
Percentile of threshold in the rainfall series	99.4th	98.4th	96th	97.4th	98.4th
95th percentile (mm/hour)	4.4	5.2	3.01	3.1	18.6
50th percentile (mm/hour)	0.2	0.4	0.1	0.5	0.5

The wavelet coherence WTC algorithm described in (Grinsted et al., 2004; Torrence, 1999) was used to validate such events of interest because it allows the plotting of contours with a level of statistical significance of 5%. This is done by constructing a base background spectrum with white or red noise, on which the true spectrum of the wavelet coherence will be compared. These normalized-based spectrums, which are Chi-square distributed χ_2^2, allow the construction of the statistical significance test. The premise, for a significance level α = 0.05 in our case, is that if the power of the WTC of the rainfall and PWV series is less than the normalized power of the wavelet coherence of two simulated base spectrum series, then the WTC wavelet power of the true signals is only generated due to random phenomena.

If instead the spectral power of the WTC of rainfall and PWV is greater than the spectral power obtained from operating the WTC of the series simulated with noise, the null hypothesis is discarded with a confidence level equal to (1 − α), 95% in our case, and it is inferred that the power of the WTC of the true signals is significantly larger than that generated by noise. The mathematical steps on how to obtain the statistical significance of the wavelet coherence WTC are very well explained in (Ge, 2008).

The WTC library (Grinsted, 2014) generates the surrogate noise series using data pairs which are the coefficients of the same lag-1 autoregressive model named AR(1). The WTC is calculated for each surrogate pair and with a Monte Carlo simulation approach the statistical significance level is estimated (Cohen, 2011). This library uses a default alpha α equal to 5% (0.05) for the significance test, in view of the fact that it provides an appropriate statistical balance to prevent Type 1 and Type 2 errors. By applying the wavelet coherence WTC, regions in time and scale (periods) can be identified with a confidence level above 95%. The time domain of these regions was used to locate the corresponding rainfall events of interest, in the time series, with the larger magnitudes or peaks.

María Sheila Fabiola Serrano Vincenti

3.2.7.6. MRA Lead–Lag Analysis for the Events of Interest

Although the continuous wavelet spectrum plots obtained through XWT and WCT have important phase information that allows us to estimate the lead or lag that the rainfall presents with respect to the PWV, it is not without the compromise of accuracy since these plots can represent entire regions of energy spanning several days, even weeks or months along with a wide range of periods in the logarithmic scale. The phase accuracy can only be improved as far as the size of the analyzed energy atom can relate to the angle of orientation of a nearby arrow. It can be argued that the continuous bivariate analysis of the XWT and WCT is not meant to provide a quantity in particular, but the dynamic behavior of two signals for a particular time-frequency range or, better said, a region of significance.

Therefore, a multiresolution analysis or MRA method was used to obtain a quantifiable result of the phase content of the two series with hourly resolution. MRA refers to the method of decomposing a signal into sub-components or frequencies; there are many methods to perform signal decomposition and one of them is through discrete wavelets transform or DWT. There are several advantages of using the DWT over a CWT; for instance, faster computation and a wide choice of wavelets to choose from. A good reading for MRA fundamentals and examples can be found in (Meyer, 1995).

The MRA decomposition of the rain and PWV series will produce different detail levels or one-dimensional scales (frequencies) for each one of them. The interesting fact is that these one-dimensional scales can be used in combination with a simple statistical cross-correlation approach to obtain the respective lag between a pair of rain and PWV scales. The larger the wavelet coefficients of the scale are, the more prominent the correlation coefficient would be and hence marking the lag between the different scales in hours.

Under this context, 20 events of interest were chosen per station which meet the following conditions: first, they have to correspond to the largest accumulated hourly rainfall data points that were detected per station; and second, that they shared with the PWV both coherence and the largest joint distribution energy over 95% confidence. Then, the final selection consisted of keeping mutually exclusive events that were not part of the same rain occurrence; in other words, that were at least 24 h apart in order to prevent overlapping with each other.

Other authors have used the cross-correlation of MRA scales of rain and PWV time series as in the case of (Sapucci et al., 2019); however, with different time resolutions and possibly with

different software tools. Note however that the results obtained in the present study are supported by a thorough continuous wavelet analysis (via cross-spectrum and coherence) and controlled by statistics in the time series as a way to make the lag selection process more objective and reliable.

3.2.8. Convection Analysis Using Satellite Images

Once the events of interest were obtained, the height of the clouds was analyzed one by one to identify deep convection through the GOES-13 satellite images in the infrared (IR) bands (10.2 to 11.2 μm), available every 30 min, both day and night, for Latin America from the Center Weather Forecast and Climate Studies (CPTEC) of the Government of Brazil available at http://satelite.cptec.inpe.br/acervo/goes16.formulario.logic?i=en, accessed on 21 march 2022 (DSA, 2020).

More than half of the incoming solar energy is absorbed by the planet's surface is re-emitted as heat to space. As some of this energy passes up through the atmosphere clouds and atmospheric gases absorb a portion of this radiation. Then, the satellite IR channel senses radiation emitted by the earth's surface, atmosphere, and cloud tops.

The IR images are colorized to bring out details in cloud patterns. In the colorized imagery used here clouds that have cloud-top temperatures colder than −80 °C are shown in pink and cloud-tops warmer than −30 °C are shown in orange. The colder the temperature, the taller the cloud-top height, and the relationship between temperature and cloud height is used to identify high, medium, or low clouds, with the more prominent heights being cumulonimbus clouds. Taller clouds correlate to more active weather such as stronger thunderstorms, and therefore a more convective activity.

3.2.9. Meteorological Anomalies of the Event's Precipitation Threshold

The meteorological anomalies of atmospheric pressure, temperature, and wind speed were analyzed per station, in order to better visualize their behavior 24 h before and after the coherent rain events. Given the Gaussian nature of these variables, determined by a previous Shapiro–Wilk normality test, each anomaly is obtained by subtracting the meteorological variable minus its mean and is then divided by the standard deviation shown in Table 5. Both

María Sheila Fabiola Serrano Vincenti

mean and standard deviation were calculated using the entire data set. Therefore, this anomaly is characterized by a dimensionless value in the interval [−3, 3]. These anomalies will complement the satellite images on the observed nature of rainfall, whether convective or stratiform, and its interactions with other local variables at each station.

The methodology and the steps of analysis applied to each station in order to obtain the relationship between rain and PWV are summarized in Figure 11. Note that atmospheric pressure was not available for Antisana (High Mountain station), therefore its derivation process is presented in Section 3.2.4.
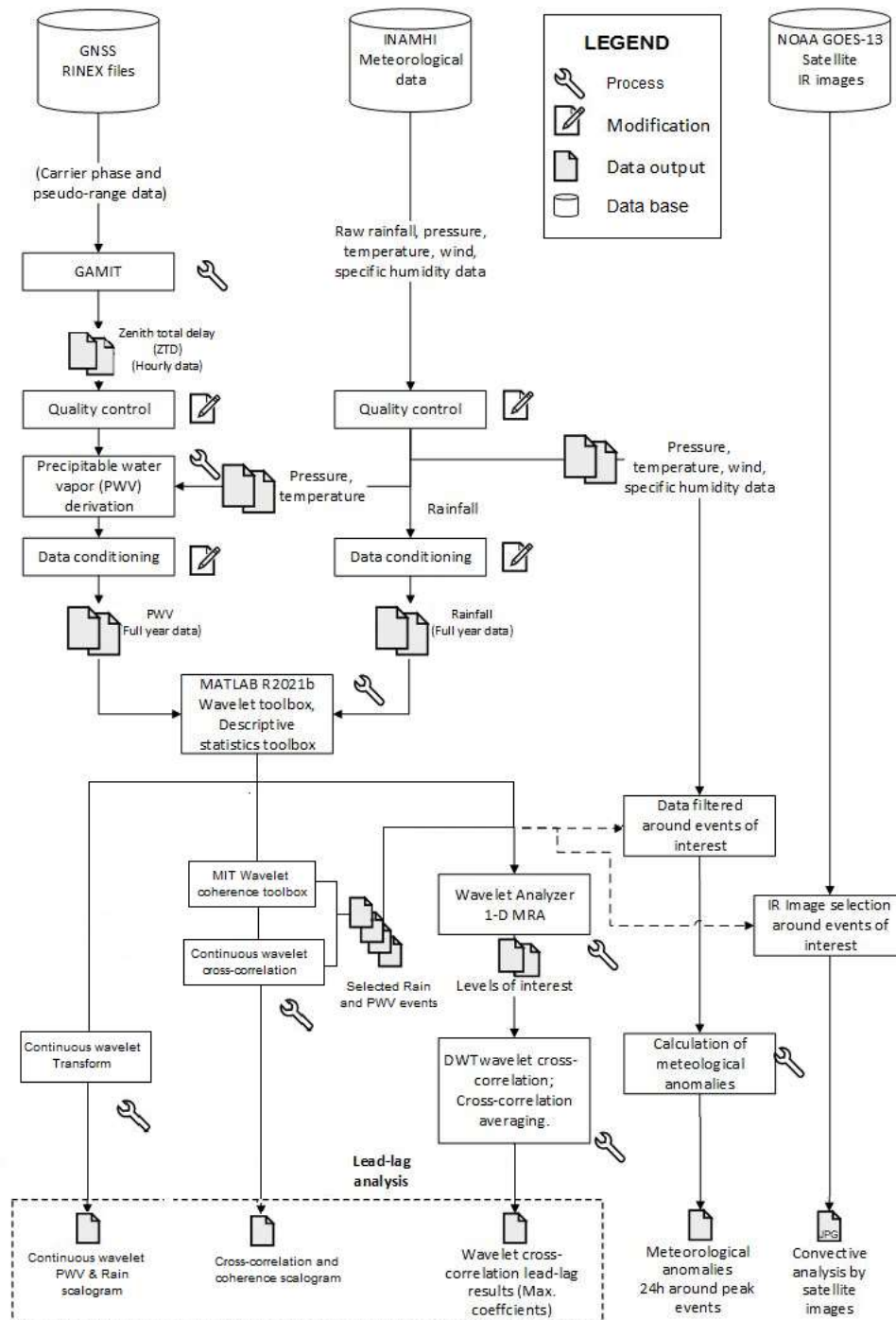
Figure 11 Data processing framework was applied to each of the five studied locations.

3.3. Results

3.3.1. Meteorological Description of the Stations' Location

The statistical analysis of every one of the available meteorological variables at each station for the year 2014, including the PWV and the missing data (NaN), is presented in Table 5. The stations are ordered from west to east, starting with the coast station (~45 m.a.s.l.), with a rain accumulation of 2640 mm/year, to the easternmost station located in the Amazon region (~553 m.a.s.l.), which is also the rainiest, with 3700 mm/year of accumulated precipitation. With increasing altitude, the annual accumulated rainfall decreases. The station with the highest altitude, located on Mount Antisana (4059 m.a.s.l.), records only precipitation of 415 mm/year.

In the same fashion, the variables of temperature and atmospheric pressure decrease with the altitude of the station, complying with Boyle's Gas Law (Houze, 2014). Regarding the behavior of the PWV, it can be seen that it decreases with altitude in most of the stations, with the exception of the high mountain station at Antisana which shows a precipitable water vapor as it is expressed in mm with a value of 21.52 ± 2.2 mm higher than the station at the Andes Valley in Quito, with 17.6 ± 1.7 mm. Despite the fact that the values of its rainfall are quite low; relative humidity, a magnitude which depends on both temperature and specific humidity, decreases at higher altitudes, except in high mountains where it increases due low temperatures. Wind registers the highest values in the high mountain station, with 4.56 ± 4.15 m/s, and a minimum in the Amazon station with 0.8 ± 0.6 m/s.

Figure 12 shows both the average hourly rainfall of the entire year 2014, including precipitation values ≥0, and the average hourly PWV per station. It can be seen that the diurnal cycle of these variables depends on the specific characteristics of the area, which produce rain with possible local origins. On the coast, land-sea breezes generated by the difference in heating between the sea and the coastal zone, which promotes convective rain at night (Mapes et al., 2003); steep-orography convective rain in the Andes valleys, which are usually present in the early morning and afternoon, and in the high mountain station, which favors rain during the day due to solar radiative forcing (Yepes et al., 2020), especially in the high mountain region (Ruiz-Hernández et al., 2021); and a probable stratiform-convective rain in the Amazon region, in which the recorded precipitation does not adjust to a particular time or schedule.
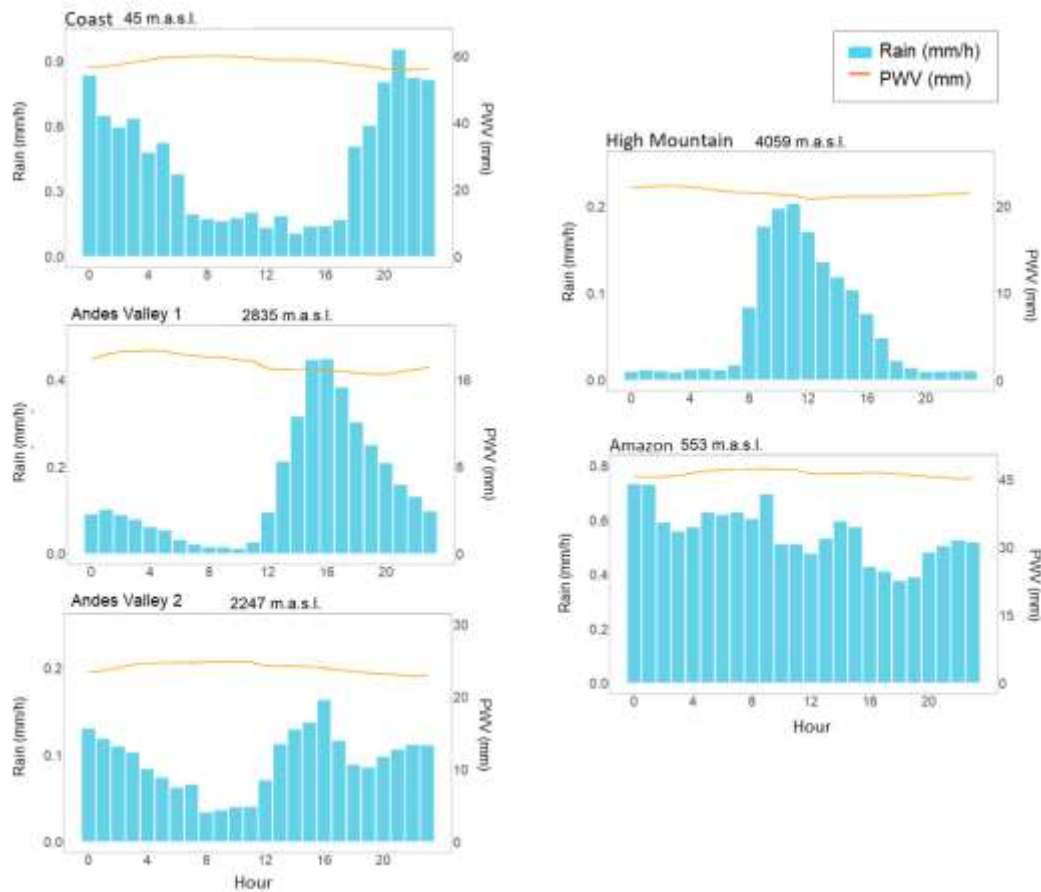
Figure 6 Diurnal cycle of rainfall and PWV. Hourly average histograms of the rainfall ≥0 (light blue) and average PWV (orange) for 2014.

Additionally, it can be noticed that the low variability of the hourly average PWV that present a slight increase during the hours when rain is less frequent.

3.3.2. Continuous and Discrete Wavelet Lead-Lag Analysis

The continuous wavelet analysis of the available rain and PWV data was performed for each station. Specifically, 12 months of data for the Andes Valley 1 and high mountain stations, and 6 months of data for the coast, Amazon region, and Andes Valley 2 stations.

As an example, Figure 13 shows a zoom-in of the wavelet plots of rain and PWV obtained from a bump-type analytic wavelet (Mathworks, 2021) for the Andes Valley 1 station for January 2014.
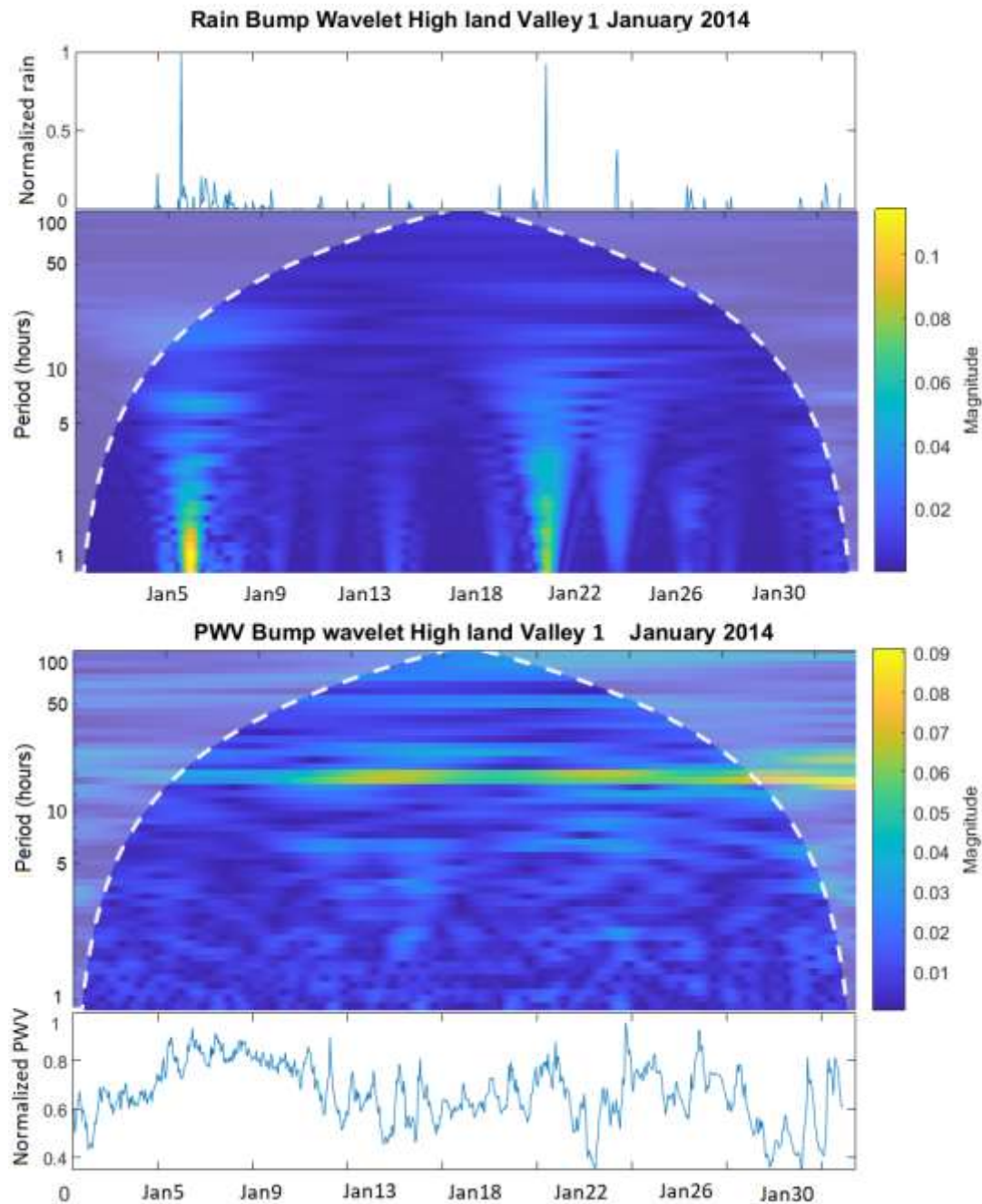
Figure 13 Zoom-in of the continuous wavelet plot of the Andes valley station 1 (Quito) for the rain (top) and PWV (bottom) for January 2014 with their respective time series.

It is evident how the presence of prominent rain events marks the magnitude of the wavelet plot, on which its lack of harmonic content can be seen except for the step-like frequency modes (periods) on 6 January and 22 January, whose intensity decreases as the period increases.

In contrast, when analyzing the CWT of the water vapor for the same time frame, a characteristic frequency of 24 h is present. In other words, throughout the time series, the

influence of the diurnal cycle over the PWV, marked by the radiative forcing of the sun, is evident.

The wavelet analysis showed the non-periodical behavior of the rain for all stations, with a few localized exceptions where rain occurrences are well concentrated in a handful of days and whose noticeable periods do not exceed 6 or 8 h as shown above. Conversely, the wavelet plot of the PWV contains some acceptable harmonic content, enough to locate a 24 h year-round period for all stations (Appendix A) and well-localized minor sub-periods during certain days or weeks as in the case of the Amazon station, (Appendix A). These periods were extracted by using the Fourier Transform in the five stations.

Although the PWV and rainfall have harmonics in common, they are not necessarily in phase. On the contrary, as can be seen in Figure 4 and will be shown in subsequent analyses, these variables are in contra-phase (when it is maximum, the other is minimum).

3.3.2.1. Cross-Spectrum Wavelet XWT and Wavelet Coherence WTC Results

As shown in Figure 13, it is hard for a simple continuous wavelet to transform to point out any possible link between the series of rain and PWV. Therefore, a continuous cross-spectrum calculation of the wavelets transforms or XWT is performed as well as their wavelet transform coherence WTC as explained in Equation (7) and Equation (9), in order to determine the time-frequency areas where both rain and PWV show a common energy activity and are covariant at the same time with a level of confidence of at least 95%.

Appendix B shows the XWT and WTC performed for each station for all data available in 2014. However, note that the study focuses on the plot areas where coherence and significant cross-spectrum shared the same time and frequency, so a zoom-in of such areas was necessary to determine the events of interest with a better resolution, as shown in Figure 14. The thick contour around the areas of maximum energy is referred to as the boundary of statistical significance of 5%.

María Sheila Fabiola Serrano Vincenti
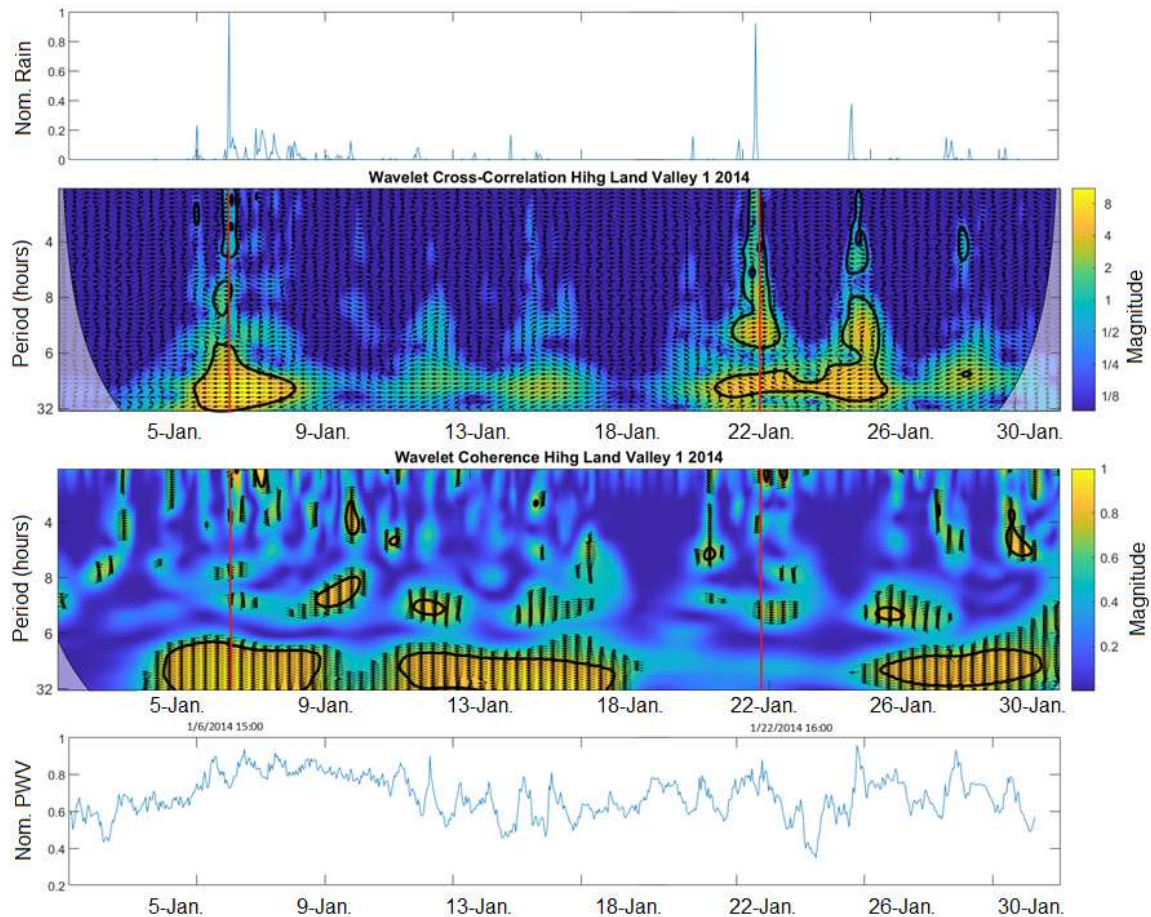
Figure 14 Zoom-in of the continuous wavelet cross-spectrum XWT and wavelet coherence scalogram WTC of the Andes Valley 1 (Quito) station for January 2014 compared with their normalized time series. Note the vertical red line which intersects the time domain for both plots and how there is coherence with a significance boundary of 5%, marked by the black line, for the 6 January event while there is none for the 22 January event.

It was found that areas with the largest power of XWT are due to the presence of important rain events. The larger the rain event, the stronger the power of the magnitude of XWT. This is because the XWT provides us with a measure of the power distribution of two signals. Hence, it is important to obtain the WTC coherence to determine which mutual events of rain and PWV have a true joint variability, as it was observed that there was no coherence with the PWV for all intense rainfall events, at least with a statistical confidence level greater than 95%.

The XWT and WTC are plotted to show periods of repetition (frequencies) of up to 32 h, with the aim of finding sub-diurnal relationships. For example, Figure 14 shows that the XWT has important periods well localized between 6 January and 9 January and also between 22

January and 26 January, being the more important ones, the periods ranging from the hours 16 to 32 in the first case, and periods centered around the hours 12 and 24 in the latter one. On the other hand, the WTC shows many areas where it can be said that the rain and PWV are covariant, the larger ones being the regions from 5 January to 9 January, 12 January to 18 January, and 26 January to 30 January, but did not consider the intense event of Jan. 22th. These regions comprise periods ranging mainly from hours 16 to 32.

The phase content of the rain and PWV can be determined from the direction of the arrows in the plot. It can be noted in Figure 6 that the only section where there is coherence and high energy activity, as matched with the above rainfall time series, corresponds to the intense event of 6 January. The arrows in this case are pointing to the left, indicating a complete out-of-phase scenario, where the PWV precedes the rain series in a half cycle as explained in Equation (8). The periods of peak energy are localized between hours 20 and 24 as seen from the XWT plot, thus the lag for the 6 January event can be calculated between hours 10 and 12, respectively.

It is interesting to note that approximately 12 h of lag will be common in all stations (Appendix B); this fact is also obtained from Figure 4, although less evidently, where high PWV values occur at the hourly level when the rain decreases, as if it were a continuous daily process where the PWV builds up, to later release itself previous to a rainfall event at particular hours during the day.

Table 4 shows the number of selected intense events with XWT and WTC with a confidence level above 95% between PWV and rainfall, among the largest selected independent rainfall events. Note that not all the intense events show XWT and WTC, i.e., not all the intense events were preceded by a PWV peak. In the same way, the rainfall threshold values are presented with their respective percentile, in addition to the maximum values and the median (50th percentile).

Additionally, Table 4 shows the climatic difference among the regions of study, not only for the quantity of accumulated rain found at each station of Table 3, but for their observed dissimilar maximum values and hourly intensity thresholds of rain (marked by the minimum rainfall value of the selected events of interest). It should be noted that all the stations register an identical minimum value of hourly accumulated precipitation equal to 0.1 mm/h, marked by the precision of the pluviometers. Note also the rapid decrease in the values of rainfall intensities since this value is close to the 50th percentile in all cases.

María Sheila Fabiola Serrano Vincenti

The coastal station is the only one where all the maximum accumulated rainfall intensity events show coherence and peak energy above the 95% confidence level boundary with the PWV, and therefore were preceded by a PWV peak. This fact may be due to the sea breeze phenomenon which is much more homogeneous and with less variability (Mapes et al., 2003) than the events that occur in the steep orography of the Andes or in the Amazon region, where various climatic factors produce intense precipitation (Segura et al., 2020). For other stations (Appendix B), these events are intercalated between the maximum values, although the largest event of the year always presents a correlation and a peak energy with the PWV.

3.3.2.2. Lead–Lag Discrete Wavelet and Convection Analysis for the Events of Interest

In order to define a more precise magnitude of the lag between the PWV and intense rain peaks, a discrete wavelet analysis was performed and a data set with as many samples as possible around the events of interest was constructed. Given the constraints that each event had to abide by, some event's data sets were as short as 9 days (the truncated ones that occurred when intense events occurred close to each other) and as long as 30 days. This was performed for each station to assure a more reliable lead–lag calculation given the hourly resolution of the samples.

Once the data sets were obtained, a wavelet cross-correlation D-WCC calculation by using a Daubechies wavelet of order 20, was implemented. This wavelet was selected due to its ability to be implemented with a large number of vanishing moments, which are useful when analyzing complex signals, as well as due to its widely available information and implementation in MRA(Daubechies, 1992).

The Wavelet Toolbox (Mathworks, 2021) was very useful to decompose each of the data sets of interest for rain and PWV. Up to five different levels of detail can be seen by selecting the db10 (Order 20 Daubechies wavelet). These levels of detail are the plot of the wavelet coefficients obtained by performing the discrete wavelet transform in dyadic scales.

The selection of the levels of detail of rain and PWV was the most taunting one, since the amplitude of these coefficients has a direct impact on the final wavelet cross-correlation calculation and therefore the lag. Nonetheless, a trade-off was necessary in some cases given the constraints of the algorithm used to calculate the wavelet cross-correlation, which required same-size datasets to be operated. That said, the results were obtained from the same level coefficients for the rain and PWV.

María Sheila Fabiola Serrano Vincenti

The most representative levels of detail, from two to four, were obtained from all data sets of rain and PWV related to each event of interest and cross-correlated with the purpose of calculating their correlation coefficients and corresponding lag. This treatment was performed for all events of interest and per station. Figure 7 shows the average D-WCC of the levels per station with their respective standard deviation. Note that the D-WCC yields a matrix of cross-correlation coefficients obtained from the different levels of decomposition of the discrete db10 wavelets of each series. For that reason, a Level (x, y, z) indicates the cross-correlation of all levels of decomposition, x for the rain and y for the PWV, from which the level z is selected for analysis or plotting.
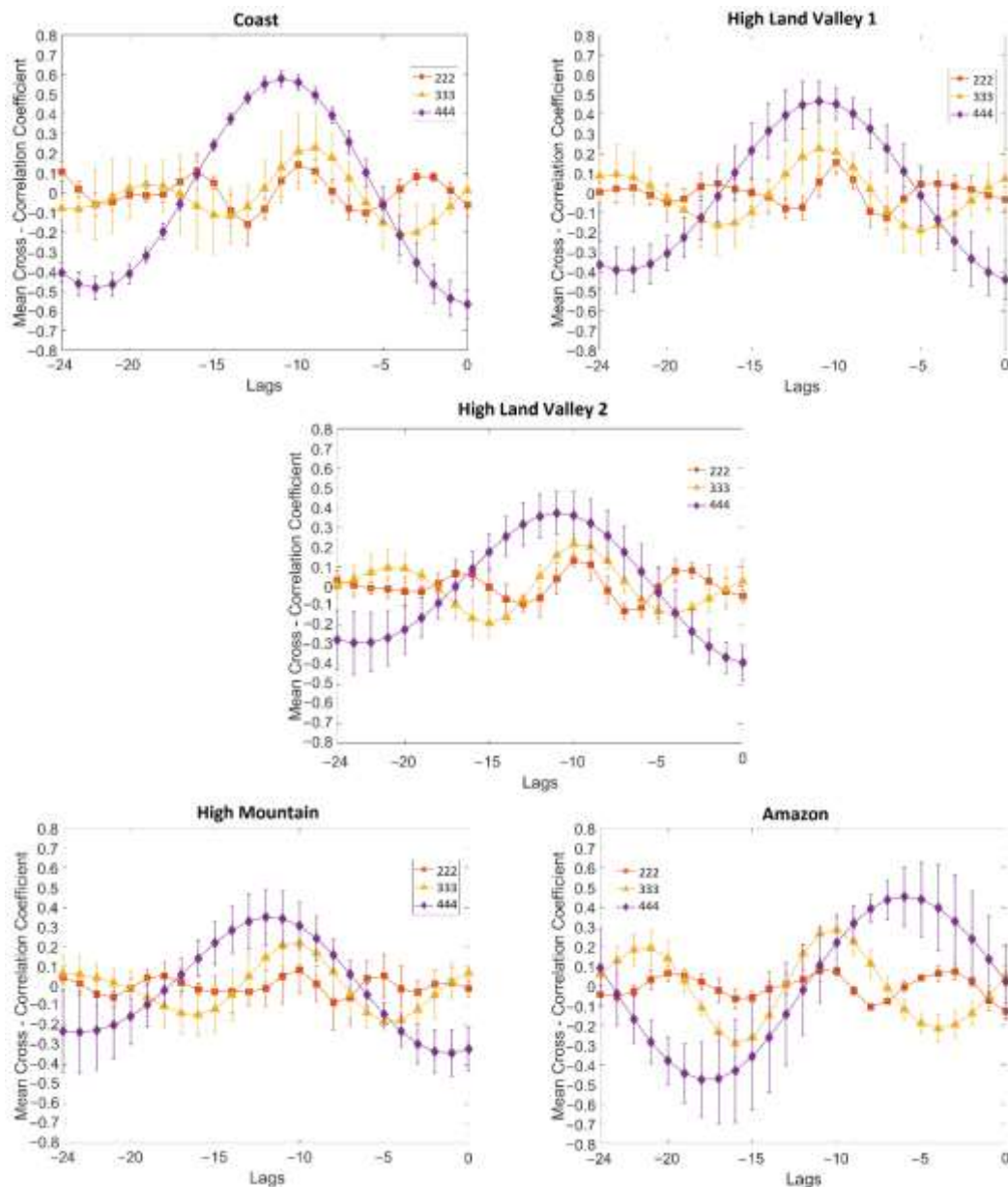
Figure 15 Average of the discrete wavelet correlation lag results registered per level of detail for the selected rainfall events presented in Table 5, including their standard deviation.

Table 7 Events of interest arranged in descending order of magnitude (first column) with the corresponding date and time, rain intensity, wavelet decomposition levels with the max. correlation, lag, control cross-correlation parameter of the time series, and the presence or not of convective clouds registered by satellite images as Y = Yes, N = No, Not Available = NA, Minimum = M. The data sets of events marked with * were truncated to prevent overlapping, for (a) Coast, (b) Andes Valley 1, (c) Andes Valley 2, (d) High mountain and (c) Amazon.

(a) Coast Station Selected Events

No.	Time	Rain intensity [mm/h]	MaxCorr	Level	Lag [hours]	Control SCC	Convective Rain
1°	10/29/2014 3:00	27.5	0.54	444	−11	−11	Y
2°	8/12/2014 19:00	22	0.6	444	−11	−10	Y
3°	8/21/2014 20:00 *	17.4	0.63	444	−11	−10	Y
4°	10/5/2014 0:00	16.9	0.56	444	−11	−10	Y
5°	8/5/2014 20:00 *	15.9	0.55	444	−11	−10	Y
6°	10/4/2014 23:00 *	12.9	0.56	444	−11	−11	N
Mean ± St.Dev					−11		

(b) Andes Valley 1 Selected Events

No.	Time	Rain intensity [mm/h]	MaxCorr	Level	Lag [hours]	Control SCC	Convective Rain
1°	1/6/2014 15:00	19.8	0.38	333	−11	−10	Y
4°	8/26/2014 17:00	14.6	0.41	333	−11	−10	NA
5°	11/17/2014 18:00	13.6	0.57	444	−12	−10	Y
7°	2/22/2014 20:00	11.5	0.57	444	−11	−10	N
8°	4/26/2014 20:00	10.9	0.57	444	−12	−10	M
9°	5/29/2014 20:00	10.9	0.46	444	−10	−10	NA
12°	4/19/2014 00:00 *	9.4	0.64	444	−11	−10	Y
13°	3/12/2014 20:00 *	9.1	0.46	333	−11	−12	Y

María Sheila Fabiola Serrano Vincenti

No.	Time	Rain intensity [mm/h]	MaxCorr	Level	Lag [hours]	Control SCC	Convective Rain
14°	2/19/2014 19:00 *	8.4	0.57	444	−11	−11	M
Mean ± St.Dev					11.1 ± 0.6		

(c) Andes Valley 2 Selected Events

No.	Time	Rain intensity [mm/h]	MaxCorr	Level	Lag [hours]	Control SCC	Convective Rain
1°	10/10/2014 19:00	16.8	0.32	444	−11	−11	Y
5°	10/28/2014 7h00	6.9	0.4	444	−9	−10	M
6°	11/9/2014 19:00	6.9	0.28	444	−10	−10	Y
9°	11/21/2014 16:00	4.7	0.409	444	−11	−13	Y
10°	10/8/2014 14:00 *	4.2	0.15	333	−12	−10	N
11°	10/19/2014 18:00	4.1	0.16	333	−11	−11	Y
13°	9/11/2014 23:00	3.6	0.53	444	−12	−10	N
Mean ± St.Dev					10.9 ± 1.1		

(d) High Mountain Station Selected Events

No.	Time	Rain intensity [mm/h]	MaxCorr	Level	Lag [hours]	Control SCC	Convective Rain
1°	5/10/2014 15:00	9.5	0.23	333	−9	−9	Y
3°	5/14/2014 11:00 *	5.99	0.12	333	−9	−10	N
5°	4/24/2014 12:00	4.1	0.5	444	−12	−12	Y
6°	3/18/2014 17:00	4.1	0.5	444	−13	−14	Y
7°	12/7/2014 15:00	4.9	0.51	444	−12	−12	Y
8°	12/28/2014 12:00	3.9	0.51	444	−12	−12	Y
Mean ± St.Dev					11.2 ± 1.7		

(e) Amazon Station Selected Events

No.	Time	Rain intensity [mm/h]	MaxCorr	Level	Lag [hours]	Control SCC	Convective Rain
1°	8/13/2014 12:00	49.1	0.56	444	−5	−6	N
2°	10/19/2014 22:00	48.6	0.412	333	−10	−10	Y
4°	7/30/2014 12:00	47.1	0.52	444	−8	−5	N
5°	8/16/2014 11:00 *	42.9	0.2	333	−3	−2	N
6°	8/4/2014 7:00 *	42.2	0.6	333	−10	−11	N
8°	8/9/2014 11:00 *	35.1	0.2	333	−3	−2	N
Mean ± St.Dev					6.5 ± 3.3		

María Sheila Fabiola Serrano Vincenti

Level 444 has been the one with the highest correlation for all stations in this study. This means that the db10 discrete wavelet transform of the rain and PWV performed down to four levels of decomposition, yielded the highest cross-correlation coefficients at the fourth level.

The average correlation at level 444 in the coast station was not only the highest, but also the least dispersed, showing the most homogeneous behavior of all five stations, and providing the highest average lag at 11 h. Levels 222 and 333 presented higher dispersion and a smaller correlation coefficient than level 444, with a lag of 10 h and 9 h, respectively.

Similarly, at the Andes Valley 1 station, the average lag was located very close to each other for the analyzed levels of decomposition. A lag of 11, 11, and 9 h were obtained for levels 444, 333, and 222, respectively. In the case of Andes Valley 2, a lag of 11, 10, and 10 h were obtained at the same levels. In the High Mountain station, a lag of 12, 10, and 10 h were obtained for the same levels. These results change when analyzing the Amazon station, where the largest lag corresponded to 6, 10, and 3 h for levels 444, 333, and 222, respectively.

Table 7 collects the individual analysis of the studied events, indicating the date and time of the event of interest, the magnitude of the recorded rainfall, as well as the level of decomposition that provided the greatest correlation. Note also that the value of the simple statistical correlation of the time series is included as a control parameter, to corroborate the value of the lag calculated with the discrete wavelet transform. The sign of the lag achieved confirms that the peak that precedes the rain is from the PWV. Finally, a convection analysis achieved from satellite images complements Table 7.

With respect to the coast station, it presents a well-localized lag at 11 h with most rain events occurring at night, which corresponds to the diurnal cycle presented in Figure 4. The major rain events are not only shared with the PWV coherence and peak energies, but also convection. It could be inferred that the sea breeze influences the dynamics between rainfall and PWV, not only because it allows uniformity in the lag but also because it could be decisive in the presence of convection on the coast.

In the case of the Andes Valley 1 station, the largest events are recorded in the afternoon and early morning, which corresponds again to its diurnal cycle. While in Andes Valley 2 station, the largest events happen in the afternoon, with exception of one event recorded at dawn. Additionally, the high mountain station presents its events only during the day. In these three Andean stations, coherence between PWV and rain does not always occur, as shown in Table 5, i.e., the PWV peak does not always precede intense rainfall. Furthermore, it can be said then that convection is present in almost all events at these stations. Therefore, the

relationship between the PWV peaks and rainfall is more complex than at the coast one station, and PWV is not always the only precursor for the generation of an intense rainfall event. Additionally, the orography of the Andes region plays an important role in stimulating this convective process.

The Amazon station has the most intense hourly accumulated rain events of the studied series, and in the Andean case not all intense rain events share coherence and peak energy with the PWV. Furthermore, in contrast with the rest of the stations, there is not a regular time during the day when rain occurs. The PWV periods found in the wavelet analysis are scattered and should be further analyzed. In the cases where there was coherence between the series, the selected events have shorter lags, which range from 5 to 10 h at most. The convection is minimal since the generation of high convective clouds is probably not required in order to achieve intense precipitation; it has already been discussed that the rain in this station may have a stratiform origin, produced by the advection of humidity masses transported through the Amazon jungle by the trade winds, differing as such from the other stations.

3.3.3. Analysis of Convective Clouds Using Satellite Images

The convection analysis presented in Table 7 has been obtained from the images in Figure 16. Note that convective events are represented by pink, blue, turquoise, yellow, and orange spots representing cloud top heights from −80 °C to −30 °C. The highest corresponds to tall vertical formations related to cumulonimbus clouds, and hence to convective rain. Only the Amazon station (Tena) shows one event of convective rain, and the rest of the occurrences correspond to low- or medium-height clouds. The analysis of the presence of high clouds tops and medium clouds, with lag evaluation, is presented in Table 7.
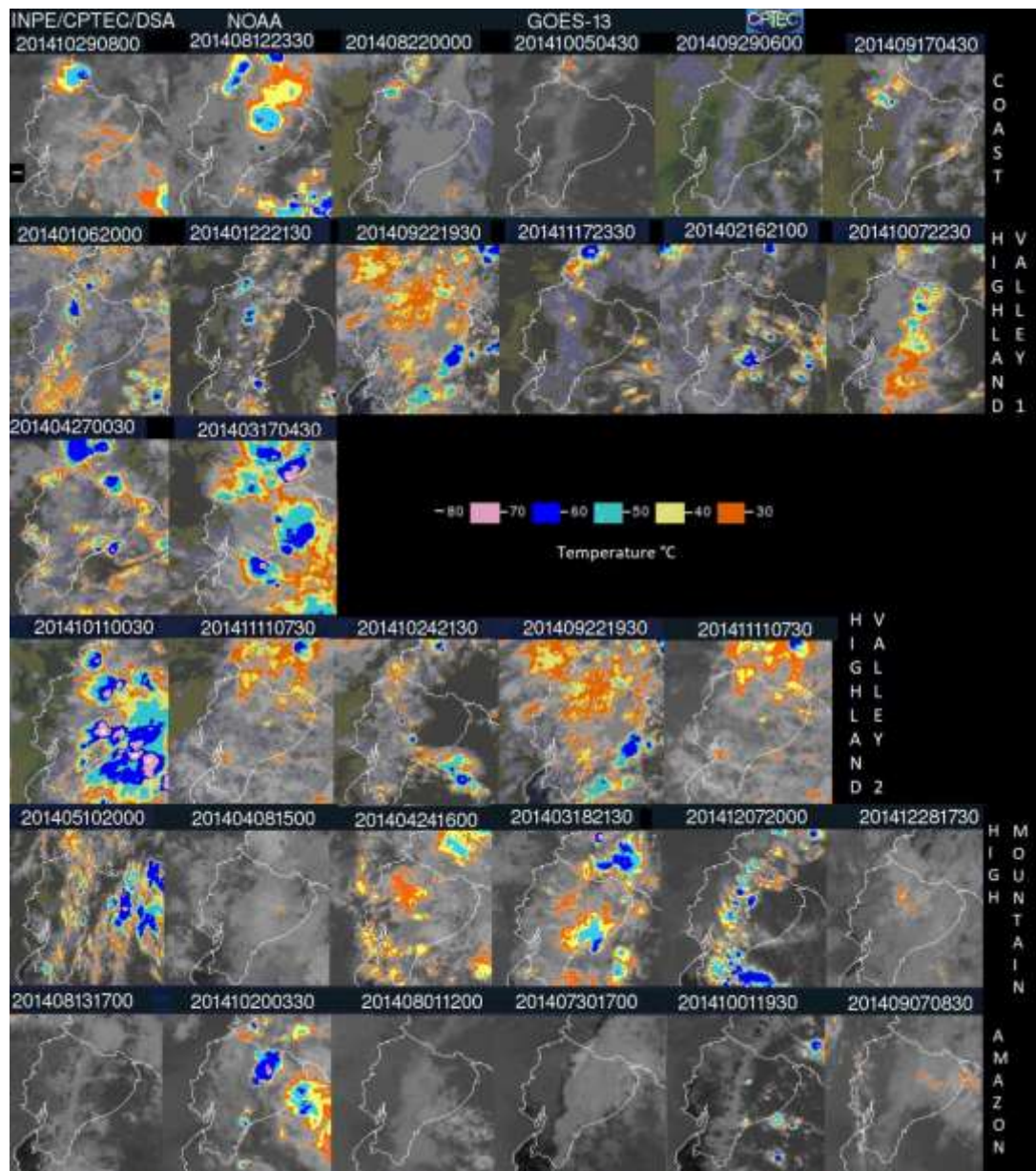
Figure 16. Satellite images corresponding to the GOES 13 in the infrared spectrum to determine the cloud top height temperature for all events per station. The timestamp of the occurrence of each image is displayed in GM.

3.3.4. Analysis of Meteorological Anomalies

Figure 17 shows the average of the meteorological anomalies, 24 h before and after the selected rain events, which were centered at 0 h. These anomalies were calculated as explained in Section 3.2.9, using the data presented in Table 5. The atmospheric pressure anomalies (Figure 17a), show different behavior between the stations of the Andes, high mountain, and coast vs. the behavior of the Amazon station. In the first ones, a decrease in

pressure is appreciated before the rainy event, which could indicate local convection caused either by the orography of the Andes, or due to the phenomenon of the sea breeze on the coast. However, in the Amazon station, this decrease in pressure is not appreciable. A phenomenon that may point to the stratiform nature of the rains in the Amazon region is caused by advective movements of moisture transport (Segura et al., 2020), in accordance with the preceding results.
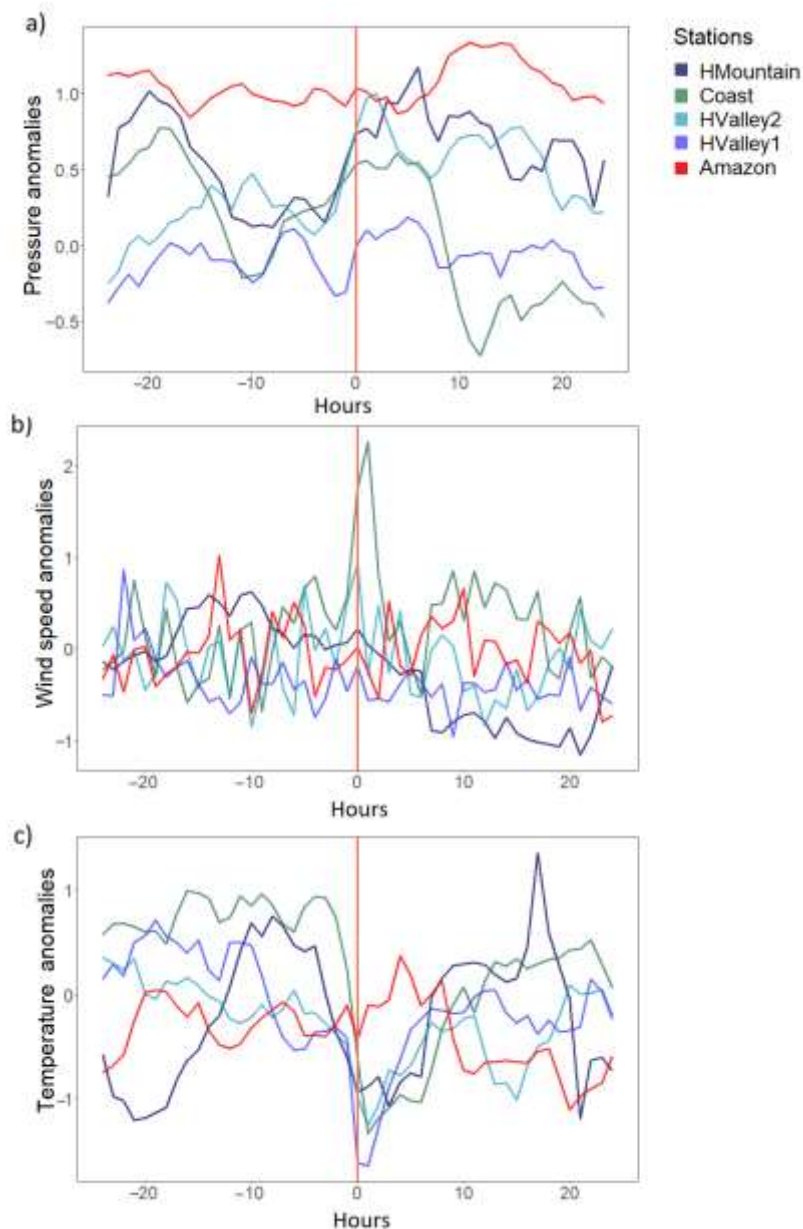
Figure 17 Average anomalies 24 h before and after each rainfall event of interest presented in Table 7: (a) atmospheric pressure (b) wind speed (c) temperature. The number of studied events per station are Coast (6), Andes Valley 1 (9), Andes Valley

In Figure 17b, a fairly regular behavior is observed with regard to wind speed, except at the coast station, where a marked increase in wind speed is seen in the presence of a rainfall event, possibly due to the effect of the sea breeze which could generate considerable winds (Mapes et al., 2003). Finally, when analyzing Figure 17c, it can be noted a generalized behavior is marked by a decrease in temperature as a result of the rain event, a phenomenon explained by the subsequent evaporation of precipitation and, therefore, absorption of latent heat from the surroundings (Mapes et al., 2003; Houze, 2014; Yepes et al., 2020; Ruiz-Hernández et al., 2021).

3.4. Discussion

Considering the results, it can be noted the difference between the relationship between the PWV and the intense rainfall for the different climatic zones of Ecuador. The most homogeneous relationship is the one found at the coast station, showing the greatest predictive potential of the PWV peaks before an intense rain event. In the Andes, these relationships are less common and despite the fact that they exist, they also obey other dynamics that must be considered for the generation of intense rain. Similarly, in the case of the Amazon region, correlations between PWV and intense rainfall events do exist, but they are due to different climatic factors than in the previous cases.

The PWV and rainfall series around the selected intense events for the coast station presented both correlation and coherence. The dispersion of the lags encountered for the PWV preceding the intense rain was minimal at the highest correlation coefficient located at 11 h. Additionally, convection was found in almost all those events. It can be said that the phenomenon of sea breeze, the main driver of precipitation events on the coast, is dynamically the most uniform aspect of the studied areas.

In the Andes stations, correlation and coherence were found in a scattered way in a number of events chosen. A total value of 64.3% of the sampled events showed correlation and coherence for the case of Andes Valley 1. This could be indicative of the presence of other dynamic factors such as the topography of the area, leaving the relationship between PWV and rainfall as an important driver of intense rain but not an essential one. The selected events of interest which presented coherence with the PWV generated a lag similar to the coast, being close to 11 h. This is also similar to Andes Valley 2, where 53% of the events presented

María Sheila Fabiola Serrano Vincenti

coherence and the most significant lag was at 11 h. The high mountain station, with 77% of events sharing coherence with the PWV, showed the most dispersed lag found at 12, 10, and 6 h.

The fact that the most common lag is found at 11 h for both the coast and Andes regions may have to do with the presence of more probable hours for precipitation (see Figure 6). This influence of the diurnal cycle favors the PWV recharge before the occurrence of an intense rainfall event, and even if it is not an intense rain, there are preferable hours for PWV to recharge in order to discharge in the rainfall most probable occurrence hours, in accordance with (Meza et al., 2020). The effect of the diurnal cycle is ubiquitous for PWV at the coast and in the Andes regions as shown in Appendix A and Figure 7, with characteristic periods of 12 h and 24 h reported (Torri et al., 2019). Furthermore, it should be noted from the cross-spectrum plot XWT that the PWV and rainfall have the same region of covariance centered at the 24 h period.

Unlike the other regions, the Amazon station shows different behavior for the PWV and rain. There are no preferred hours for precipitation (Figure 6) and the 24 h period is not unique, with other major periods appearing for PWV (Appendix B). Furthermore, the winds are minimal (Table 5), in addition to the fact that local convection is not necessary for the formation of intense rain. The main cause of rain in this region could be the advective transfer of moist air masses by trade winds (Campozano et al., 2016; Segura et al., 2020; Vargas et al., 2022) and its subsequent orographic forcing when touching the eastern mountain range of Ecuador (see Figure 9). This transport is responsible for both the generation of water vapor peaks and rain, but due to the greater frequency of rainy events, this lag between PWV peaks and rain is shorter: 6, 10, and 3 h for the wavelet levels analyzed, and not necessarily due to the diurnal cycle. These results are in agreement with other studies in the region (Adams et al., 2013b). Satellite images and the study of meteorological anomalies (Figure 17) show that there is no evidence of a significant decrease in atmospheric pressure before the rainfall event, which reinforces the fact that stratiform rainfall is preponderant at the Amazonian station.

3.5. Conclusions

This study shows the potential of harmonic analysis tools to characterize the relationship between the PWV and rainfall, with the aim of possibly anticipating the occurrence of an intense rainfall event in different climatic zones. The area that showed the most consistent relationships between these variables was the coast, where all rainfall events were preceded

by peaks in water vapor within 95% of confidence. This fact points to the importance of the PWV dynamics for the occurrence of intense rain events, in accordance with previous investigations carried out in other coastal areas such as Portugal (Benevides et al., 2015), Italia (Bonafoni & Biondi, 2016), and China (Yao et al., 2017; Zhao et al., 2020). Nevertheless, the reported lag is different, which may be due not only to the choice of short windows in the study of this lag, but also to other dynamic factors of each locality that must be considered.

In contrast, in the Andes, this relationship loses consistency, since only 53% to 77% of intense rainfall events were preceded by water vapor peaks. These types of studies in high-altitude areas have been little reported (Meza et al., 2020) and show the importance not only of the appearance of PWV peaks, but of other factors such as the topography in the occurrence of intense rainfall events. Likewise, in the Amazon region, where 66.7% of rain events preceded by PWV peaks were reported, the importance of other atmospheric phenomena is also appreciated, such as the advective transport of moisture carried by the trade winds from the Amazon. The reported lag was also found in previous studies (Adams et al., 2013b).

This study identified two types of lag, the first around 11 h, and found both on the coast and in the Andean areas, which had rainfall characterized by convectivity and determined rainfall occurrence times. Smaller lag (around 6 h) was recorded in the Amazon area, which coincides with previous research (Adams et al., 2013b; Sapucci et al., 2019). In this region, no preferred times for the occurrence of rain were identified, and most of the intense events were stratiform. It is possible that one of these factors or both influences the duration of the reported lag.

Therefore, the present study is an attempt to understand the relationships between PWV and rainfall in regions that have been identified for this purpose as relevant and with different climatology. Likewise, the integration of not only high-resolution regional models but also re-analyses of atmospheric circulation would be of great help to interpret other physical connections deemed important and necessary. Further progress would be the improvement of the tropospheric models through the adequate assimilation of GNSS-PWV data. All these factors will improve not only the better prediction of severe intense rain events, but also their comprehension in a complementary way to the traditional ones.

It is also important to highlight the contrast between the different harmonic analysis tools used in this study. Despite the featureless results provided by the spectrogram of a single variable wavelet analysis, when combined with statistics in a bivariate approach to obtain the cross-spectrum and coherence scalograms, the results are more significant. Time-frequency regions that would go otherwise unnoticed are now evident with a level of statistical significance including the phase content (lead or lag) between rain and PWV. The discrete wavelet analysis

using Daubechies 20 was performed to validate the results of the cross-spectrum scalogram and to determine a measurable quantity that could be compared to the traditional statistical correlation method used as a control parameter.

Further research with data sets containing longer periods of time, perhaps years if not decades, and sampled to better temporal resolutions, are necessary in order to increase the understanding of the relationship between rain and PWV, improve the harmonic content of the results, and accuracy of the lag in the aforementioned geographical areas. Although this is not always possible in stations located in remote or hard-to-access areas, refurbishing, maintaining, and expanding the current meteorological network will allow better data collection and data logging in the long term.

In Figure 9b, a fairly regular behavior is observed with regard to wind speed, except at the coast station, where a marked increase in wind speed is seen in the presence of a rainfall event, possibly due to the effect of the sea breeze which could generate considerable winds (Mapes et al., 2003). Finally, when analyzing Figure 9c, it can be noted a generalized behavior is marked by a decrease in temperature as a result of the rain event, a phenomenon explained by the subsequent evaporation of precipitation and, therefore, absorption of latent heat from the surroundings (Houze, 2014; Mapes et al., 2003; Ruiz-Hernández et al., 2021; Yepes et al., 2020).

Chapter 4

WRF modeled PWV and in-situ GNSS-PWV comparison, and their relationship with intense precipitation at high altitude stations situated in the Equatorial Tropical Andes

The precipitation modeling is still a challenge in mountain regions due complex terrain, the scarcity of instrumental measurements and difficulties of the models to reproduce all the atmospheric processes. The present research aims to evaluate the ability of the numerical weather research and forecasting model (WRF) to estimate the behavior of the precipitable water vapor (PWV) obtained from the global navigation satellite signal (GNSS) tropospheric delay and its influence on the estimation of precipitation. The study is carried out in the highland regions of the equatorial tropical Andes Valley (close to 2800 m.s.n.m) and High Mountain (close to 4000 m.s.n.m). The evaluation of the WRF model is carried out at the hourly level and with a horizontal resolution of 1x1 km during the year 2014 with in-situ hourly meteorological data of precipitation and in-situ PWV-GNSS. Both statistical techniques and harmonic analysis by the means of bi-variate wavelet and multiresolution analysis are used to convey the results. It is shown that the main discrepancy of WRF model is related to the non-reproducibility of the periods of charge and discharge of the atmosphere: PWV measurements show high values when precipitation probabilities are low with 12 hours of lag, while the model present a simultaneous behavior between PWV and precipitation increases with no delay.

Related publication:

4.1 Introduction

Disasters produced by extreme precipitation are still considered the deadliest global natural hazards (Pielke et al., 2013) in spite of the improvement in the understanding of precipitation phenomena and the use of many forecasting tools. Several authors have identified the fundamental limitations of atmospheric predictability as the main problem of Numerical Weather Prediction (NWP) models (Herman & Schumacher, 2016; Nielsen & Schumacher, 2016; Schumacher, 2016), especially when the precipitation event of interest occurs on small spatial and temporal scales, which is characteristic of extreme precipitation (intensities over 90th, 95th percentile) (Pendergrass, 2018; Harp & Horton, 2022).

 When small spatial features of the region are difficult to be well represented, as in mountainous areas of highly stepped zones, nonnegligible biases in the NWP simulations of precipitation were reported (Junquas et al., 2018; Ruiz-Hernández et al., 2021). Numerous factors could explain these biases, either at the synoptic scale with errors in the available moisture in the atmosphere (Precipitable Water Vapor; PWV) or in the physical processes that allow precipitation based on available moisture. Although bias corrections exist (Heredia et al., 2019), few studies explain on the causes of them.

NWP models has focused on the study of the PWV derived from Global Navigation Satellite System (GNSS), since it produces a significant enhancement of the forecasting of precipitation and its intense events when it is directly assimilated into the prediction model (Shoji et al., 2009; Risanto et al., 2021). In addition, the estimation of PWV has become much more accessible thanks to the methodologies that derive it from delays of the GNSS signals, which provide an accurate calculation in real time, under any climatic condition, and at relatively low cost, noting that the only condition is to have a sufficiently dense network of GNSS sensors adapted to the orography of the zone (Walpersdorf et al., 2007; Brenot et al., 2014).

It is clear that the study of the PWV is a key factor for the estimation of precipitation. Additionally, there is evidence of a new type of relationships between PWV and precipitation that could not only help the correct modeling of precipitation, but also to the understanding of new atmospheric processes that could trigger it. For example, it can be mentioned the presence of periods of loading and unloading of the atmospheric PWV (Torri et al., 2019; Meza et al., 2020), and which could be subject to the precipitation patterns of the study area (Serrano -Vincenti et al., 2022). Likewise, the appearance of increases in PWV before the occurrence of heavy precipitation events have been reported in various areas of the planet such as

Portugal (Benevides et al., 2015), China (H. Li et al., 2020, 2021; L Li et al., 2022), the Amazon (Adams et al., 2013; Sapucci et al., 2019), and various areas of the equatorial Andes region, although not always the occurrence of a PWV peak necessarily triggers a heavy precipitation event (Serrano-Vincenti et al., 2022); these physical relationships are recent and could be not yet assimilated into the NWP models.

Under this context, the objective of this work is to find through the comparison of the data between simulations of the Weather Research and Forecasting (WRF) in non-forecasting climate mode. WRF resolution was 1km and 1h spatial and temporal resolution respectively. WRF model was compared with in-situ observations of PWV-GNSS and observed precipitation, and the physical interaction of them. Possible errors in the PWV modeling could be pinpointed in order to improve the estimation of precipitation and intense precipitation. The data available for this analysis has been sampled in an hourly scale in the year 2014 from locations with great orographic complexity such as the High Mountain station (Antisana) and Andean Tropical Valley (Quito). The evaluation of the WRF model performance will be carried out through wavelet harmonic analysis tools and statistics at the seasonal and diurnal dynamic cycle levels. Also, the harmonic analysis of observed and modeled hourly intense precipitation events was performed.

4.2 Data, area of study and methodology

4.2.1 GNSS network data of tropospheric delays and in-situ meteorological data

The PWV was derived from zenith tropospheric delay (ZTD) data of the GNSS stations maintained by the Geophysics Institute of the National Polytechnic School of Ecuador (Instituto Geofísico de la Escuela Politécnica Nacional https://www.igepn.edu.ec/red-nacional-de-geodesia), which export this information to the database of UNAVCO https://www.unavco.org/data/gps-gnss/gps-gnss.html, (accessed on 21 March 2019). Two GNSS stations were chosen from this database due to their complex terrain, their proximity to the meteorological stations and due to the fact that a WRF model was previously validated for this area of study (Heredia et al., 2019).

The meteorological data for this study were obtained from Quito Atmospheric Monitoring Metropolitan Network (Red Metropolitana de Monitoreo Atmosférico de Quito, REMMAQ) and Institute of Meteorology and Hydrology of Ecuador (Instituto Nacional de Meteorología e Hidrología—INAHMI) INAMHI in collaboration with French Institute of Research for Development (Institut de recherche pour le développement, IRD) networks, for the Quito and Antisana stations respectively.

Figure 18 and Table 8 show the location and data used for this study. The Antisana HMA meteorological station (ORE) is located 4.75 km from the ASEC GPS station, while the meteorological station in Quito HV1 (BELI) is located 6.42 Km from the GPS station QUEM.
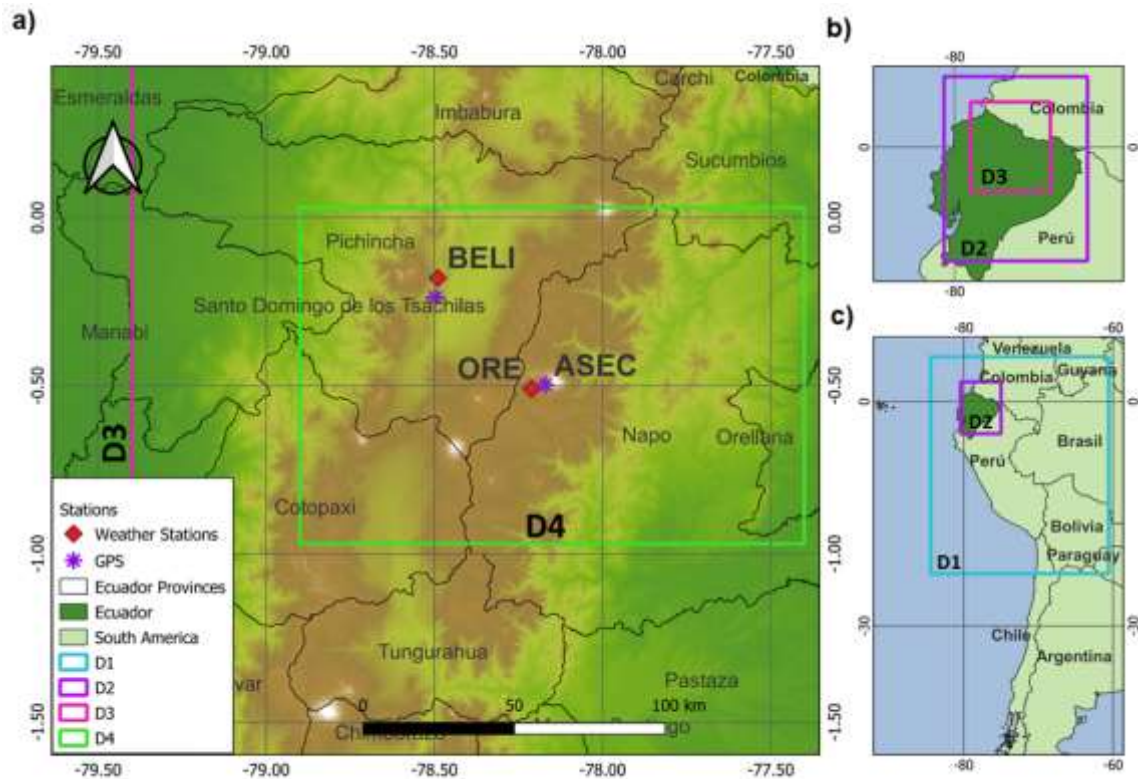
Figure 18 a) Location of the studied area model domain D4 of 1km resolution, white areas represent glaciers, and the numbered color rectangles in a), b) and c) indicate the model domains represented by D4 (green), D3 (magenta), D2 (purple), and D1 (light blue); b) country level domains D2 and D3 of 9 and 3 km resolution respectively, and c) continental level including D1 and D2 whit 27 and 9 km resolution respectively.

Table 8 Weather and GPS in-situ stations, and WRF model description.

Data source	Data type	LON (°)	LAT (°)	Altitude m.a.s.l	Variables	Time span
Antisana (ORE)	Meterological in-situ station	-78.2112	-0.5092	4059	Precipitation [mm/h] Temperature [°C]	30 min (2005-2018)
Quito (BELI)	Meterological in-situ station	-78.49	-0.18	2835	Precipitation [mm/h] Temperature [°C]	Hourly (2013-2020)

					Pressure [hPa]	
ASEC QUEM	GPS in-situ Station	-78.1704 -78.4973	-0.4973 -0.237	4571.1 3061.06	Derived Tropospheric Water Vapor PWV [mm]	Hourly 2014
WRF	Regional Model	1km grid points	1km grid points	Gridded	Precipitation [mm/h] Precipitable Water [mm]	Hourly (2014-2015)

4.2.2 WRF modeled data

The simulations used in this study were previously used and validated in Heredia et al. (2018). The model is WRF version 3.7.1 (Skamarock et al., 2008) and the period is 2014-2015, but only the outputs of the year 2014 were used in this study. The WRF model is non-hydrostatic and uses 30 vertical sigma coordinates following the terrain. Nested domains of 27, 9, 3 and 1 km were run with a first domain forced by NCEP-FNL (1°x1°). Only the outputs of the fourth domain at 1km were used in this study. The temporal resolution of this domain is 1h. The parameterizations were: Yonsei University scheme (Hong et al. 2006) as the planetary boundary layer option, with a wind topographic correction for the complex surface terrain (Jiménez and Dudhia 2012), that has already been used in previous studies using the WRF Model in the Andes (Heredia et al., 2018; Mourre et al., 2016). The microphysical parameterization was obtained from Lin et al. (1983), and the cumulus scheme from Grell and Dévényi (2002).

4.2.3 Quality control of in situ-data of ZTD and precipitation

Complete hourly ZTD series for the entire year 2014 were generated after the elimination of the repeated values at session boundaries at 06:00 and 18:00 UTC, which account for 1% to 2% of the total data. Then, the quality control carried out for these meteorological data series followed the standard for hourly data treatment (Climate World & Programme, 1986): First to the bias error, which creates discontinuities due to sensors' deviation, by adding a constant

value to the discontinuous period so that the moving averages are maintained, and next to the outliers, which were eliminated outside the interval [μ ± 4σ] (Lanzante, 1996). Only in the case of precipitation, outliers greater than [μ + 10σ] (Lewis et al., 2018)were eliminated since being characterized by non-normal distributions (Cho et al., 2004; Lanzante, 1996; Serrano-Vincenti et al., 2020) could present higher precipitation intensities. Once the outliers were eliminated, no padding was performed in these meteorological data.

It should be noted that the surface atmospheric pressure in the case of the Antisana station HMA, had to be estimated from the time series available at the Quito HV1 station minus an offset of 156.1 hPa, which corresponded to the height difference pressure average (Garratt, 1992). This was due to the absence of barometric pressure parameter at the in-situ station.

4.2.4 Estimation of the PWV from ZTD

First of all, the ZTD information has to be obtained from GNSS data which requires the computation of the GPS phase signals with several models with the purpose of finding the location of the ground base stations with great accuracy. These models as explained in (Herring et al., 2018) take into consideration the deformation of the earth's crust due to the ocean tides, also known as ocean loading, atmospheric loading, and the influence of solar and lunar movements on Earth's atmosphere (atmospheric tides). All these models used for obtaining the hourly ZTD were processed with GAMIT/GLOBK R10.7, which parameterizes the ZTD as a stochastic variation of the Saastamoinen model with a stepwise linear interpolation, at the Institute of Earth's Sciences, ISTerre, of the University Grenoble Alpes, using the ISTerre computing cluster IST-OAR.

The ZTD has two components: a Zenith Hydrostatic Delay ZHD due to a hydrostatic or total troposphere, which generally explains 90% of ZTD (Fernández L.I., Meza A.M., 2009), and a wet component or Zenith Wet Delay ZWD due to the tropospheric water vapor, which is a strong variable and related to the PWV that explains the remaining 10% of the total delay. Equation (13) (Bevis et al., 1992, 1994).

$$ZTD = ZHD + ZWD \quad (13)$$

The ZHD in [m] is obtained from the surface atmospheric pressure in [hPa] at the station's position in latitude and ellipsoidal height, while the ZWD has a direct relationship with the PWV as noted in(Bevis et al., 1994) as shown in Equation 14. However, in similar climatology

---

María Sheila Fabiola Serrano Vincenti

studies, a $\kappa_{E\&D}$ constant is used as presented by (Emardson & Derks, 1992), which only depends on the surface temperature Ts and has a precision of 1.15% Equation (14):

$$PWV_{E\&D} = \kappa_{E\&D} \cdot ZWD \quad (14)$$

Therefore, the PWV can be derived from these two equations at each location of interest by estimating the GNSS ZTD and the use of ground meteorological observations like pressure and surface temperature as detailed in (Serrano-Vincenti et al., 2022).

4.2.5 Wavelet based bi-variate analysis of PWV and Precipitation Data: observed versus modeled

4.2.5.1 The bi-variate continuous wavelet transform

This study doesn't use the wavelet transform in the strict sense but an extension of its application for a bi-variate case as explained in (Rösch & Schmidbauer, 2018), where the Continuous Wavelet Transform CWT is applied to two time series in order to find their common frequencies, their location in time, and the common spectral power in two-dimensional graphs known as wavelet cross- spectrum. The geometric interpretation of this operation is analogous to the calculation of the covariance, as shown in Equation 15:

$$W_{xy}(u,s) = S\left(W_x^{\star}(u,s).W_y(u,s) \right) \quad (15)$$

This bi-variate covariance is a function of the CWT of each single variable "*x*" and "*y*", which are noted as $W_x(u,s)$ and $W_y(u,s)$ where u and s are the time-location and frequency scale parameters, respectively. The asterisk represents the complex conjugate of the wavelet transform of the first series and *S* is a smoothing operator in time and frequency (Mathworks, 2021), thus obtaining the typical definition of the covariance of two complex random variables (Park, 2018) since the operated wavelets are analytical.

The time series to be analyzed with this wavelet bi-variate approach are the PWV and precipitation, both in the observed (in-situ) case and in the modeled case. Additional, PWV and precipitation, both the observed and in the modeled case, because provides a set of tools which are able to pin-point features of interest between these two-time series even if they appear to lack harmonic content (Mallat, 1999). Additionally, Fourier transform was applied to

PWV and precipitation, in the observed and modeled case in order to determine the characteristics periods of each series and its coincidences (Yang et al., 2006).

4.2.5.2 The cross-wavelet power spectrum

From Equation 15 the cross-wavelet power spectrum XWT is defined as:

$$\text{XWT}(u,s) = \left| W_{xy}(u,s) \right| \quad (16)$$

and the instantaneous relative phase angle between x_t and y_t, in the time-frequency space, is determined by calculating the complex argument of the bivariate transform. See (Grinsted et al., 2004).

$$\theta(u,s) = \arg\left(W_{xy}(u,s) \right) \quad (17)$$

This angle can be seen as a lead of lag between the variables in the time-frequency plane which in the case of our study will be the existing lag between precipitation and PWV.

Hence, from Equation (4) and (5), the XWT unveils common high energy regions between the operated series, in addition to providing valuable information on the relationship of phase between these two.

4.2.5.2.3 Wavelet coherence

The XWT, as traditionally seen from a statistical cross-correlation approach, provides the highest concentration of energy or peaks where the two variables are delayed one to another. However, is insufficient to analyze the correlations between a pair of processes in the time-frequency space, especially when the series has flat characteristics due to data loss or padding. This is the case of the PWV where some data has been left unpadded versus the pronounced peaks of precipitation.

Wavelet transform coherence (WTC) in the other hand, reveals the local correlation of the two series in time and frequency throughout their entire extension, with a localized phase behavior even in time and frequency segments that are not necessarily those detected by the XWT, but that could render useful information on the dynamics of these meteorological variables (Serrano-Vincenti et al., 2022).

The WTC is derived by normalization of the XWT, also known as the magnitude square coherence, as shown in Equation (18):

$$WTC_{xy}(u,s) = \frac{\left|S\left(W_x^\star(u,s) \cdot W_y(u,s)\right)\right|^2}{S(|Wx(u,s)|^2) \cdot S(|Wy(u,s)|^2)} \quad (18)$$

The smoothing operator "S" is also present as explained in (Torrence & Compo, 1998) and (Cohen, 2011).

4.2.5.4 Data conditioning, processing and interpretation of the bi-variate analysis

The missing data of the derived PWV was padded using a linear interpolation for all the empty data points since the variability of the series was almost smooth and regular in nature rather than irregular and with harsh discontinuities as in the case of the precipitation series for which a zero padding approach was used for the missing data. No erroneous data or outliers however were identified in any of the modeled WRF series.

Then, these conditioned series of the derived PWV and observed precipitation, were normalized using a simple min-max method, dividing the data points of the series by the maximum recorded value in order to preserve their original envelope but at the same time allowing to be mutually processed under the same energy conditions.

The analysis was performed using the implemented libraries for XWT and WTC from the Wavelet Coherence toolbox by Grinsted (Grinsted, 2014) run in (Matlab R 2021b), since they also provide regions of statistical significance which come handy for detecting areas of particular covariance between the analyzed series which are not the result of random variability. Also, these libraries apply the Morlet wavelet for the calculation of the XWT, providing a good balance location both in time and in frequency.

The XWT and WTC plot 2-dimensional scalograms with phase information, as explained in Equation (17), using arrows to determine the lead-lag angle in the interval of [−π, π]. An absolute value less than π/2 indicates that the two series tend to move in phase, having an ideal phase when the arrows point to the right horizontally. If the series tend to lag one another, the absolute value of the angle must be greater than π/2. If both signals oppose each other or are in a complete out-of-phase, these arrows are horizontal but left-oriented (Percival & Walden, 2000).

María Sheila Fabiola Serrano Vincenti

4.2.6 Lead-lag analysis of the intense precipitation

4.2.6.1 On the selection of intense precipitation events

As the study requires the analysis of the relationship between an intense precipitation occurrence and its corresponding PWV jump, both, observed and modeled intense precipitation where analyzed. Using the same criteria found in (Serrano-Vincenti et al., 2022), 20 events of interest were chosen per station which met the following conditions: first, they have to correspond to the largest accumulated hourly precipitation data points and arranged in an ordinal way. Second, PWV and precipitation shared both, correlation and coherence with the largest joint distribution energy over 95% confidence, and third that they must be kept mutually exclusive or not a part of the same precipitation occurrence; in other words, that were at least 24 hours apart in order to prevent the overlapping of the samples surrounding the peak event during processing.

4.2.6.2 Lead-lag multi resolution analysis (MRA) of the events of interest

A discrete wavelet transform using Daubechies 20 wavelet via the Multi-Resolution Analysis (MRA) method was applied to the regions where the events of interest lied, so to obtain a quantifiable result of the lag of the two series with hourly resolution.

It can be argued that the continuous bivariate analysis of the XWT and WCT is not meant to provide a quantity in particular, but the dynamic behavior of two signals for a particular time-frequency range or, better said, a region of significance (Serrano-Vincenti et al., 2022). Therefore, the MRA method of decomposing a signal into sub-components or frequencies comes useful as it allows faster computation and a wide choice of wavelets to choose from (Meyer, 1995).

The MRA decomposition of the precipitation and PWV series will produce different detail levels or one-dimensional scales (frequencies) for each one of them. Detail levels were tested from 1 to 5 for x serie, y and for the conjunction xy, in the form: 111, 222, 333, 444, 555; very low and high levels were removed due the creation of mathematical artifacts.

4.2.7 Statistical analysis of the model behavior for seasonal and diurnal cycles

Complementary to the harmonic analysis performed to the series, the statistical analysis of the WRF models of precipitation and PWV for the same year 2014 both seasonal (monthly) and hourly (diurnal cycle) was performed in order to verify the reproducibility of the dynamic behavior of the atmosphere in the model; this included frequency histograms, Q-Q plots and Taylor diagrams.

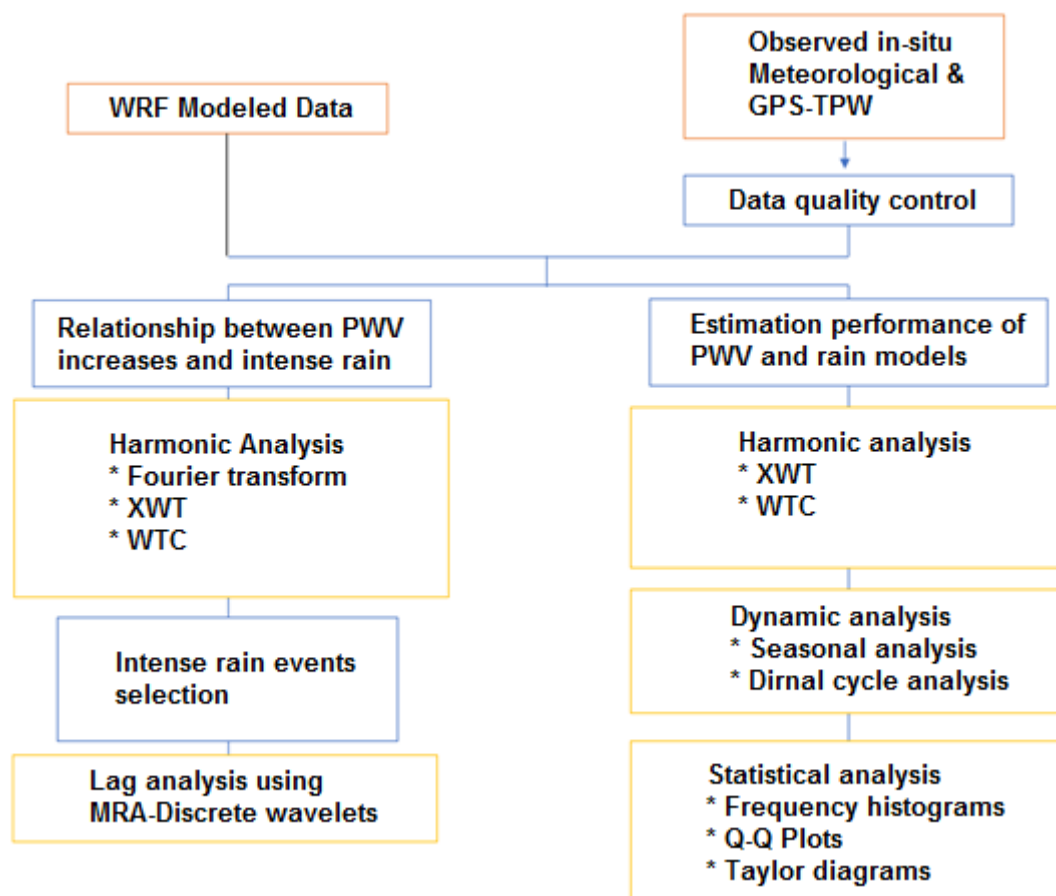The methodology applied to this study is presented in Figure 19.

Figure 19 Sequence of the applied methodology for this study

4.3 Results

4.3.1 Statistical indicators

The main statistical indicators of the precipitation and PWV series are presented in Table 9. It is important to note that the selected stations are located in the Andes, however, Quito station (HV1) is situated in a valley at lower altitude than the high-mountain station of Antisana (HMA), which shows less accumulated precipitation per year but shows the higher amount of observed PWV. As a reference, Table 10 shows the missing data of the series per month.

Table 9 Principal statistical indicators of the studied stations for observed and modeled PWV and precipitation, in 2014 year.

		HV1 2835 m.a.s.l			HMA 4059 m.a.s.l			
	PRECIP Obs. [mm.h^{-1}]	WRF-PRECIP [mm.h^{-1}]	**PWV Obs. [mm]**	WRF-PWV [mm]	**PRECIP Obs. [mm.h^{-1}]**	WRF-PRECIP [mm.h^{-1}]	**PWV Obs. [mm]**	WRF-PWV [mm]
% EMPTY VALUES	**9,45%**	0%	**9,48%**	0%	**6,19%**	0%	**9,75%**	0%
MEAN	**0,18**	0,054	**17,57**	16,64	**0,68**	0,074	**21,52**	9,24
MAX	**19,80**	13,88	**25,64**	25,59	**9,45**	18,53	**31,41**	14,96
MIN	**0,1**	1.93E-05	**7,30**	4,22	**0.101**	1.78E-05	**14,84**	1,86
SD	**1,23**	0,84	**3,08**	3,71	**0,64**	0,13	**2,20**	2,48
Q1	**0,1**	0,005	**15,36**	14,37	**0,21**	0,47	**19,91**	7,62
Q2	**0,4**	0,19	**17,85**	16,84	**0,52**	0,65	**21,70**	9,41
Q3	**1,3**	0,1	**19,93**	19,48	**1,12**	1,16	**23,17**	11,02
Acumm. Precipitation	**1224.5**		**-**	-	**415.7**		**-**	-

Table 10 Percent of the missing data per month of the observed in-situ precipitation and PWV. The highest missing values correspond to the months of April and May.

	HV1		HMA	
	PRECIP	**PWV**	**PRECIP**	**PWV**
JAN	1,77%	1,77%	0,00%	1,77%
FEB	10,71%	10,71%	0,00%	10,71%
MAR	2,69%	2,69%	1,48%	2,96%
APR	**13,19%**	**13,19%**	**13,06%**	**13,19%**
MAY	**36,29%**	**36,29%**	**31,45%**	**36,29%**
JUN	6,11%	6,11%	1,53%	6,11%
JUL	0,27%	0,27%	2,96%	3,23%
AUG	6,72%	6,72%	6,18%	6,72%
SEP	9,17%	9,17%	3,06%	9,17%
OCT	6,59%	6,59%	6,05%	6,59%

María Sheila Fabiola Serrano Vincenti

NOV	8,47%	8,75%	3,19%	8,75%
DEC	**11,48%**	**11,48%**	**4,51%**	**11,48%**

4.3.2 Harmonic relationship between modeled and observed PWV and intense precipitation

The first step in the harmonic analysis of the precipitation and PWV series was to perform the Fourier transform of the observed and modeled data in 2014 year, in order to find common characteristic periods or frequency modes. The most notable periods are shown in Table 11, and Appendix I and J, for both the observed series and the modeled ones. Were found periods around 24 and 12 hours, which show that the model correctly matched the periodicities of the observed series. In the case of Quito, one more period appears around 8 hours (7.9 hours) for the observed precipitation and 8 for the modeled precipitation. This extra period corresponds to an extra cycle not very evident apart of the described for precipitation in HV1 explained in Section 4.3.3. However, its presence is weak, since its spectral density is significatevly low (marked with *), which indicates that this period does not occur frequently, but it is remarkable that the model can detect it, also with low spectral density.

Table 11 Common characteristic periods obtained by Fourier analysis for the observed and modeled series of precipitation and PWV for 2014 year (* indicates a low spectral density).

	Observed Rain		Modeled Rain		PWV-GNSS		Modeled PWV	
	Periods [hours]	Spectral density	Periods [hours]	Spectral density	Obs. PWV-GNSS	Spectral density	WRF PWV	Spectral density
HV1	24 12	0.13 0.07	23.8 12	0.39 0.13	24 12*	2.2 0.38	23.8 11.7*	0.38 0.9
HMA	22.5 13.9*	0.045 0.01	24 12.2*	0.02 0.01	24 12*	0.42 0.16	24 12*	1.16 0.18

Next, a bi-variate wavelet analysis is carried out to determine the relationship between the increases in water vapor and heavy precipitation, based on both the XWT (cross-spectrum) and WTC (coherence). This analysis was performed over the entire series for the year 2014, which is presented in Annex 1-2. Although for better visualization, the three rainiest months of March, April and May, were chosen for the next figures.

Figure 20a shows the cross-spectrum of the observed precipitation for the HV1 station whose more visible energy regions depend on the intensity of the precipitation events, while the

coherence can be spread over smaller precipitation events as appreciated in Figure 20b. Figúre 20a and Figure 20b show a characteristic period around 24 hours with a lag of the precipitation signal with respect to that of water vapor, in a half period, as indicated by the left pointed arrows (Ec. 17). That said, both signals are practically in anti-phase; in other words, when the precipitation reaches its greatest value, the vapor is at its minimum.

By looking at Figures 20c and 20d, of the modeled series, the same period of 24 hours is present, however, the principal difference with the observed in-situ series is that the precipitation and PWV are in phase, indicated by the right direction of the arrows. In other words, the increments in the modeled PWV correspond to the modeled precipitation increments. As in the observed case, the cross-spectrum of the modeled precipitation follows the most intense precipitation events.

The situation is similar for the HMA station: there is a marked but less visible characteristic period around the 24 hour and the arrows in the observed in-situ series (Fig. 11 a, b) are mostly left oriented or in anti-phase. In the modeled series, the arrows are oriented mainly to the right (Fig. 11 c, d) showing that precipitation and PWV are in phase.

Figure 20 Zoom-in of the continuous wavelet cross-spectrum XWT and wavelet coherence scalogram of Precipitation and PWV in HV1. (a) XWT observed, (b) observed coherence, (d) XWT modeled (d) modeled coherence.
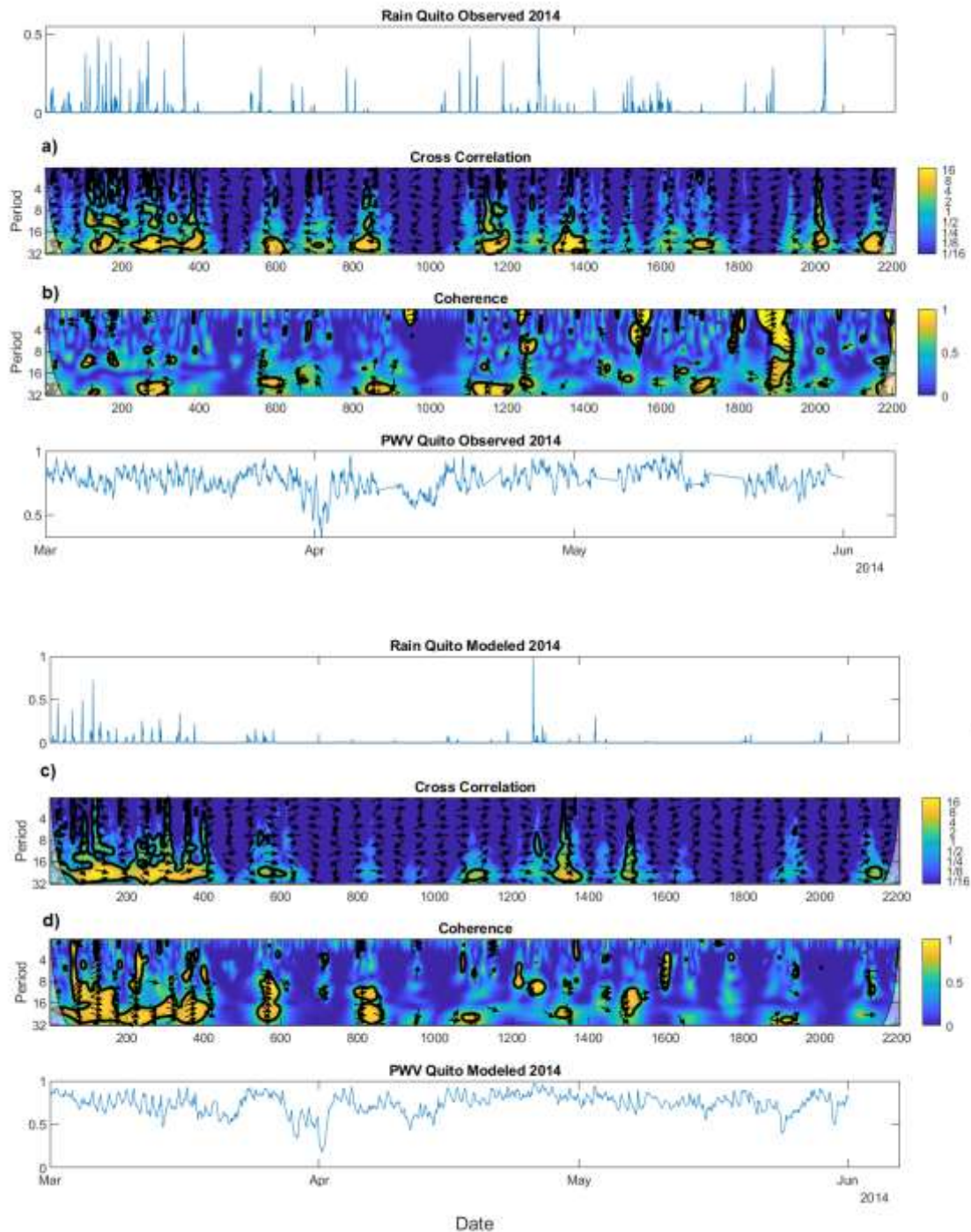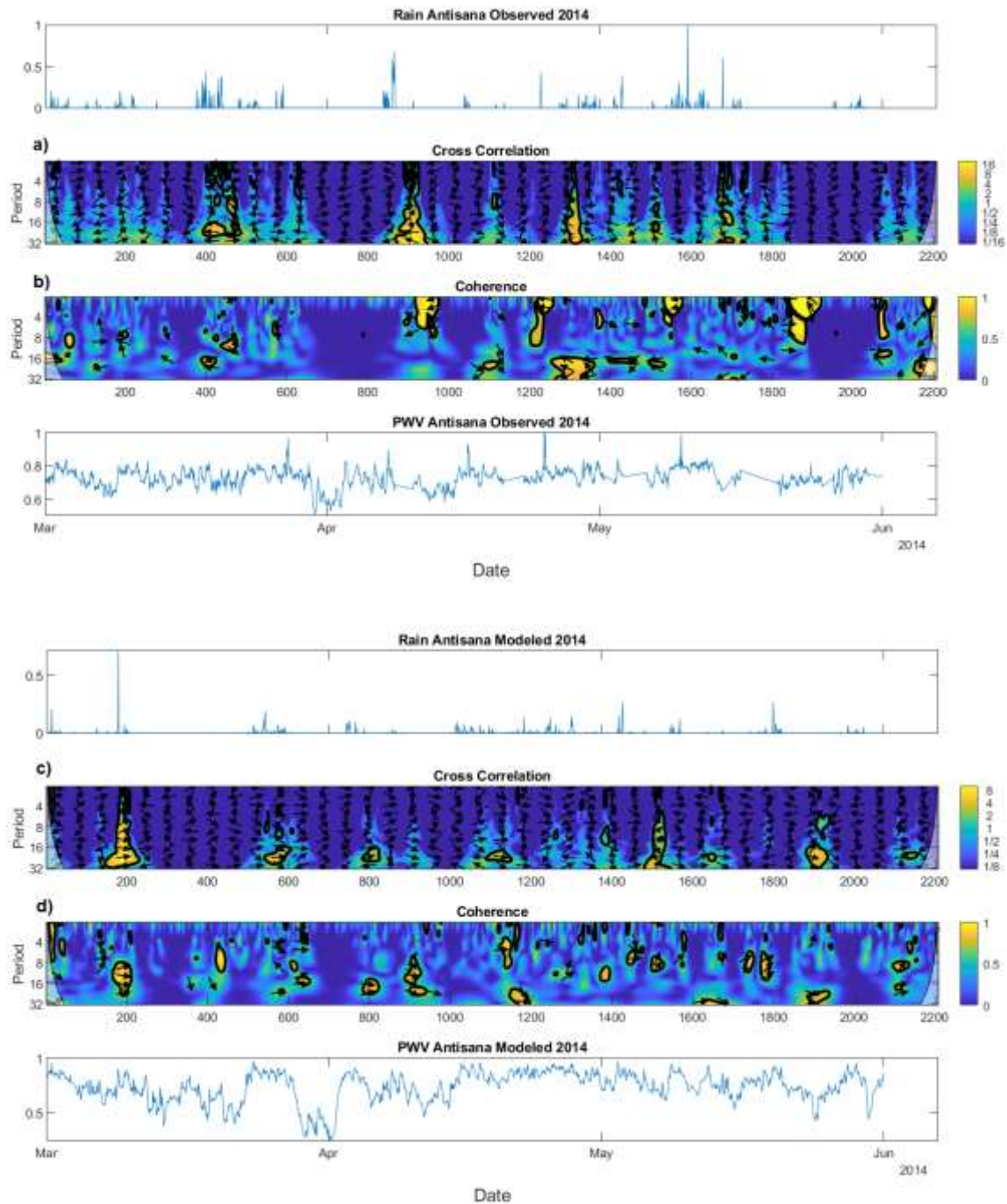
Figure 21 Zoom-in of the continuous wavelet cross-spectrum XWT and wavelet coherence scalogram and PWV in HMA. (a) XWT observed, (b) observed coherence, (d) XWT modeled (d) modeled coherence.

4.3.2.1 Seasonal and daily behavior of modeled and observed data

When analyzing the dynamic behavior of the model vs. using the in-situ data, in the case of PWV the monthly averages of the observed and modeled values are compared. Thus, the bimodal behavior of the rainy season in Quito HV1 is reproduced, and its subsequent increase in PWV in the rainy months around of March and September (Figure 22a). Similar behavior can be seen in the case of Antisana HM1, where despite the fact that there is no bimodality in the precipitation (Ruiz-Hernández et al., 2021), the model moderately adjusts to the behavior, but it systematically underestimates the monthly values (Figure 22b).

However, in the case of the daytime cycle, the first discrepancy between the model and the observed PWV data for the case of HV1 can be seen (Fig. 22c). While the model finds the highest values in the evening hours, the observed data shows that the highest values are obtained at midnight. Finally, in the case of HMA (Fig. 24d) at the daytime level, any other discrepancy between the modeled and observed values is hardly significant, other than the visible bias of the data.

In the case of precipitation we see that, at the seasonal level, the model reproduces precipitation fairly well for both the case of HV1 and HMA (Fig 23 a, b) with respect to the bimodal behavior that determines the rainiest periods. However, the modeling of this behavior is not perfect, in the case of HV1, the precipitation in the months of April, May and September is underestimated, while it is overestimated in the months of November and December. In the case of HMA, the model is even less exact, underestimating the month of May, and overestimating the months from June to December.
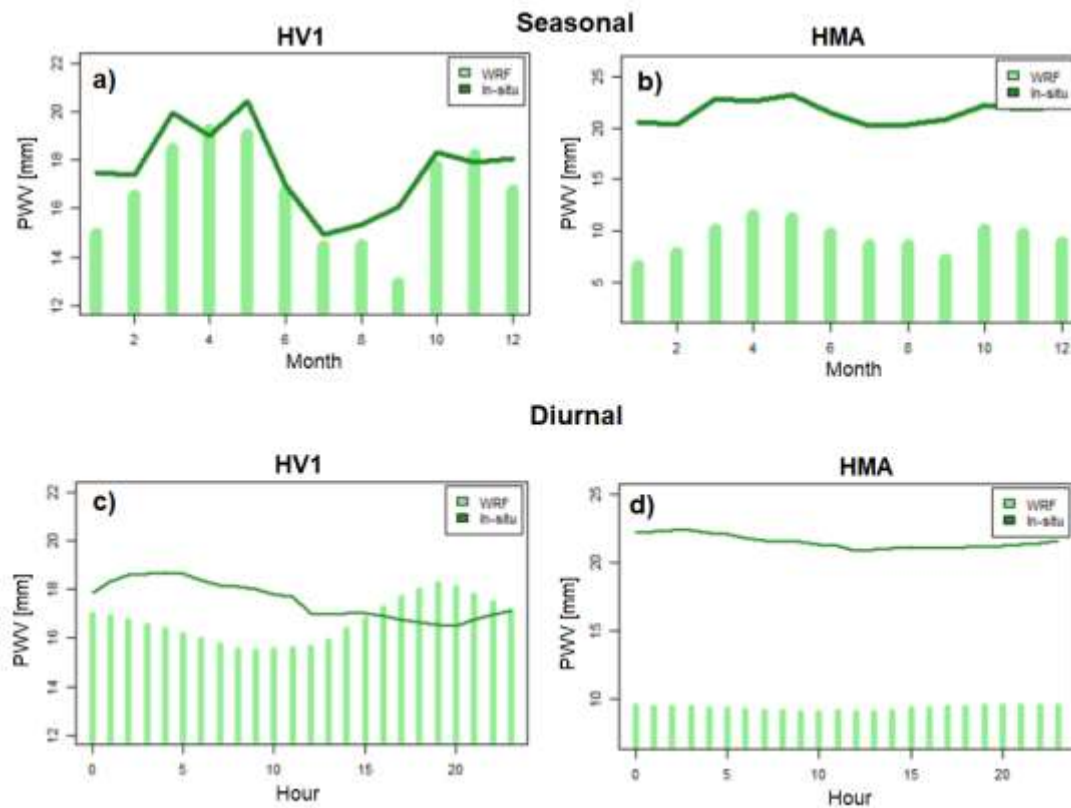
Figure 22 Comparison of the behavior of the mean PWV observed and modeled by WRF at the seasonal level for HV1, Quito (a), HMA, Antisana (b). And at the hourly level or daytime cycle for HV1 (c) and HMA (d).

At the hourly level, the results are different. WRF modeling reproduces the daytime precipitation cycle quite well (Fig. 23c), with a weak under-estimation during the early morning precipitation at 0h00. It is assumed that the origin of this precipitation is given by advective transport (Campozano et al., 2016). The rainiest hours of the day is found around 5:00 p.m. and is correctly simulated by the model.

In the case of HMA (Fig. 23d), the model represents an opposite behavior, the rainiest hours that occur around 11:00 a.m. are those with the least precipitation, while at night, in which the observed precipitation is not presented, are those in which the model indicates more precipitation. It can be seen that the more accurate the PWV estimate is, the better the precipitation estimate will be. This diurnal cycle bias was already described in Junquas et al. (2022) in this region as to be related to a westward shift of the nocturnal precipitation hotspot pattern in the Andes-Amazon transition region.
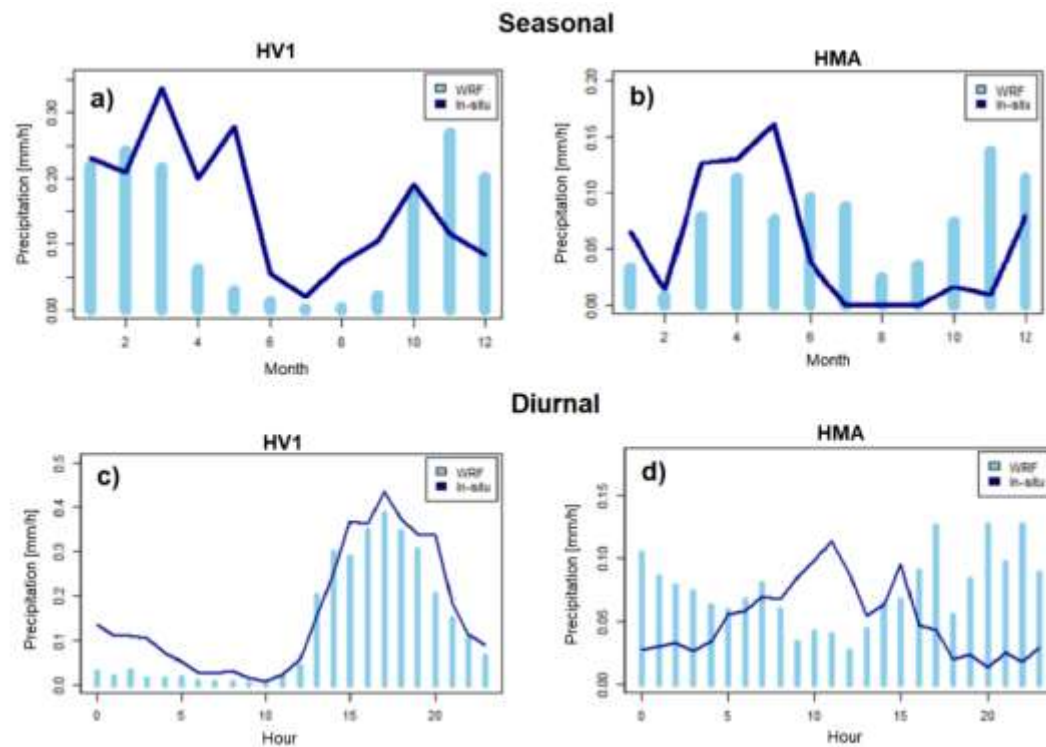
Figure 23 Comparison of the behavior of the precipitation observed and modeled by WRF at the seasonal level for HV1, Quito (a), HMA, Antisana (b). As well as at the hourly level or day cycle for HV1 (c) and HMA (d).

4.3.2.2 Analysis by event, with the discrete wavelet transform

The detailed analysis of heavy precipitation events that show correlation and coherence with PWV increases is presented below in Section 4.5.1. The events are presented in Table 12 for observed and modeled intense precipitation events. The ordinal section of the table indicates the order of the intensity of the event, from the most intense (1st) to the least intense, in descending order. From the sample of 20 events, 9 events selected for HV1 and 6 HMA events remained. For WRF modeled, 9 events were selected for HV1, and 8 for HMA.

Table 12 Selected intense heavy precipitation events consistent with increases in PWV, for observed and modeled data.

HV1 (Quito) Observed			HMA (Antisana) Observed		
No.	Time	Precipitation intensity [mm/h]	No.	Time	Precipitation intensity [mm/h]
1°	1/6/2014 15:00	19.8	1°	5/10/2014 15:00	9.5
4°	8/26/2014 17:00	14.6	3°	5/14/2014 11:00	5.99
5°	11/17/2014 18:00	13.6	5°	4/24/2014 12:00	4.1
7°	2/22/2014 20:00	11.5	6°	3/18/2014 17:00	4.1
8°	4/26/2014 20:00	10.9	7°	12/7/2014 15:00	4.9
9°	5/29/2014 20:00	10.9	8°	12/28/2014 12:00	3.9
12°	4/19/2014 00:00	9.4			
13°	3/12/2014 20:00	9.1			
14°	2/19/2014 19:00	8.4			
HV1 (Quito) Modeled			HMA (Antisana) Modeled		
No.	Time	Precipitation intensity [mm/h]	No.	Time	Precipitation intensity [mm/h]
1°	4/25/2014 17:00	13.88	1°	11/5/2014 20:00	18.53
2°	12/7/2014 14:00	12.03	2°	11/20/2014 17:00	16.57
3°	10/30/2014 14:00	10.15	3°	3/8/2014 22:00	13.27
5°	2/28/2014 15:00	8.998	4°	12/18/2014 19:00	8.3
6°	1/28/2014 14:00	8.58	6°	12/18/2014 20:00	8.07
8°	11/26/2014 15:00	7.69	7°	11/21/2014 23:00	6.63
9°	12/11/2014 15:00	7.27	8°	1/9/2014 0:00	6.14
10°	3/4/2014 18:00	6.88	9°	12/11/2014 14:00	5.75
11°	3/1/2014 22:00	6.31			

Using discrete wavelet transform MRA, the events of interest presented in Table 12 are decomposed under different levels of detail using the criteria found in Section 4.5.2, and cross-correlated independently with their PWV counterparts. Figure 24 shows the average of the correlation coefficients obtained at 2, 3 and 4 levels of detail of the operated series. The observed data is plotted in (Fig. 24 a, c) and the modeled data in (Fig. 24 b, d). In the case of the observed series, it is evident that the intense precipitation events were preceded by PWV peaks, showing a lag around 11 hours, which corresponds to the results found in Figures 20 a, b and Figure 21 a, b.

María Sheila Fabiola Serrano Vincenti

For the modeled series the lag is considerably reduced to 3 hours, as is the case of HV1 (Fig 24 b, d), which is consistent to the right arrows shown in Figures 20 b, c, as an almost in-phase condition. In the case of HMA, the lag is less, which is consistent with the phase information provided by the right arrows in Figure 20 a,d. Therefore, the main discrepancy of the model with the observed in PWV and precipitation during precipitation intense events is precisely the lag between the increases of PWV and precipitation.
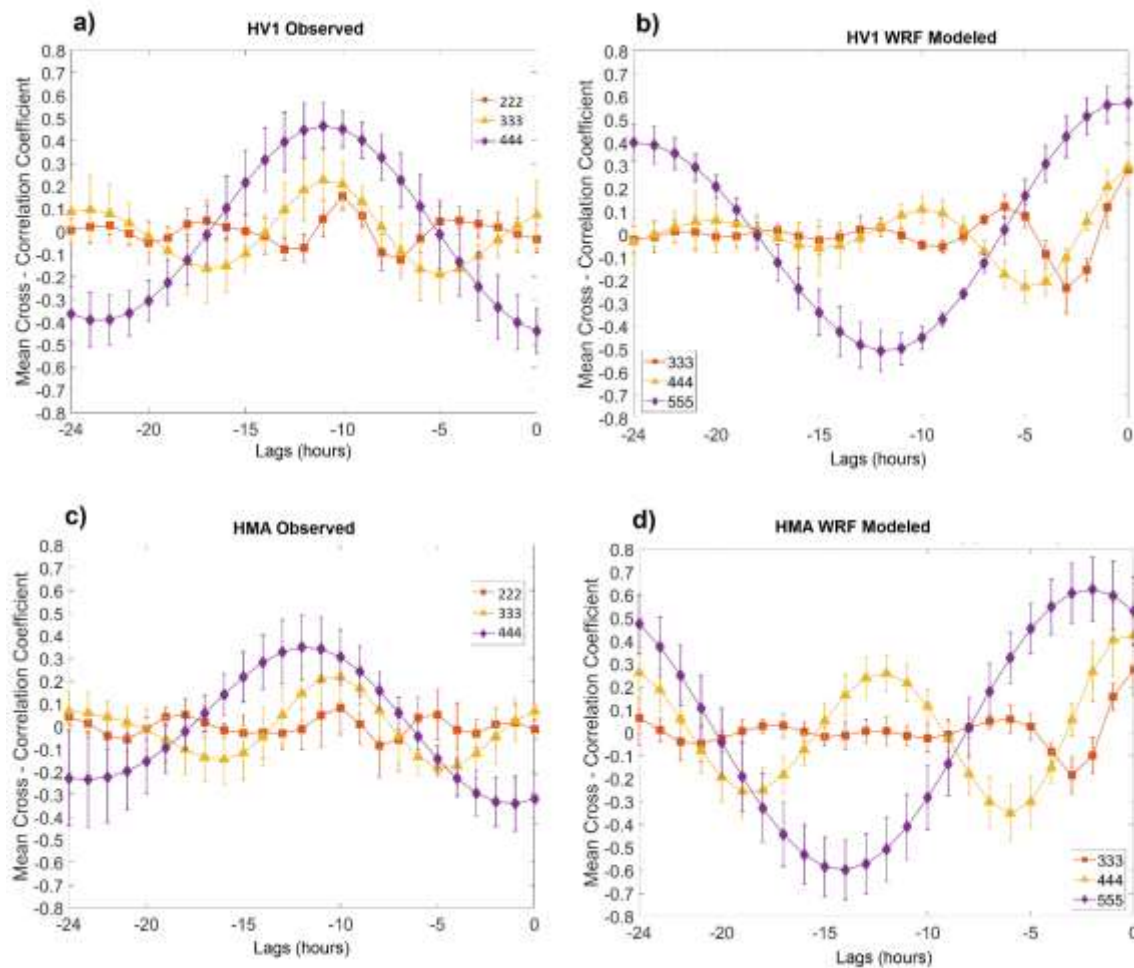

Figure 24 Average of the discrete wavelet correlation lag results recorded per level of detail for the selected precipitation events presented in Table 12, including their standard deviation.

4.3.3 General performance of the WRF model for the estimation of precipitation and PWV

Frequency histograms and Q-Q plots for the 8760 hourly data for 2014, of the compared values for the WRF model and the observed data for PWV are presented in Figure 25. In the case of HV1, an adequate location of the data is appreciated (Fig. 25a), with better

performance in high values, Q-Q plot of Fig. 25c. However, in the case of HMA, a systematic underestimation or bias error is evident (Fig. 25b), in addition to an underestimation at high PWV values (Fig. 25.c).
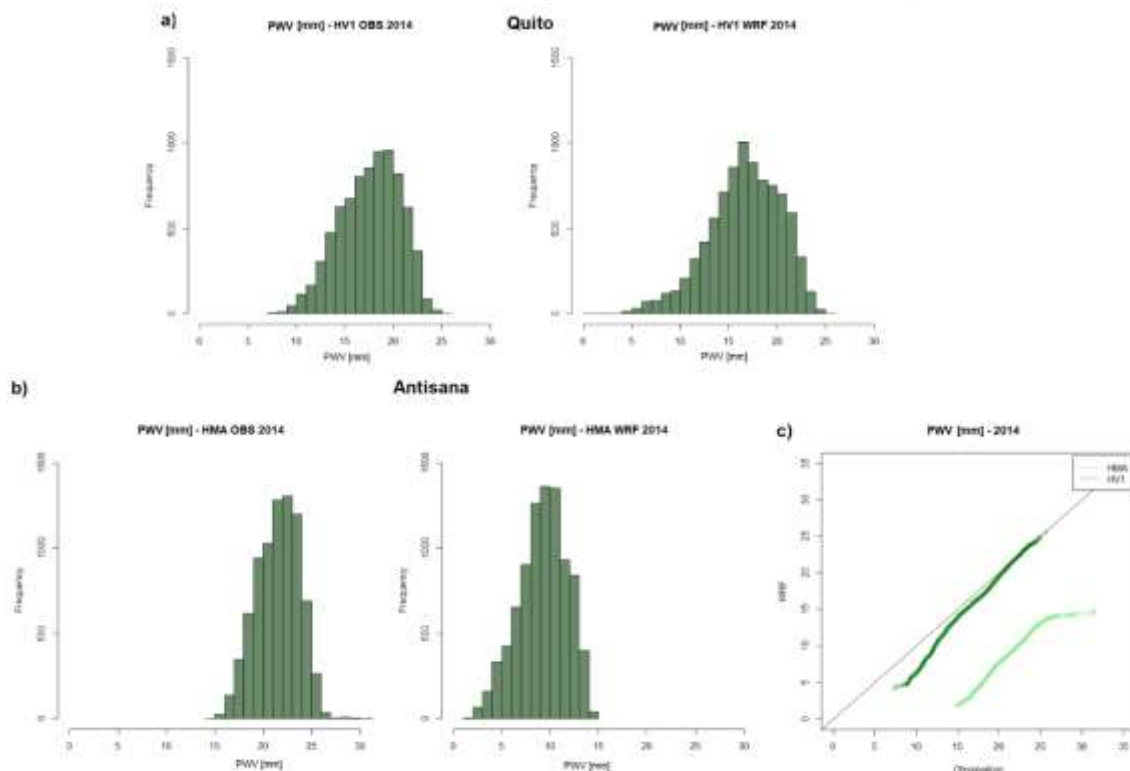

Figure 25 Frequency histograms of PW-GPS data in HV1 (Quito) and HMA (Antisana), note the bias error in the case of Antisana (b), having hourly data for the year 2014, there is a total of 8760 data.

When analyzing precipitation, Figure 26 presents the frequency histograms on a logarithmic scale, due to non-parametric frequency distributions such as exponential and log-normal (Serrano-Vincenti et al., 2020)(Fig. 26 a, b) and Q-Q plots of the modeled and in situ data (Fig 26c). These frequency distributions are determined by the sensitivity of the instrument that determines the minimum value that can be detected, which in the case of HV1 is 0.1 mm/h, while in HMA it is 0.001mm/h, while in the case of modeled values are 1.9 x 10-5 mm/h and 1.78 x 10-5 mm/h (Table 9).

When analyzing the Q-Q plots (Fig. 26c), as mentioned in (Heredia et al., 2019b), the pre-performance of the model decreases when modeling the most intense precipitation events, in the case of Quito (HV1), it can be seen an underestimation in extreme events, while in the case of Antisana HMA, these events are overestimated.

María Sheila Fabiola Serrano Vincenti

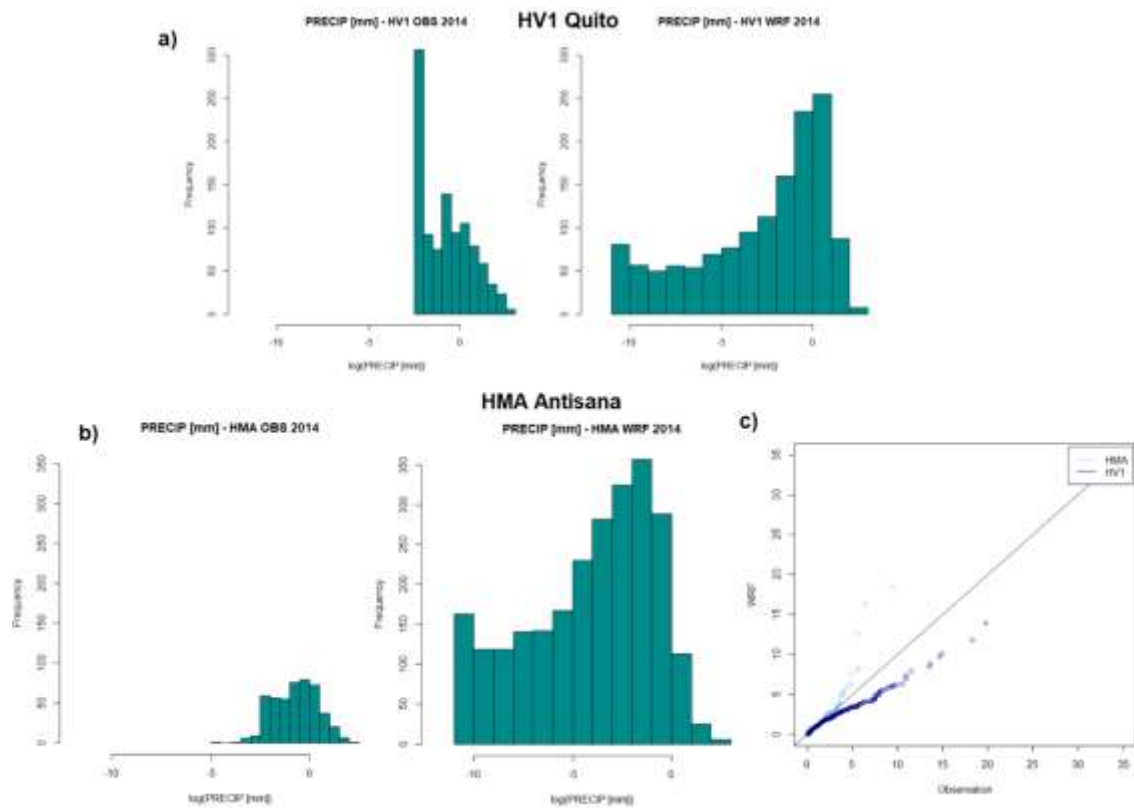Figure 26 Frequency histograms on a logarithmic scale of precipitation data in HV1 (Quito) and HM1 (Antisana), having hourly data for the year 2014, there is a total of 8760 data.

Finally, Figure 4.27shows the Taylor diagram of all the available meteorological variables. The variable in which the model has the lowest performance is precipitation, followed by PWV, and that the model's performance is better for HV1.
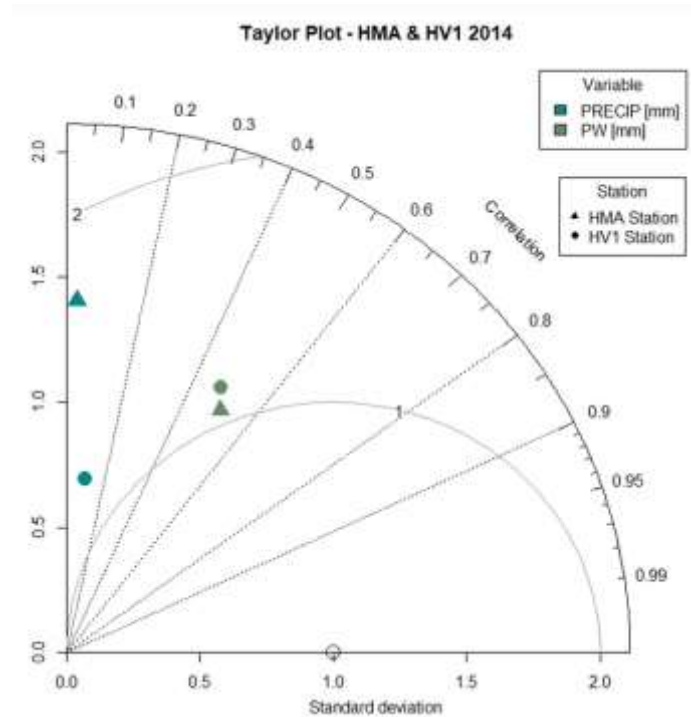
Figure 27 Taylor plot of WRF vs. on-site data for precipitation and PWV.

4.3.3.1 Harmonic analysis on observed and modeled values.

In order to understand the harmonic behavior of the observed signals vs. the modeled ones, the XWT (cross-spectrum) and WTC (coherence) between the observed signal, the modeled signal and the error of the observation that corresponds to the modeled value, minus the observed one, are presented. For ease, the three rainiest months are also shown, although the analysis was performed for the complete series in Annex 3-6.

In the case of PWV, it is necessary to indicate that the harmonic analysis will seek to find coherence between the periods of the observed and modeled series, regardless of other types of errors such as bias, detected in past sections. In this case, there is a greater correlation and coherence between the modeled and observed values, around the 24-hour period. In addition, the generally out-of-phase state between the modeled and observed events is evident, that is, the modeled signal is 12 hours out of phase with the observed one, or in other words, when there is an increase in PWV, the model represents a valley or it does not reproduce the loading and unloading periods of the observed PWV. For HMA (Fig. 29), the 24-hour characteristic period is also presented, however, the anti-phase condition is no longer so evident, being mostly dispersed and difficult to characterize, even though the coherence is less than in the case of HV1.

María Sheila Fabiola Serrano Vincenti

Figure 28 Normalized time series of observed and modeled PWV in HV1. As well as the XWT (a) and WCT (b) between the observed and modeled values. And the XWT (c) and WCT (d) between the observed values and their error.

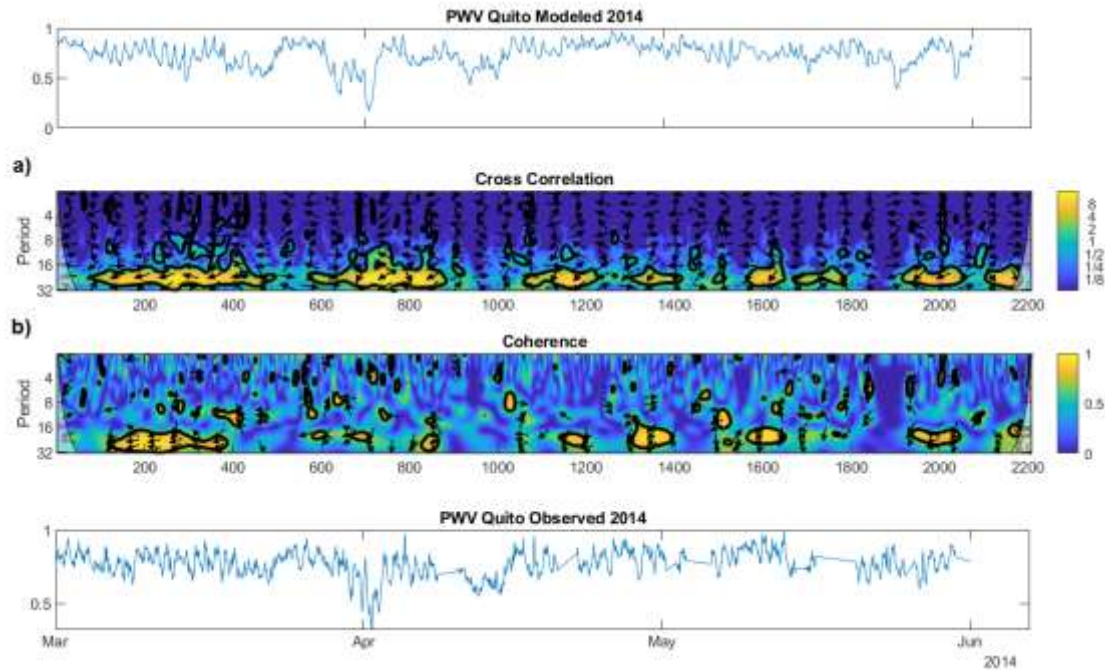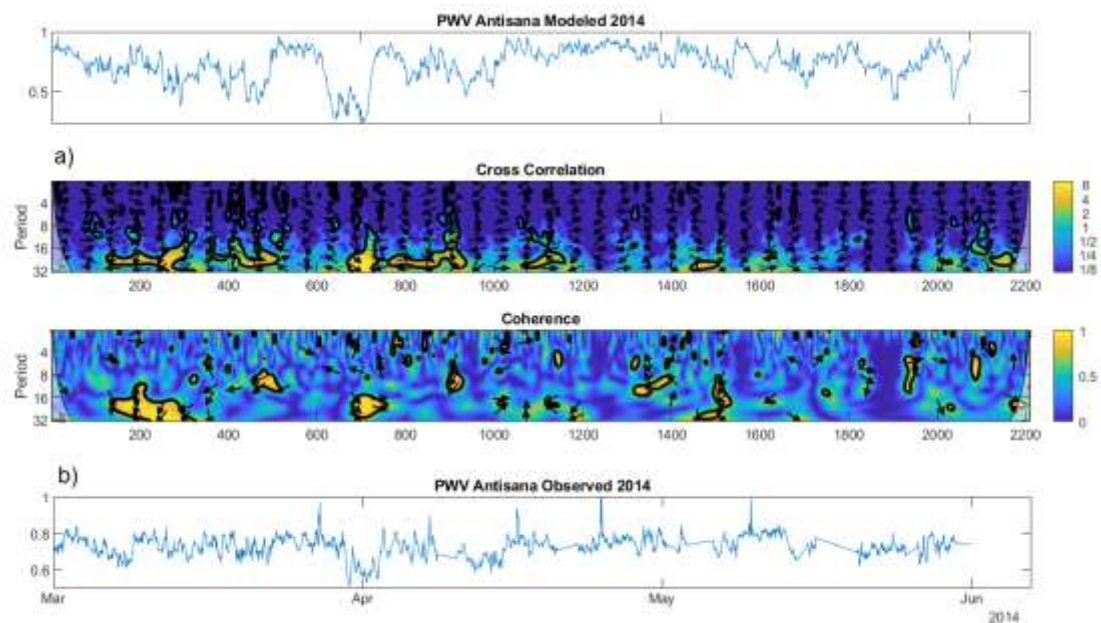
Figure 29 Normalized time series of observed and modeled PWV in HMA. As well as the XWT (a) and WCT (b) between the observed and modeled values. And the XWT (c) and WCT (d) between the observed values and their error

In the case of the HV1 precipitation Fig. 30, a better coherence WCT and correlation XWT between the observed signal and the modeled one can be seen, as shown by the direction of

the arrows to the right, which corresponds to the phase state between the signals. That is to say that the model detects the observed precipitation in the correct way, in its seasonal dynamic behavior and daytime cycle, as shown in Fig. 23 a, c. In the case of the HMA (Fig. 31), the correlation and coherence are lost, a behavior that also shows the low reproduction of the model in the seasonal and daytime cycle (Fig. 23 b, d). In both cases, the error in HV1 and HMA occurs most of the time, and it is out of phase because it arises from the subtraction of the analyzed values.
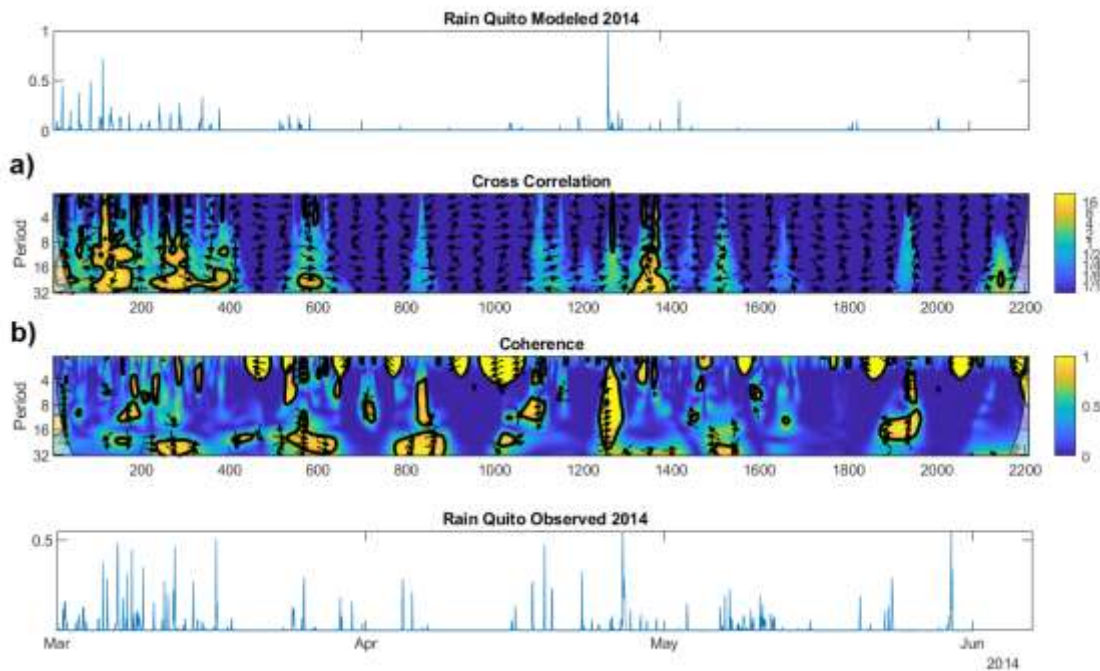

Figure 30 Normalized time series of observed and modeled precipitation in HV1. As well as the XWT (a) and WCT (b) between the observed and modeled values. And the XWT (c) and WCT (d) between the observed values and their error.
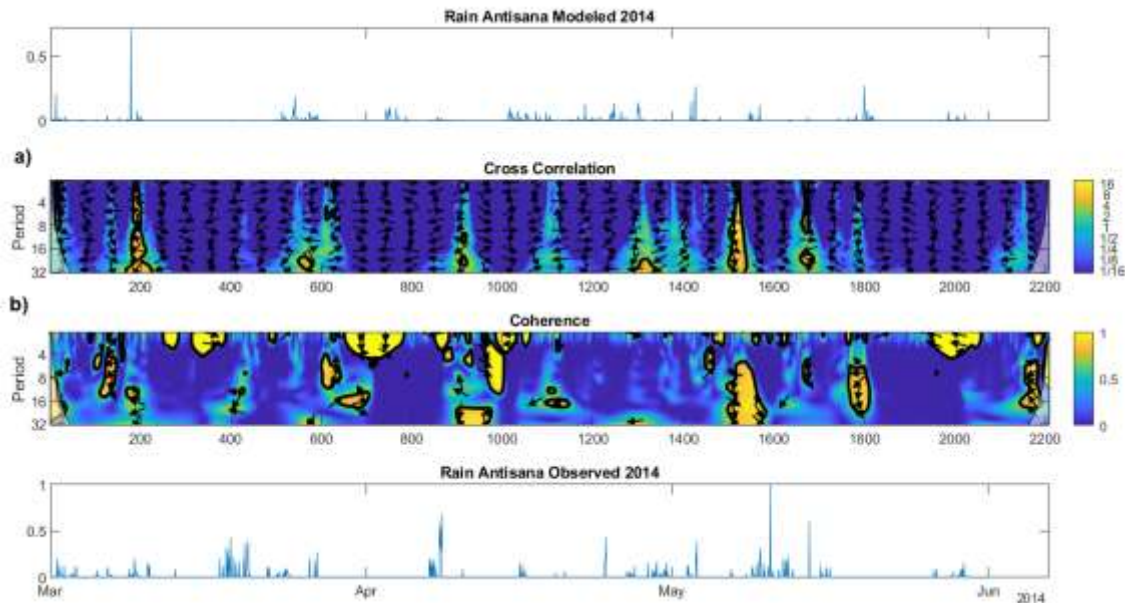
Figure 31 Normalized time series of observed and modeled precipitation, in HMA. As well as the XWT (a) and WCT (b) between the observed and modeled values. And the XWT (c) and WCT (d) between the observed values and their error.

4.4 Discussion

This research work focuses on two points that can improve atmospheric modeling through NWPs, such as WRF. The first is the verification of relationships between PWV increases and intense precipitation, and its possible optimization on the predictive potential of these events, as several research point (Benevides et al., 2015; Bonafoni et al., 2019; H. Li et al., 2020). And the second is to evaluate the statistical and harmonic relationship between the observed and modeled values. Harmonic analysis allows us to find coherent behaviors between the observed and modeled signals, despite other statistical discrepancies, such as bias errors or the presence of outliers (Chui, 1992; Daubechies, 1992; Mallat, 1999). In this way, it will be possible to locate the main weaknesses of the model and propose strategies to improve it.

In the first place, regarding the relationship between PWV and precipitation in observed and modeled data, there are two findings: the first are the periods of charge and discharge of PWV in the atmosphere, that is to say that the PWV recharges and discharge naturally in the atmosphere (Meza et al., 2020; Torri et al., 2019). However, these periods are not uniform for all the study areas, but rather depend on the local daytime precipitation cycle: the PWV recharges during the least rainy hours of the day, and discharges at the most likely hours of precipitation (Serrano-Vincenti et al., 2022) as is shown at the Figure 12 and 23.

María Sheila Fabiola Serrano Vincenti

In the case in which the precipitation cycle is unimodal, as in HMA, where the rains occur mainly during the day and around 11:00 a.m., the characteristic periods will be principally at 24 and, secondary at 12 hours (Torri et al., 2019), but if the rainfall diurnal cycle it is bimodal, as in the case of HV1, the periods of 24 and 12 hours are more noticeable, where the most probable hours of precipitation are at 5:00 p.m. and there is a light peak around 0:00 a.m.. Also, the Fourier Transform, pointing an additional characteristic extra-period of 8 hours, as is shown in Table 11 which appears whit a low spectral density, i.e with low occurrence.

It is noticeably to see how the model correctly finds the characteristic periods observed for both HV1 and HMA (Table 11). However, when analyzing the observed precipitation time series and PWV relation, they are out of phase (Fig 20, 21; a) and b)), from the harmonic analysis it is found how the modeled and observed PWV signals are mostly in phase, that is to say that the PWV increases correspond to the precipitation increments (Fig 20, 21; c) and d)), and not out of phase as found in the observed data. The incorporation of these periods of charge and discharge of the PWV can be one of the main contributions for the improvement of the model and one of the reasons why real-time PWV assimilation improves precipitation estimation (Risanto et al., 2021; Shoji et al., 2009).

On the other hand, regarding the relationship between the "peaks" or sudden increases in water vapor before heavy precipitation events, it is necessary to indicate that these do not always occur, since they are subject to the presence of other factors (Benevides et al., 2019; Sapucci et al., 2019; Serrano-Vincenti et al., 2022), as shown in the Table 12, despite the fact that the 20 most intense precipitation events were taken, after eliminating those not independent, only 9 precipitation events in HV1 and 6 in HMA, presented PWV peaks that preceded them, so it is still difficult to incorporate these findings systematically into a model, and more research is needed on the factors that produce these peaks and the lag that exists between them (Bonafoni et al., 2019).

Regarding the seasonal behavior of the model, in the case of the PWV, HV1 has an adequate reproduction of the seasonal values is appreciated (Fig. 22a). In HMA there is a systematic bias error of around 15 mm (Fig. 22b). This error then leads to an estimate of precipitation by the model that is better for HV1 (Fig. 23 a, c) and less adequate for HMA (Fig 23 b, d), as is also shown in a comparative way by the Taylor diagram (Fig. 27).

When analyzing the daytime cycle, the findings found in in the past sections are reinforced, identifying that the lag occurs on the PWV values, since the model cannot correctly estimate the PWV charge and discharge values (Fig. 22c). When performing the harmonic analysis, it can be clearly seen how the observed values of the correlation and coherence of the observed

and modeled values are mostly in out-of-phase (arrows to the left) (Fig 19 a and b). When analyzing the behavior of PWV in HMA, one of the strengths of harmonic analysis is seen, which allows us to analyze coherence and correlation between the observed and modeled data, regardless of systematic errors such as bias. In this case, it can be seen that the error is not only bias, but it is difficult to identify if the signals are in phase or out of phase, since the direction of the arrows is very dispersed (Fig 29 a and b).

From these behaviors is appreciated that the better the estimate of the PWV, the better the estimate of precipitation, which is modeled quite adequately in HV1, both at the seasonal level and in the daytime cycle, while the estimate of precipitation for the case of HMA, it still has important shortcomings both at the seasonal level (Fig. 23b) and the daytime cycle (Fig. 23d). On the other hand, the correct estimation of the PWV diurnal cycle has been one of the main challenges in its modeling, finding that WRF has a good performance, but in low-lying areas (Alshawaf et al., 2015; González et al., 2013), where the transport of humidity is given on a synoptic scale, while the inclusion of the steep terrain variable may lead to errors in its estimation (Qian et al., 2020).

Finally, the statistical analysis shows us that when analyzing the modeled precipitation intensities, the model fails to estimate the highest values, underestimating them in the case of HV1 and overestimating them in the case of HMA (Fig 25c). Correctly estimating heavy precipitation events is one of the most difficult NWP problems to solve (Herman & Schumacher, 2016). On the other hand, in the estimation of PWV for HV1, there are quite adequate results for high values and an underestimation occurs in small values (Fig. 25c); while in HMA, in addition to the bias errors and the no-coherence between the modeled series, the model significantly underestimates the high PWV values (Fig. 26c).

4.5 Conclusions

In light of the results, the importance of the correct estimation of PWV becomes important for an adequate modeling of precipitation events. The great contribution of the inclusion PWV-GNSS data into processes of NWP results in a significant improvement of the estimation of precipitation, and above all, intense precipitation (Risanto et al., 2021; Shoji et al., 2009). However, it is evident that these advantages are possible in areas where both, the modeling processes and the in-situ meteorological platforms, are served with running and continuously interconnected GPS geodetic networks with significant computing power. And these

conditions are not available for every countries or regions. This is precisely the limitation of this study; first, not to being able to work with more locations in the study area, and to have longer time series which may represent characteristics of different climatologies in a proper way. In the same way, the resolution of the data determine the periods found, a better temporal resolution could analyze with better detail shorter periods.

Therefore, the methods for the analysis and discussion of the results provided in this research, as a way to interpret the main discrepancies in the estimation of PWV can be very useful, especially when locating the presence of atmospheric charge and discharge periods, which depend on the daytime cycle of the atmosphere and precipitation as well as on the area, which could be as complex as the studied in this research. Several studies focus on the correct modeling of rainfall linked to topography or statistical strategies to improve the estimation of the model (Heredia et al., 2019; Mourre et al., 2016); or focus on PWV modeling (González et al., 2013; Qian et al., 2020). But the study of rainfall and PWV interactions, as well as their causal relationships, are necessary not only for a better understanding of their particular dynamics, but also for their proper modeling and forecasting.

Also, even if the model could not be modified, the post-hoc addition and/or correction of these founded interactions as the charge and discharge periods, could be a valuable strategy in a future work of the process of estimating the PWV and hence in precipitation estimation. Finally, satellite data, as well as the use of a wider network of GNSS and meteorological data, will be useful in a future work, where other strategies as the inclusion of artificial intelligence accompanied with NWP will be able to make a difference in improving the prediction and forecasting of extreme precipitation events.

Chapter 5

5.1 Synthesis

The present investigation has sought to give a deep insight into the relationships between PWV and intense rain over the Eastern Pacific Ocean and Ecuadorian regions. The instruments and techniques that have been used for this purpose are varied, and therefore, achieved different objectives which are summarized below.

First of all, in **Chapter 2** it should be mentioned the empirical relationship between the PWV and the maximum precipitation values over the ocean, which was independent of the analyzed time and space; due to the nature of satellite scanning, there were no continuous series in time or space over the study area. The high resolution used in this part of the study ensures that the series was not filled with modeled data, and therefore, these values are directly observed by the satellite as often called level 2A12 data.

In **Chapter 2**, two main findings are presented: the first one shows the superior performance of the PWV to spatially locate rain events than SST, and to describe their maximum values. The empirical description of the maximum rainfall values associated with the PWV was described by an exponential model, closely followed by the power law model described by Peters and Neelin (2006), and by a stretched exponential. It is important to indicate that the proposed model found does not limit the presence of rain of lower intensities to those indicated by the exponential, but that the maximums achieve a degree of intensity depending on the amount of PWV present in the atmosphere at a given moment.

In **Chapter 3**, estimated PWV values from GNSS stations on land are already available, a technique considered accurate and reliable (Businger et al., 1996). The main challenge was locating meteorological stations that work at the same resolution (hourly in this case) than their neighboring GNSS stations. At the end, five sites with different climatologies were selected: Coast, Andes Valley (two sites), High Mountain and Amazon.

By having continuous time series, PWV and rainfall cause-effect relationships were analyzed using harmonic analysis tools, firstly to establish periods of charge and discharge of the PWV whose maximum values coincided with the minimum values of rainfall. That is, when both periods of interest were in antiphase, noting that the local climate of the area influenced the outcome: in those places where the most probable and maximum values of rain occurred at night, such as on the Coast (Esmeraldas), the maximum values PWV occurred during the day;

conversely, in places like the High Mountain (Antisana), where rain most likely occurs during the day, the PWV values occur at night.

Also, after analyzing the intense rain event independently, it was found a similarity with previous remarks like the ones found in Sapucci et al. (2019), where it is said that PWV peaks appear before intense rain events, although they do not occur all the time but on occasion: sometimes the PWV peak occur without triggering any rain or intense rain event, while other times intense rain takes place without the evidence of a previous PWV peak (Table 3. 4). However, is important to note that when this phenomenon occurs, the hourly difference or lag between PWV and rain is 11 hours, in congruity with the phase analysis of the covariant periods of PWV and rain which are in opposition.

In order to check whether these interactions are correctly assimilated in models such as WRF, **Chapter 4** analyses the relationship between PWV and rainfall and the NWP model of Heredia et al. (2019), the latter validated for high altitude areas (more than 2000 m.s.n.m). Despite the good performance of the model, which is capable of reproducing the climatological behavior of rainfall over the studied areas fairly well, the estimation of the PWV presents errors, frist in the estimation of the true magnitude, mainly bias deviation, as was described in Antisana station. Also the second discrepancy occurs when describing their diurnal cycle. Apparently, the WRF model considers that the behavior of PWV and rain are in phase or with a minimum lag (less than 3 hours as shown in Figure 24), instead of the results obtained by the observed data which state the opposite.

5.2 Future research

This dissertation has opened up new questions and challenges that lead to possible future research topics.

For instance, it has been identified the need for more PWV data outside of the methodology that involves the use of GNSS and in-situ weather stations. It is a fact that insufficient coverage, both geographically and temporally, limits research on these topics. For this reason, it is necessary to have PWV data from the new constellation of satellite products, such as the one launched by the Global Precipitation Measurement (GPM) which include already PWV products -although still at level 2-, that must be compared with in-situ data, such as the one presented in this research, for data validation or to find limitations in its use.

It is known that the good performance that satellite products have in the estimation of variables such as rain and water vapor, is limited to monthly and even daily scales, but results at hourly scales are not as reliable, especially when high altitude stations are analyzed or those that

correspond to steep topographies like the ones chosen in this study (Panegrossi et al., 2017). In this way, carrying out a validation or correction of the data is a work that would allow the extension of this study to areas with diverse climatology.

Once data is available from a larger number of study areas, it will be possible to analyze the incorporation of an empirical model that can predict the maximum rainfall values associated with different PWV values, but with land data associated with different climatologies. Having a relationship that determines the maximum rainfall values is undoubtedly an important tool in the analysis of extreme events and their possible forecast.

Finally, both findings, the exponential empirical model between PWV and maximum rainfall, as well as the periods of charge and discharge of the PWV as a function of the diurnal cycle of rainfall, should be carefully reviewed for their incorporation into NPW models, either as new relationships from atmospheric physics, or as post-hoc processes that could improve the correct estimation of PWV and thus rainfall and intense rain events.

References

Adams, D. K., Gutman, S. I., Holub, K. L., & Pereira, D. S. (2013a). GNSS observations of deep convective time scales in the Amazon. *Geophysical Research Letters*, *40*(11), 2818–2823. https://doi.org/10.1002/grl.50573

Adams, D. K., Gutman, S. I., Holub, K. L., & Pereira, D. S. (2013b). GNSS observations of deep convective time scales in the Amazon. *Geophys. Res. Lett*, *40*, 2818–2823. https://doi.org/10.1002/grl.50573

Ahmed, F., & Neelin, J. D. (2018). Reverse engineering the tropical precipitation-buoyancy relationship. *Journal of the Atmospheric Sciences*, *75*(5), 1587–1608. https://doi.org/10.1175/JAS-D-17-0333.1

Albán, M. F. (2019). *Técnica de posicionamiento puntual preciso (PPP) aplicada a la estimación de vapor de agua precipitable (PWV) de la estación EPEC* [ESPE]. http://repositorio.espe.edu.ec/bitstream/21000/14672/1/T-ESPE-040083.pdf

Alshawaf, F., Fuhrmann, T., Knöpfler, A., Luo, X., Mayer, M., Hinz, S., & Heck, B. (2015). Accurate Estimation of Atmospheric Water Vapor Using GNSS Observations and Surface Meteorological Data. *IEEE Transactions on Geoscience and Remote Sensing*, *53*(7), 3764–3771. https://doi.org/10.1109/TGRS.2014.2382713

Arakawa, A. (2006). Scaling tropical rain. *Nature Physics*, *2*(June), 373–374. https://doi.org/10.1038/nphys323

Ayala, M. F., D., C.-V., & Tierra, A. (2018). Relación espacio-temporal entre estaciones utilizadas para el relleno de datos de precipitación en Chone, Ecuador. *Revista Geografica Venezolana*, *59*(2), 298–313. https://doi.org/http://www.redalyc.org/articulo.oa?id=347760473005 Cómo

Baty, F., Ritz, C., Charles, S., Brutsche, M., Flandrois, J.-P., & Delignette-Muller, M.-L. (2015). A Toolbox for Nonlinear Regression in R : The Package nlstools . *Journal of Statistical Software*, *66*(5). https://doi.org/10.18637/jss.v066.i05

Bell, T. L., Kundu, P. K., & Kummerow, C. D. (2001). Sampling errors of SSM/I TRMM rainfall averages: Comparison with error estimates from surface data and a simple model. *Journal of Applied Meteorology*, *40*(5), 956. https://doi.org/10.1175/1520-0450(2001)040<0938:seosia>2.0.co;2

Bellenger, H., Yoneyama, K., Katsumata, M., Nishizawa, T., Yasunaga, K., & Shirooka, R. (2015). Observation of Moisture Tendencies Related to Shallow Convection. *Journal of the Atmospheric Sciences*, *72*. https://doi.org/10.1175/JAS-D-14-0042.1

Benedetto, J. (1997). *Harmonic analysis and applications*. Boca Raton, FL, USA: CRC Press. Inc. https://doi.org/https://doi.org/10.1201/9781003068839

Benevides, P., Catalao, J., & Miranda, P. M. A. (2015). On the inclusion of GPS precipitable water vapour in the nowcasting of rainfall. *Natural Hazards and Earth System Sciences Discussions*, *3*(6), 3861–3895. https://doi.org/10.5194/nhessd-3-3861-2015

Benevides, P. ;, Catalao, J., & Nico, G. (2019). Neural network approach to forecast hourly intense rainfall using GNSS precipitable water vapor and meteorological sensors. *Remote Sensing*, *11*(8). https://doi.org/10.3390/rs11080917

Bernstein, D. N., & Neelin, J. D. (2016). Identifying sensitive ranges in global warming precipitation change dependence on convective parameters. *Geophysical Research Letters*, *43*(11), 5841–5850. https://doi.org/10.1002/2016GL069022

Betts, A. K., & Ridgway, W. (1989). Climatic equilibrium of the atmospheric convective boundary layer over a tropical ocean. In *Journal of the Atmospheric Sciences* (Vol. 46, Issue 17, pp. 2621–2641). https://doi.org/10.1175/1520-0469(1989)046<2621:CEOTAC>2.0.CO;2

Bevis, M. ;, Businger, S. ;, & Chiswell, S., ;Herring, T.A.; Anthes, R.A.; Rocken C.; Ware, R. H. (1994). GPS meteorology: mapping zenith wet delays onto precipitable water. *J. of Appl. Meteorol*, *33*, 379–386. https://doi.org/https://www.jstor.org/stable/26186685

Bevis, M. ;, Businger, S. ;, Herring, T. A. ;, Rocken, C. ;, Anthes, R. A. ;, & Ware, R. H. (1992). GPS meteorology: remote sensing of atmospheric water vapor using the Global Positioning System. *J. Geophys. Res.*, *97*, 787–801. https://doi.org/92JD01517. 0148-0227/92/92JD-01517505.00

Bonafoni, S., & Biondi, R. (2016). The usefulness of the Global Navigation Satellite Systems (GNSS) in the analysis of precipitation events. *Atmospheric Research*, *167*, 15–23. https://doi.org/10.1016/j.atmosres.2015.07.011

Bonafoni, S., Biondi, R., Brenot, H., & Anthes, R. (2019). Radio occultation and ground-based GNSS products for observing, understanding and predicting extreme events: A review. *Atmospheric Research*, *230*(April), 104624. https://doi.org/10.1016/j.atmosres.2019.104624

Brenot, H., Ducrocq, V., Walpersdorf, A., Champollion, C., & Caumont, O. (2006). GPS zenith delay sensitivity evaluated from high-resolution numerical weather prediction simulations of the 8-9 September 2002 flash flood over southeastern France. *J. Geophys. Res*, *111*, 15105. https://doi.org/10.1029/2004JD005726

Brenot, H., Walpersdorf, A., Reverdy, M., van Baelen, J., Ducrocq, V., Champollion, C., Masson, F., van Baelen, J., Doerflinger, E., Collard, P., & Giroux, P. (2014). A GPS network for tropospheric tomography in the framework of the Mediterranean hydrometeorological observatory Cévennes-Vivarais (southeastern France). *Atmospheric Measurement Techniques*, *7*(2), 553–578. https://doi.org/10.5194/amt-7-553-2014

Bretherton, C. S., Peters, M. E., & Back, L. E. (2004). Relationships between water vapor path and precipitation over the tropical oceans. *Journal of Climate*, *17*(7), 1517–1528. https://doi.org/10.1175/1520-0442(2004)017<1517:RBWVPA>2.0.CO;2

Businger, S., Chiswell, S. R., Bevis, M., Duan, J., Anthes, R. A., Rocken, C., Ware, R. H., Exner, M., VanHove, T., & Solheim, F. S. (1996). The Promise of GPS in Atmospheric Monitoring. *Bulletin of the American Meteorological Society*, *77*(1), 5–18. https://doi.org/10.1175/1520-0477(1996)077<0005:TPOGIA>2.0.CO;2

Buytaert, W., Celleri, R., Willems, P., Bièvre, B. de, & Wyseure, G. (2006). Spatial and temporal rainfall variability in mountainous areas: A case study from the south Ecuadorian Andes. *Journal of Hydrology*, *329*(3–4), 413–421. https://doi.org/10.1016/j.jhydrol.2006.02.031

Calori, A., Santos, J. R., Blanco, M., Pessano, H., Llamedo, P., Alexander, P., & de la Torre, A. (2016). Ground-based GNSS network and integrated water vapor mapping during the development of severe storms at the Cuyo region (Argentina). *Atmospheric Research*, *176–177*, 267–275. https://doi.org/10.1016/j.atmosres.2016.03.002

Campozano, L., Célleri, R., Trachte, K., Bendix, J., & Samaniego, E. (2016). Rainfall and Cloud Dynamics in the Andes: A Southern Ecuador Case Study. *Advances in Meteorology*, *2016*, 1–15. https://doi.org/10.1155/2016/3192765

Campozano, L., Trachte, K., Célleri, R., Samaniego, E., Bendix, J., Albuja, C., & Mejia, J. F. (2018). Climatology and Teleconnections of Mesoscale Convective Systems in an Andean Basin in Southern Ecuador: The Case of the Paute Basin. *Advances in Meteorology*, *2018*, 1–13. https://doi.org/10.1155/2018/4259191

Cho, H.-K., Bowman, K. P., & North, G. R. (2004). A Comparison of Gamma and Lognormal Distributions for Characterizing Satellite Rain Rates from the Tropical Rainfall Measuring Mission. *Journal of Applied Meteorology*, *43*(11), 1586–1597. https://doi.org/10.1175/jam2165.1

Chui, C. K. (1992). *An Introduction to Wavelets* (U. NY, Ed.). New York. Academic Press.

Clauset, A., Shalizi, C. R., & Newman, M. E. J. (2009). Power-law distributions in empirical data. *SIAM Review*, *51*(4), 661–703. https://doi.org/10.1137/070710111

Climate World, & Programme. (1986). *Guidelines on the quality control of surface climatological data* (Issue 111).

Cohen, E. (2011). *A statistical study of wavelet coherence for stationary and nonstationary processes* [Imperial College, London, United Kingdom]. https://www.ma.imperial.ac.uk/~eakc07/publications/THESIS.pdf

Daubechies, I. (1992). *Ten Lectures on Wavelets* (C.-N. R. C. S. in A. Mathematics, Ed.). SIAM. https://doi.org/https://doi.org/10.1137/1.9781611970104

Davis, C. A., Manning, K. W., Carbone, R. E., Trier, S. B., & Tuttle, J. D. (2003). Coherence of Warm-Season Continental Rainfall in Numerical Weather Prediction Models. *Monthly Weather Review- USA*, *131*, 2667–2679.

Dickman, R. (2003). Rain, power laws, and advection. *Physical Review Letters*, *018701*(2002). https://doi.org/https://doi.org/10.1103/PhysRevLett.90.108701

DSA. (2020). *Centre of Wether Forecast and Climate Studies*. DSA Satellite Division and Environmental Systems. Brazil Ministry of Science (BMS). http://satelite.cptec.inpe.br/acervo/goes.formulario.logic

Emardson, T. ;, & Derks, H. (1992). On the Relation Between the Wet Delay and the Integrated Precipitable Water Vapour in the European Atmosphere. *Meteorol. Appl.*, *6*, 1–12. https://doi.org/https://doi.org/10.1017/S1350482700001377

Erazo, B., Bourrel, L., Frappart, F., Chimborazo, O., Labat, D., Dominguez-Granda, L., Matamoros, D., & Mejia, R. (2018). Validation of satellite estimates (Tropical Rainfall Measuring Mission, TRMM) for rainfall variability over the Pacific slope and Coast of Ecuador. *Water (Switzerland)*, *10*(2). https://doi.org/10.3390/w10020213

Fernández L.I., Meza A.M., N. M. P. (2009). Determinación del contenido de vapor de agua precipitable (PWV) a partir de mediciones GPS: primeros resultados en Argentina. *Geoacta*, *34*(2), 35–57. https://doi.org/http://www.scielo.org.ar/scielo.php?script=sci_arttext&pid=S1852-77442009000200001&lng=es&nrm=iso

Furuzawa, F. a., & Nakamura, K. (2005). Differences of Rainfall Estimates over Land by Tropical Rainfall Measuring Mission (TRMM) Precipitation Radar (PR) and TRMM Microwave Imager (TMI)—Dependence on Storm Height. *Journal of Applied Meteorology*, *44*(3), 367–383. https://doi.org/10.1175/JAM-2200.1

Gamache, J. F., & Houze, R. A. (1983). Water Budget of a Mesoscale Convective System in the Tropics. In *Journal of the Atmospheric Sciences* (Vol. 40, Issue 7, pp. 1835–1850). https://doi.org/10.1175/1520-0469(1983)040<1835:WBOAMC>2.0.CO;2

Garratt, R. J. (1992). *The atmospheric boundary layer* (C. Press, Ed.). Elsevier: Amsterdam, The Netherlands. https://doi.org/https://doi.org/10.1016/0012-8252(94)90026-4

Ge, Z. (2008). Significance tests for the wavelet cross spectrum and wavelet linear coherence. *Annales Geophysicae*, *26*(12), 3819–3829. https://doi.org/10.5194/angeo-26-3819-2008

González, A., Expósito, F. J., Pérez, J. C., Díaz, J. P., & Taima, D. (2013). Verification of precipitable water vapour in high-resolution WRF simulations over a mountainous archipelago. *Quarterly Journal of the Royal Meteorological Society*, *139*(677), 2119–2133. https://doi.org/10.1002/qj.2092

Gray, W. M. (1998). The formation of tropical cyclones. *Meteorology and Atmospheric Physics*, *67*(1–4), 37–69. https://doi.org/10.1007/BF01277501

Grecu, M., & Anagnostou, E. N. (2001). Overland Precipitation Estimation from TRMM Passive Microwave Observations. *Journal of Applied Meteorology*, *40*(8), 1367–1380. https://doi.org/10.1175/1520-0450(2001)040<1367:OPEFTP>2.0.CO;2

Grinsted, A. (2014). *Wavelet Coherence*. http://www.glaciology.net/wavelet-coherence

Grinsted, A., Moore, J. C., & Jevrejeva, S. (2004). Application of the cross wavelet transform and wavelet coherence to geophysical time series. *Nonlinear Processes in Geophysics*, *11*(5/6), 561–566. https://doi.org/10.5194/npg-11-561-2004

Guan, L., & Kawamura, H. (2004). Merging satellite infrared and microwave SSTs: Methodology and evaluation of the new SST. *Journal of Oceanography*, *60*(5), 905–912. https://doi.org/10.1007/s10872-005-5782-5

Guevara Díaz, J. M. (2014). Uso correcto de la correlación cruzada en Climatología: el caso de la presión atmosférica entre Taití y Darwin. *Terra*, *30*, 79–102. https://doi.org/http://ve.scielo.org/scielo.php?script=sci_arttext&pid=S1012-70892014000100005&lng=es&tlng=es.

Haiden, T., & Kahlig, P. (1994). Modelling extreme precipitation events. *Österreichische Wasser- Und Abfallwirtschaft*, *October*.

Harp, R. D., & Horton, D. E. (2022). Observed Changes in Daily Precipitation Intensity in the United States. *Geophysical Research Letters*, *49, 19*(19). https://doi.org/https://doi.org/10.1029/2022GL099955

Heisenberg, W. (1927). Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Zeitschrift Fur Physik*, *11*, 561–566. https://doi.org/https://link.springer.com/article/10.1007%2FBF01397280

Heredia, M. B., Junquas, C., Prieur, C., & Condom, T. (2018). New statistical methods for precipitation bias correction applied to WRF model simulations in the Antisana Region, Ecuador. *Journal of Hydrometeorology*, *19*(12), 2021–2040. https://doi.org/10.1175/JHM-D-18-0032.1

Heredia, M. B., Junquas, C., Prieur, C., Heredia, M. B., Junquas, C., Prieur, C., Condom, T., & Statistical, N. (2019). *New Statistical Methods for Precipitation Bias Correction Applied to WRF Model Simulations in the Antisana Region (Ecuador) To cite this version : HAL Id : hal-01971425*.

Herman, G. R., & Schumacher, R. S. (2016). Extreme Precipitation in Models: An Evaluation. *Weather and Forecasting*, *31*(6), 1853–1879. https://doi.org/10.1175/WAF-D-16-0093.1

Herring, T. A., King, R. W., & McClusky, S. C. (2018). *Introduction to GAMIT / GLOBK Basic framework : GAMIT Basic framework : GLOBK GAMIT Structure* (Issue July 2018, pp. 1–16). Massachusetts Institute of Technology: Cambridge, MA, USA. https://doi.org/http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.534.6776&rep=rep1&type=pdf

Holloway, C. E., & Neelin, J. D. (2009). Moisture Vertical Structure, Column Water Vapor, and Tropical Deep Convection. *Journal of the Atmospheric Sciences*, *66*(6), 1665–1683. https://doi.org/10.1175/2008JAS2806.1

Houze, R. A. (2014). *Cloud Dynamics*. Academic Press, UK. https://www.elsevier.com/books/cloud-dynamics/houze-jr/978-0-12-374266-7

Ilbay-yupa, M., & Barragán, Ricardo Zubieta, Lavado-Casimiro, W. (2019). Regionalización de la precipitación, su agresividad y concentración en la cuenca del Río Guayas, Ecuador. *La Granja; Revista de Ciencias de La Vida.*, *30*(2), 57–76.

Jade, S., & Vijayan, M. S. M. (2008). GPS-based atmospheric precipitable water vapor estimation using meteorological parameters interpolated from NCEP global reanalysis data. *Journal of Geophysical Research Atmospheres*, *113*(3), 1–12. https://doi.org/10.1029/2007JD008758

Jakob Themeßl, M., Gobiet, A., & Leuprecht, A. (2011). Empirical-statistical downscaling and error correction of daily precipitation from regional climate models. *International Journal of Climatology*, *31*(10), 1530–1544. https://doi.org/10.1002/joc.2168

Jauregui, Y. R., & Takahashi, K. (2018a). Simple physical-empirical model of the precipitation distribution based on a tropical sea surface temperature threshold and the effects of climate change. *Climate Dynamics*, *50*(5–6), 2217–2237. https://doi.org/10.1007/s00382-017-3745-3

Jauregui, Y. R., & Takahashi, K. (2018b). Simple physical-empirical model of the precipitation distribution based on a tropical sea surface temperature threshold and the effects of climate change. *Climate Dynamics*, *50*(5–6), 2217–2237. https://doi.org/10.1007/s00382-017-3745-3

Johnson, N. C., & Xie, S. P. (2010). Changes in the sea surface temperature threshold for tropical convection. *Nature Geoscience*, *3*(12), 842–845. https://doi.org/10.1038/ngeo1008

Jonkman, S. N. (2005). Global Perspectives on Loss of Human Life Caused by Floods. *Natural Hazards*, *34*(2), 151–175. https://doi.org/10.1007/s11069-004-8891-3

Junquas, C., Heredia, M. B., Condom, T., Ruiz-Hernández, J. C., Campozano, L., Dudhia, J., Espinoza, J. C., Menegoz, M., Rabatel, A., & Sicart, J. E. (2022). Regional climate modeling of the diurnal cycle of precipitation and associated atmospheric circulation patterns over an Andean glacier region (Antisana, Ecuador). *Climate Dynamics*, *58*(11–12), 3075–3104. https://doi.org/10.1007/s00382-021-06079-y

Junquas, C., Takahashi, K., Condom, T., Espinoza, J. C., Chavez, S., Sicart, J. E., & Lebel, T. (2018). Understanding the influence of orography on the precipitation diurnal cycle and the associated atmospheric processes in the central Andes. *Climate Dynamics*, *50*(11–12), 3995–4017. https://doi.org/10.1007/s00382-017-3858-8

Khairoutdinov, M., & Randall, D. (2006). High-Resolution Simulation of Shallow-to-Deep Convection Transition over Land. *Journal of the Atmospheric Sciences*, *63*(12), 3421–3436. https://doi.org/10.1175/JAS3810.1

Kummerow, C., Hong, Y., Olson, W. S., Yang, S., Adler, R. F., McCollum, J., Ferraro, R., Petty, G., Shin, D.-B., & Wilheit, T. T. (2001). The Evolution of the Goddard Profiling Algorithm (GPROF) for Rainfall Estimation from Passive Microwave Sensors. *Journal of Applied Meteorology*, *40*(11), 1801–1820. https://doi.org/10.1175/1520-0450(2001)040<1801:TEOTGP>2.0.CO;2

Kummerow, C., Olson, W. S., & Giglio, L. (1996). *A Simplified Scheme for Obtaining Precipitation and Vertical Hydrometeor Profiles from Passive Microwave Sensors*. *34*(5), 1213–1232.

Kuo, Y., Neelin, D., & Mechoso, R. (2017). Tropical Convective Transition Statistics and Causality in the Water Vapor – Precipitation Relation. *Journal of the Atmospheric Sciences*, *74*, 915–931. https://doi.org/10.1175/JAS-D-16-0182.1

Labbouz, L., van Baelen, J., & Duroure, C. (2015). Investigation of the links between water vapor field evolution and rain rate based on 5 years of measurements at a midlatitude site. *Geophysical Research Letters*, *42*(21), 9538–9545. https://doi.org/10.1002/2015GL066048

Lanzante, J. R. (1996). Resistant, robust and non-parametric techniques for the analysis of climate data: Theory and examples, including applications to historical radiosonde station data. *International Journal of Climatology*, *16*(11), 1197–1226. https://doi.org/10.1002/(SICI)1097-0088(199611)16:11<1197::AID-JOC89>3.0.CO;2-L

Lenderink, G., & Fowler, H. J. (2017). Hydroclimate: Understanding rainfall extremes. *Nature Climate Change*. https://doi.org/10.1038/nclimate3305

Leon, D. C., French, J. R., Lasher-Trapp, S., Blyth, A. M., Abel, S. J., Ballard, S., Barrett, A., Bennett, L. J., Bower, K., Brooks, B., Brown, P., Charlton-Perez, C., Choularton, T., Clark, P., Collier, C., Crosier, J., Cui, Z., Dey, S., Dufton, D., … Young, G. (2016). The Convective Precipitation Experiment (COPE): Investigating the Origins of Heavy Precipitation in the Southwestern United Kingdom. *Bull. Amer. Meteor. Soc. Bulletin of the American Meteorological Society*, *97*(6), 1003–1020.

Lewis, E., Quinn, N., Blenkinsop, S., Fowler, H. J., Freer, J., Tanguy, M., Hitt, O., Coxon, G., Bates, P., & Woods, R. (2018). A rule based quality control method for hourly rainfall data and a 1 km resolution gridded hourly rainfall dataset for Great Britain: CEH-GEAR1hr. *Journal of Hydrology*, *564*(July), 930–943. https://doi.org/10.1016/j.jhydrol.2018.07.034

Li, H., Wang, X., Wu, S., Zhang, K., Chen, X., Qiu, C., Zhang, S., Zhang, J., Xie, M., & Li, L. (2020). Development of an improved model for prediction of short-term heavy precipitation based on gnss-derived pwv. *Remote Sensing*, *12*(24), 1–22. https://doi.org/10.3390/rs12244101

Li, L., Zhang, K., Wu, S., Li, H., Wang, X., Hu, A., Li, W., Fu, E., Zhang, M., & Shen, Z. (2022). An Improved Method for Rainfall Forecast Based on GNSS-PWV. *Remote Sensing*, *14*(17), 4280. https://doi.org/10.3390/rs14174280

Li, Y., Guan, K., Schnitkey, G. D., DeLucia, E., & Peng, B. (2019). Excessive rainfall leads to maize yield loss of a comparable magnitude to extreme drought in the United States. *Global Change Biology*, *25*(7), 2325–2337. https://doi.org/10.1111/gcb.14628

Lovejoy, S., & Mandelbrot, B. B. (1985). Fractal properties of rain, and a fractal model. *Tellus A*, *37 A*(3), 209–232. https://doi.org/10.1111/j.1600-0870.1985.tb00423.x

Mackern, M. V., Mateo, M. L., Camisay, M. F., & Morichetti, P. v. (2020). Tropospheric products from high-level GNSS processing in Latin America. *In Beyond 100: The Next Century in Geodesy: Proceedings of the IAG General Assembly, Montreal, Canada*, 227–240.

Mallat, S. (1999). *A Wavelet Tour of Signal Processing*.

Manabe, S., Han, D., Holloway, J. (1974). The Seasonal Variation of the Tropical Circulation as Simulated by a Global Model of the Atmosphere. *Journal of the Atmospheric Sciences*, *31*.

Mapes, B., Warner, T., & Xu, M. (2003). Diurnal patterns of rainfall in northwestern South America. Part II: Model simulations. *Monthly Weather Review*, *131*(5), 813–829. https://doi.org/10.1175/1520-0493(2003)131<0813:DPORIN>2.0.CO;2

Martinez-Villalobos, C., & Neelin, J. D. (2018). Shifts in Precipitation Accumulation Extremes During the Warm Season Over the United States. *Geophysical Research Letters*, *45*(16), 8586–8595. https://doi.org/10.1029/2018GL078465

Mathias Vuille, Raymond S. Bradley, and F. K. (2000). Climate Variability in the Andes of Ecuador and Its Relation to Tropical Pacific and Atlantic Sea Surface Temperature Anomalies. *Climate, Journal Of*, *Hastenrath 1981*, 2520–2535.

Mathworks. (2021). *Choose a wavelet*. Philadelphia, PA, USA,.

Meyer, Y. (1995). *Wavelets and Operators* (N. Y. C. U. Press., Ed.). The Mathematical Gazette, The Mathematical Association: Leicester, UK,. https://doi.org/https://doi.org/10.2307/3620106

Meza, A., Mendoza, L., Natali, M. P., Bianchi, C., & Fernández, L. (2020). Diurnal variation of precipitable water vapor over Central and South America. *Geodesy and Geodynamics*, *11*(6), 426–441. https://doi.org/10.1016/j.geog.2020.04.005

Mourre, L., Condom, T., Junquas, C., Lebel, T., E. Sicart, J., Figueroa, R., & Cochachin, A. (2016). Spatio-temporal assessment of WRF, TRMM and in situ precipitation data in a tropical mountain environment (Cordillera Blanca, Peru). *Hydrology and Earth System Sciences*, *20*(1), 125–141. https://doi.org/10.5194/hess-20-125-2016

NASDA. (2001). *TRMM Data Users Handbook*.

Nathan, R., Jordan, P., Scorah, M., Lang, S., Kuczera, G., Schaefer, M., & Weinmann, E. (2016). Estimating the exceedance probability of extreme rainfalls up to the probable maximum precipitation. *Journal of Hydrology*, *543*, 706–720. https://doi.org/10.1016/j.jhydrol.2016.10.044

Neelin, J. D., Peters, O., & Hales, K. (2009a). The transition to strong convection. *Journal of the Atmospheric Sciences*, *66*(8), 2367–2384. https://doi.org/10.1175/2009JAS2962.1

Neelin, J. D., Peters, O., & Hales, K. (2009b). The transition to strong convection. *Journal of the Atmospheric Sciences*, *66*(8), 2367–2384. https://doi.org/10.1175/2009JAS2962.1

Newman, M. E. J. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary Physics*, *46*(5), 323–351. https://doi.org/10.1080/00107510500052444

Nielsen, E. R., & Schumacher, R. S. (2016). Using Convection-Allowing Ensembles to Understand the Predictability of an Extreme Rainfall Event. *Mon. Wea. Rev. Monthly Weather Review*, *144*(10), 3651–3676.

Olson, W. S., Kummerow, C. D., Hong, Y., & Tao, W. K. (1999). Atmospheric latent heating distributions in the tropics derived from satellite passive microwave radiometer measurements. *Journal of Applied Meteorology*, *38*(6), 633–664. https://doi.org/10.1175/1520-0450(1999)038<0633:ALHDIT>2.0.CO;2

Padullés, R., Kuo, Y.-H., Neelin, J. D., Turk, F. J., Ao, C. O., & de la Torre Juárez, M. (2022). Global tropical precipitation relationships to free tropospheric water vapor using Radio Occultations. *Journal of the Atmospheric Sciences*, 1–46. https://doi.org/10.1175/jas-d-21-0052.1

Panegrossi, G., Rysman, J. F., Casella, D., Marra, A. C., Sanò, P., & Kulie, M. S. (2017). CloudSat-based assessment of GPM microwave imager snowfall observation capabilities. *Remote Sensing*, *9*(12). https://doi.org/10.3390/rs9121263

Park, K. I. (2018). *Fundamentals of Probability and Stochastic Processes with Applications to Communications* (Springer, Ed.). Berlin/Heidelberg, Germany. https://doi.org/doi:https://doi.org/10.1007/978-3-319-68075-0

Pendergrass, A. G. (2018). What precipitation is extreme? *Science*, *360*(6393). https://doi.org/10.1126/science.aat1871

Percival, D. B., & Walden, A. T. (2000). Wavelet Methods for Time SeriesAnalysis. In U. Cambridge (Ed.), *Wavelet Methods for Time SeriesAnalysis*. Cambridge University Press, Hoboken, NJ, USA,. https://doi.org/10.1017/cbo9780511841040

Peters, O., Hertlein, C., Christensen, K., & Hertlein, C. (2002). A Complexity View of Rainfall. *Physical Review Letters*, *88*(1), 4. https://doi.org/10.1103/PhysRevLett.88.018701

Peters, O., & Neelin, J. D. (2006). Critical phenomena in atmospheric precipitation. *Nature Physics*, *2*(6), 393–396. https://doi.org/10.1038/nphys314

Pielke, R. A., Wilby, R., Niyogi, D., Hossain, F., Dairuku, K., Adegoke, J., Kallos, G., Seastedt, T., & Suding, K. (2013). Dealing with Complexity and Extreme Events Using a Bottom-Up, Resource-Based Vulnerability Perspective. *Extreme Events and Natural Hazards: The Complexity Perspective*, 345–359. https://doi.org/10.1029/2011GM001086

Porras, L., Romero, R., Amores, M., Estrella, C., Tierra, A., & Zavala, M. (2016). Diferencias obtenidas del contenido de vapor de agua en ciertas regiones del Ecuador entre productos satelitales y valores calculados de las estaciones meteorológicas de la REGME. *Simposio SIRGAS 2016 Ecuador y Taller SIRGAS (GT-III)*.

Qian, X., Yao, Y., Wang, H., Zou, L., Li, Y., & Yin, J. (2020). Validation of the WRF model for estimating precipitable water vapor at the ali observatory on the Tibetan plateau. *Publications of the Astronomical Society of the Pacific*, *132*(1018), 1–10. https://doi.org/10.1088/1538-3873/abc22d

María Sheila Fabiola Serrano Vincenti

Risanto, C. B. ;, Castro, C. L. ;, Arellano, A. F., Jr. ;, Moker, J. M., Jr. ;, & Adams, D. K. (2021). The Impact of Assimilating GPS Precipitable Water Vapor in Convective-Permitting WRF-ARW on North American Monsoon Precipitation Forecasts over Northwest Mexico. *Monthly Weather Review*, *149*(9), 3013–3035. https://doi.org/https://doi.org/10.1175/MWR-D-20-0394.1

Roderick, T. P., Wasko, C., & Sharma, A. (2019). Atmospheric Moisture Measurements Explain Increases in Tropical Rainfall Extremes. *Geophysical Research Letters*, 1375–1382. https://doi.org/10.1029/2018GL080833

Romero, R., Pilapanta, C., Porras, L., & Tierra, A. (2019). Prediction of Precipitable Water Vapor With a Neural Network From the Ecuadorian Gnss and Meteorological Data. *Revista Geoespacial*, *15*(2), 1. https://doi.org/10.24133/geoespacial.v15i2.1349

Rösch, A., & Schmidbauer, H. (2018). *WaveletComp: Computational Wavelet Analysis. R package version 1.1.* (pp. 1–38). R core Team: Vienna, Austria. https://doi.org/http://www.hs-stat.com/projects/WaveletComp/WaveletComp_guided_tour.pdf%0Ahttps://cran.r-project.org/package=WaveletComp

Ruiz-Hernández, J. C., Condom, T., Ribstein, P., le Moine, N., Espinoza, J. C., Junquas, C., Villacís, M., Vera, A., Muñoz, T., Maisincho, L., Campozano, L., Rabatel, A., & Sicart, J. E. (2021). Spatial variability of diurnal to seasonal cycles of precipitation from a high-altitude equatorial Andean valley to the Amazon Basin. *Journal of Hydrology: Regional Studies*, *38*(October). https://doi.org/10.1016/j.ejrh.2021.100924

Sajith, V., Adegoke, J. O., Raghavan, S. K., Mohan, R., Vinod, K., & Preenub, P. N. (2007). Evaluation of daily and diurnal signals of total precipitable water (TPW) over the Indian Ocean based on TMI retrieved 3-day composite estimates and radiosonde data. *International Journal of Climatology*, *27*(November 2006), 761–770. https://doi.org/10.1002/joc

Sapucci, L. F., Machado, L. A. T., de Souza, E. M., & Campos, T. B. (2019). Global Positioning System precipitable water vapour (GPS-PWV) jumps before intense rain events: A potential application to nowcasting. *Meteorological Applications*, *26*(1), 49–63. https://doi.org/10.1002/met.1735

Schroeder, A., Basara, J., Marshall Shepherd, J., & Nelson, S. (2016). Insights into atmospheric contributors to urban flash flooding across the United States using an analysis of rawinsonde data and associated calculated parameters. *Journal of Applied Meteorology and Climatology*, *55*(2), 313–323. https://doi.org/10.1175/JAMC-D-14-0232.1

Schumacher, R. (2016). *Heavy Rainfall and Flash Flooding* (Vol. 1, Issue September 2018). https://doi.org/10.1093/acrefore/9780199389407.013.132

Segura, H., Espinoza, J. C., Junquas, C., Lebel, T., Vuille, M., & Garreaud, R. (2020). Recent changes in the precipitation-driving processes over the southern tropical Andes/western Amazon. *Climate Dynamics*, *54*(5–6), 2613–2631. https://doi.org/10.1007/s00382-020-05132-6

Serrano S. (2016). *Fenómenos Críticos En Datos De Precipitación Lluviosa Intensa Detectados Con Radar Y Microondas , En La Zona De Influencia Del Fenómeno Del Niño Sobre El Ecuador ,* [Escuela Politécnica Nacional]. http://bibdigital.epn.edu.ec/handle/15000/14107

Serrano-Vincenti, S., Condom, T., Campozano, L., Escobar, L. A., Walpersdorf, A., Carchipulla-Morales, D., & Villacís, M. (2022). *Harmonic Analysis of the Relationship between GNSS Precipitable Water Vapor and Heavy Rainfall over the Northwest Equatorial Coast, Andes, and Amazon Regions. 13,* 1809. https://doi.org/10.3390/atmos13111809

Serrano-Vincenti, S., Condom, T., Campozano, L., Guamán, J., & Villacís, M. (2020). An Empirical Model for Rainfall Maximums Conditioned to Tropospheric Water Vapor Over the Eastern Pacific Ocean. *Frontiers in Earth Science*, *8*(July), 1–12. https://doi.org/10.3389/feart.2020.00198

Sguerso, D., Labbouz, L., & Walpersdorf, A. (2016). 14 years of GPS tropospheric delays in the French–Italian border region: comparisons and first application in a case study. *Applied Geomatics*, *8*(1), 13–25. https://doi.org/10.1007/s12518-015-0158-z

Shoji, Y., Kunii, M., & Saito, K. (2009). Assimilation of Nationwide and Global GPS PWV Data for a Heavy Rain Event on 28 July 2008 in Hokuriku and Kinki, Japan. *SOLA*, *5*, 45–48. https://doi.org/10.2151/sola.2009-012

Stoica, P., & Moses, R. (2005). Spectral Analysis of Signals [Book Review]. In *IEEE Signal Processing Magazine* (Vol. 24, Issue 1). Prentice Hall, Inc. https://doi.org/10.1109/msp.2007.273066

Takahashi, K., & Dewitte, B. (2016). Strong and moderate nonlinear El Niño regimes. *Climate Dynamics*, *46*(5–6), 1627–1645. https://doi.org/10.1007/s00382-015-2665-3

Takahashi, K., Montecinos, A., Goubanova, K., & Dewitte, B. (2011). ENSO regimes: Reinterpreting the canonical and Modoki El Nio. *Geophysical Research Letters*, *38*(10), 1–5. https://doi.org/10.1029/2011GL047364

Torrence, C. (1999). Interdecadal Changes in the ENSO-Monsoon System. Journal of Climate. *Journal of Climate*, *12*(8), 2679–2690. https://doi.org/https://doi.org/10.1175/1520-0442(1999)012<2679:ICITEM>2.0.CO;2

Torrence, C., & Compo, G. (1998). *A Practical Guide toWavelet Analysis*.

Torri, G., Adams, D. K., Wang, H., & Kuang, Z. (2019). On the diurnal cycle of GPS-derived precipitable water vapor over sumatra. *Journal of the Atmospheric Sciences*, *76*(11), 3529–3552. https://doi.org/10.1175/JAS-D-19-0094.1

Vargas, D., Pucha-Cofrep, D., Serrano-Vincenti, S., Burneo, A., Carlosama, L., Herrera, M., Cerna, M., Molnár, M., Jull, A. J. T., Temovski, M., László, E., Futó, I., Horváth, A., & Palcsu, L. (2022). ITCZ precipitation and cloud cover excursions control Cedrela nebulosa tree-ring oxygen and carbon isotopes in the northwestern Amazon. *Global and Planetary Change*, *211*(April 2021). https://doi.org/10.1016/j.gloplacha.2022.103791

Villacís, M., Vimeux, F., & Taupin, J. D. (2008). Analysis of the climate controls on the isotopic composition of precipitation (δ18O) at Nuevo Rocafuerte, 74.5°W, 0.9°S, 250 m, Ecuador. *Comptes Rendus - Geoscience*, *340*(1), 1–9. https://doi.org/10.1016/j.crte.2007.11.003

Vincent, E. M., Lengaigne, M., Madec, G., Vialard, J., Samson, G., Jourdain, N. C., Menkes, C. E., & Jullien, S. (2012). Processes setting the characteristics of sea surface cooling induced by tropical cyclones. *Journal of Geophysical Research: Oceans*, *117*(2). https://doi.org/10.1029/2011JC007396

Wagnon, P., Lafaysse, M., Lejeune, Y., Maisincho, L., Rojas, M., & Chazarin, J. P. (2009). Understanding and modeling the physical processes that govern the melting of snow cover in a tropical mountain environment in Ecuador. *Journal of Geophysical Research Atmospheres*, *114*(19), 1–14. https://doi.org/10.1029/2009JD012292

Walpersdorf, A., Bouin, M., Bock, O., & Doerflinger, E. (2007). Assessment of GPS data for meteorological applications over Africa : Study of error sources and analysis of positioning accuracy. *Journal of Atmospheric and Solar-Terrestrial Physics*, *69*, 1312–1330. https://doi.org/10.1016/j.jastp.2007.04.008

Wang, N. Y., Liu, C., Ferraro, R., Wolff, D., Zipser, E., & Kummerow, C. (2009). Trmm 2A12 land precipitation product - Status and future plans. *Journal of the Meteorological Society of Japan*, *87 A*, 237–253. https://doi.org/10.2151/jmsj.87A.237

Warner, T. (2010). *Numerical Weather and Climate Prediction.* (C. U. Press., Ed.). Cambridge. https://doi.org/doi:10.1017/CBO9780511763243

Wentz, F. J., Ashcroft, P., & Gentemann, C. (2001). Post-launch calibration of the TRMM microwave imager. *IEEE Transactions on Geoscience and Remote Sensing*, *39*(2), 415–422. https://doi.org/10.1109/36.905249

Wilcox, E. M., & Donner, L. J. (2007). The frequency of extreme rain events in satellite rain-rate estimates and an atmospheric general circulation model. *Journal of Climate*, *20*(1), 53–69. https://doi.org/10.1175/JCLI3987.1

Wu, C.-M., Stevens, B., & Arakawa, A. (2009). What Controls the Transition from Shallow to Deep Convection? *Journal of the Atmospheric Sciences*. https://doi.org/10.1175/2008JAS2945.1

Yang, M. ;, Yao, T. ;, Gou, X. ;, Wang, H. ;, & Neumann, T. (2006). Wavelet analysis reveals periodic oscillations in a 1700 year ice-core record from Guliya. *Annals of Glaciology*, *43*(1), 132–136. https://doi.org/10.3189/172756406781812375

Yao, Y., Shan, L., & Zhao, Q. (2017). Establishing a method of short-term rainfall forecasting based on GNSS-derived PWV and its application. *Scientific Reports*, *7*(1), 1–11. https://doi.org/10.1038/s41598-017-12593-z

Yeh, T. K., Hong, J. S., Wang, C. S., Chen, C. H., Chen, K. H., & Fong, C. T. (2016). Determining the precipitable water vapor with ground-based GPS and comparing its

yearly variation to rainfall over Taiwan. *Advances in Space Research*, *57*(12), 2496–2507. https://doi.org/10.1016/j.asr.2016.04.002

Yepes, J., Mejía, J. F., Mapes, B., & Poveda, G. (2020). Gravity Waves and Other Mechanisms Modulating the Diurnal Precipitation over One of the Rainiest Spots on Earth: Observations and Simulations in 2016. *Monthly Weather Review*, *148*(9), 3933–3950. https://doi.org/10.1175/MWR-D-19-0405.1

Zhao, Q., Ma, X., & Yao, Y. (2020). Preliminary result of capturing the signature of heavy rainfall events using the 2-d-/4-d water vapour information derived from GNSS measurement in Hong Kong. *Advances in Space Research*, *66*(7), 1537–1550. https://doi.org/10.1016/j.asr.2020.06.013

Appendix

Appendix A

Characteristic periods of (a) PWV. Note the marked period of 24 h and the secondary one of 12. In the Amazon region, the spectral density is lower and there are two more characteristic periods of 8 and 16 h (b) rain. Note the lower spectral density in comparison with PWV, and the characteristic period of around 24 h.
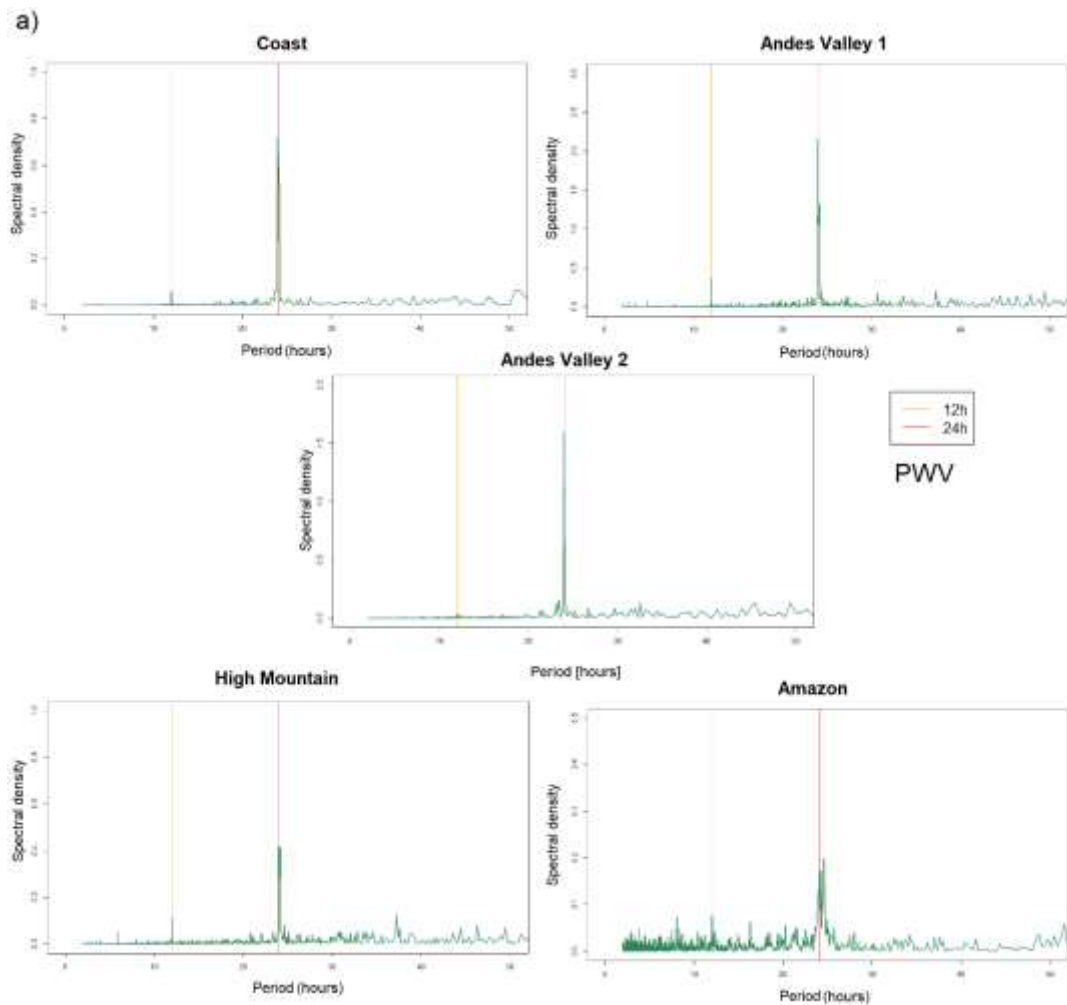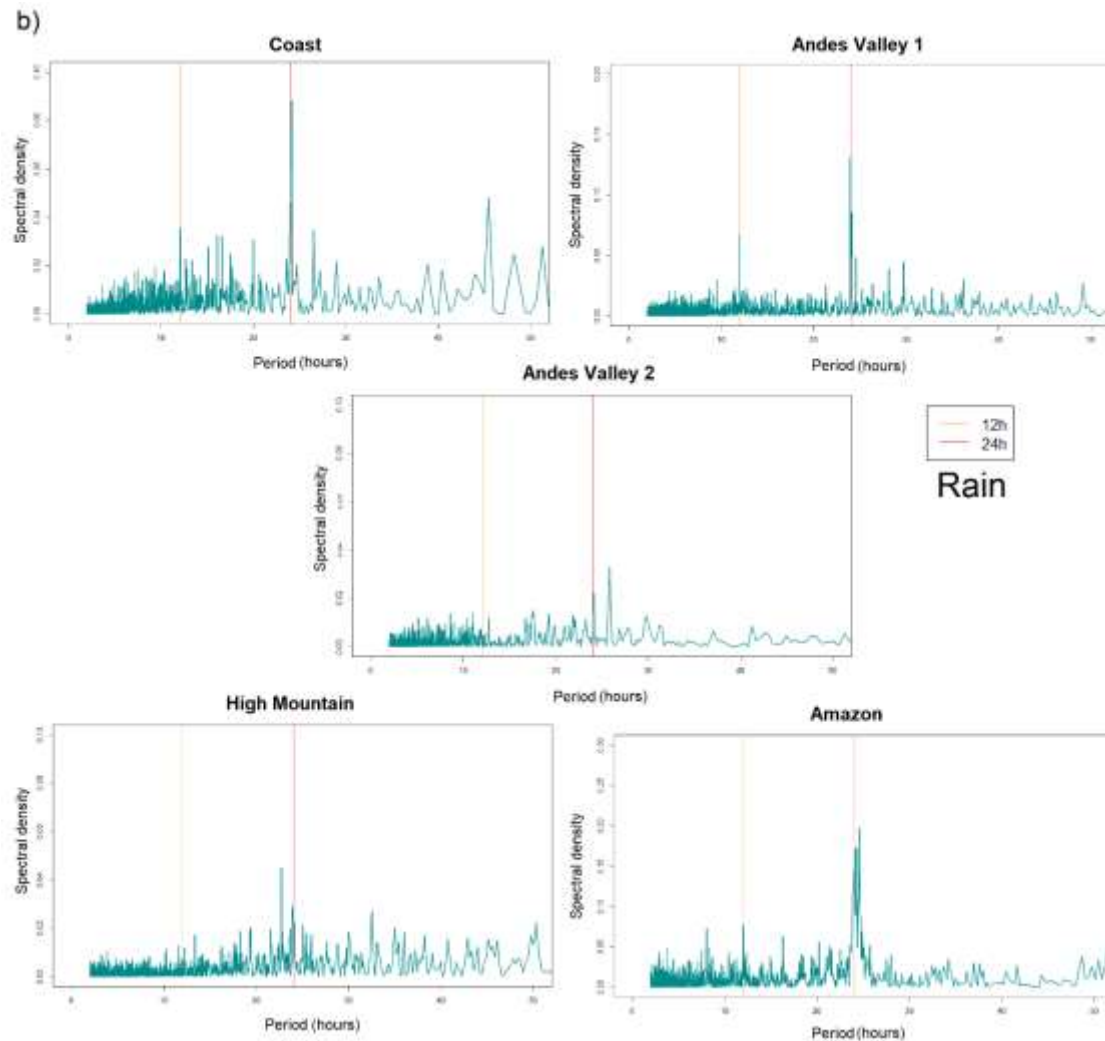
Figure A.1. Characteristic periods of (a) PWV. Note the marked period of 24 h and the secondary one of 12. In the Amazon region, the spectral density is lower and there are two more characteristic periods of 8 and 16 h (b) rain. Note the lower spectral density in comparison with PWV, and the characteristic period of around 24 h.

Appendix B

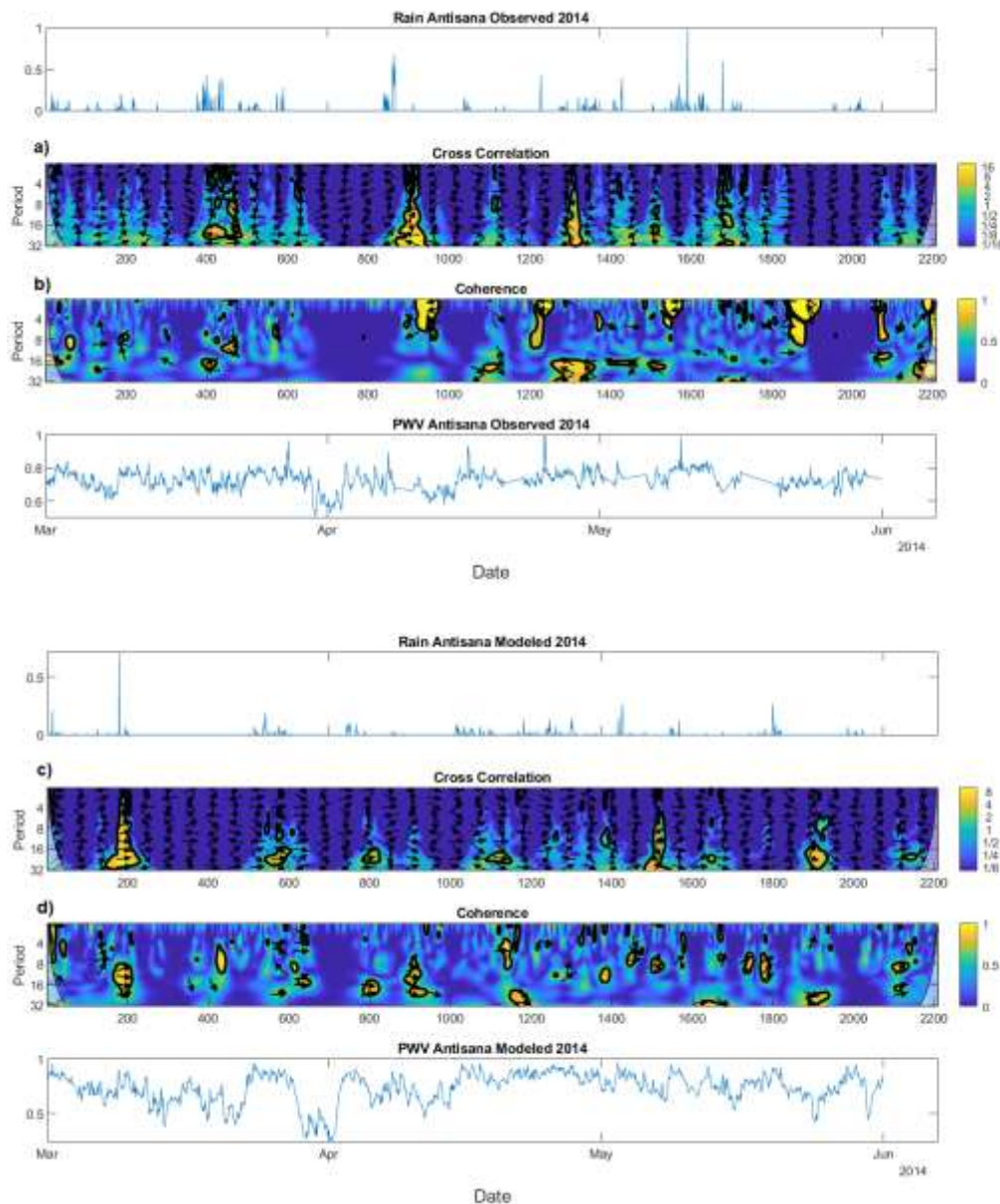Wavelet Cross-Spectrum (XWT) and Wavelet Transform Coherence (WTC) of the stations for the year 2014 for the Coast, Andes Valley 1, Andes Valley 2, High Mountain and Amazon.
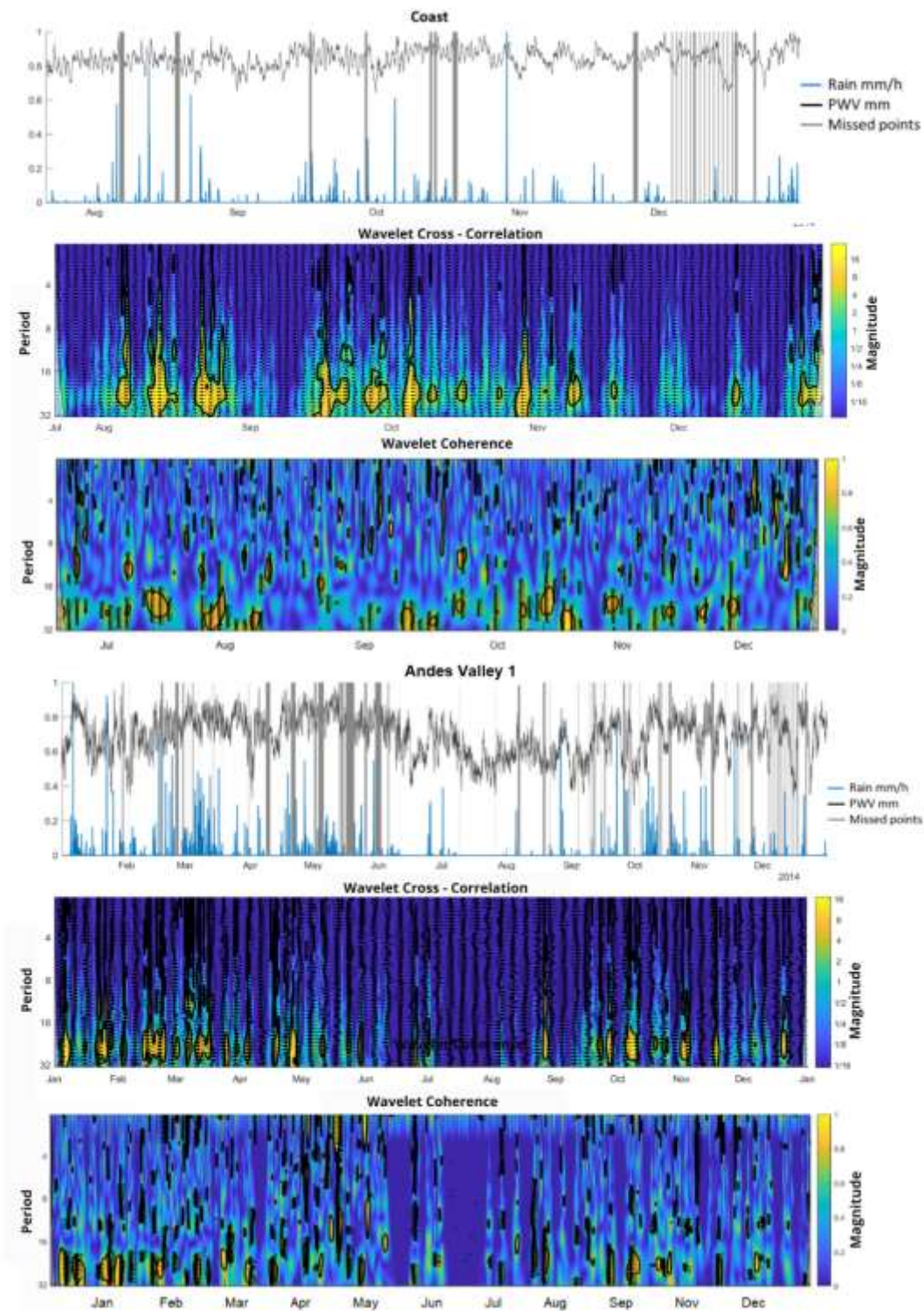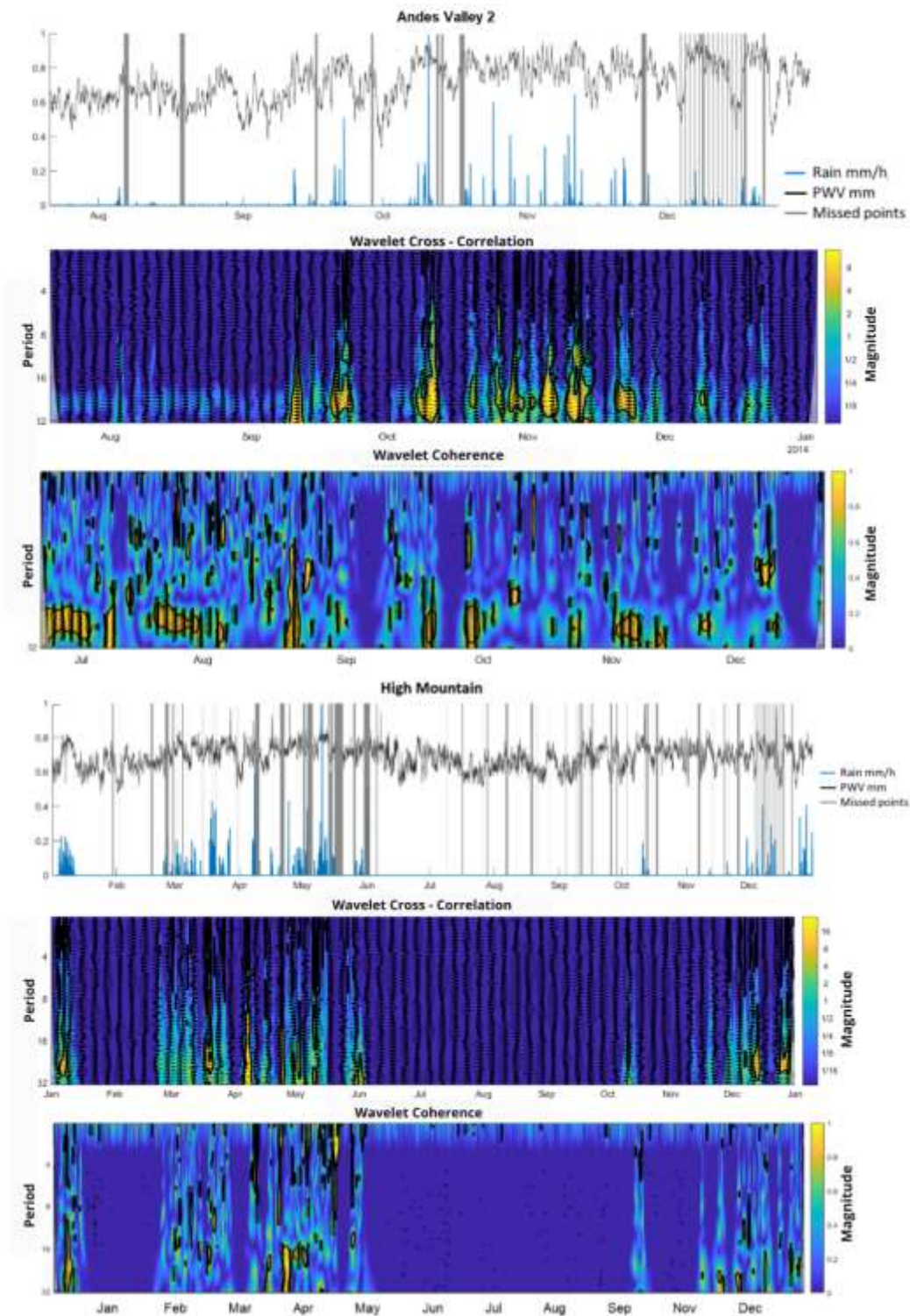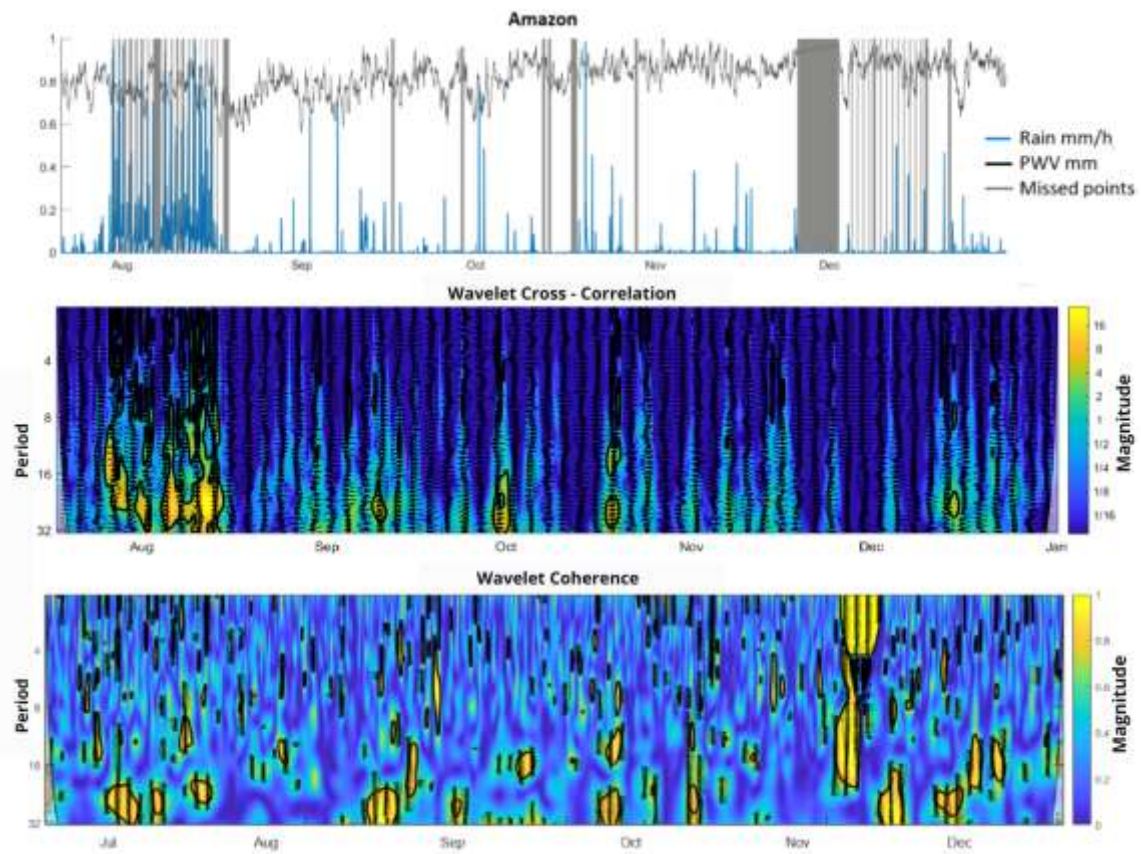
Appendix C

Continuous wavelet cross-spectrum XWT and wavelet coherence scalogram for rain and PWV in HV1 for entire year 2014. (a) XWT observed, (b) observed coherence, (d) XWT modeled (d) modeled coherence.
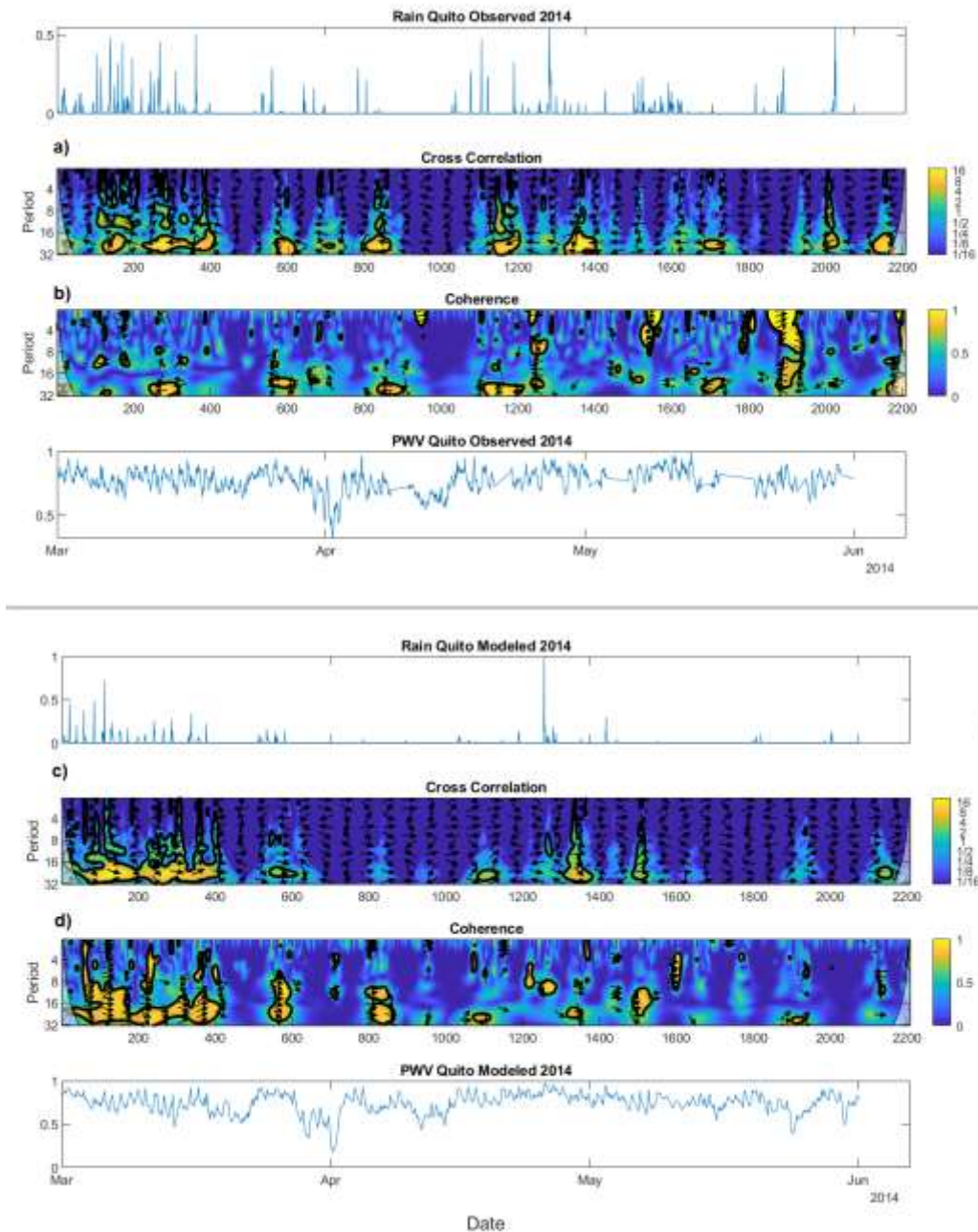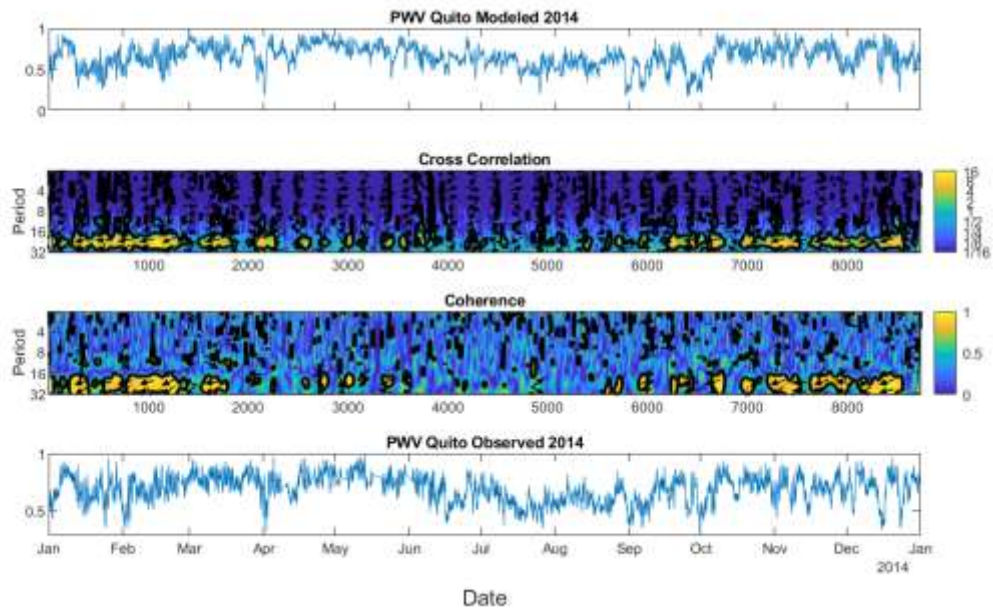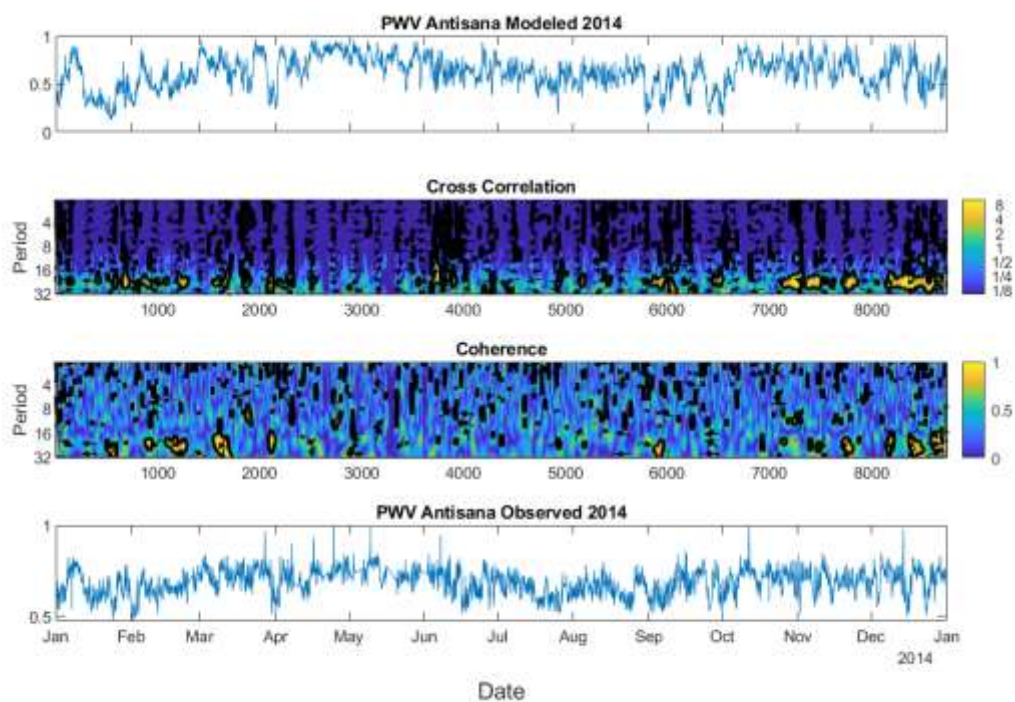
Appendix D

Continuous wavelet cross-spectrum XWT and wavelet coherence scalogram of rain and PWV in HMA for entire year 2014. (a) XWT observed, (b) observed coherence, (d) XWT modeled (d) modeled coherence.
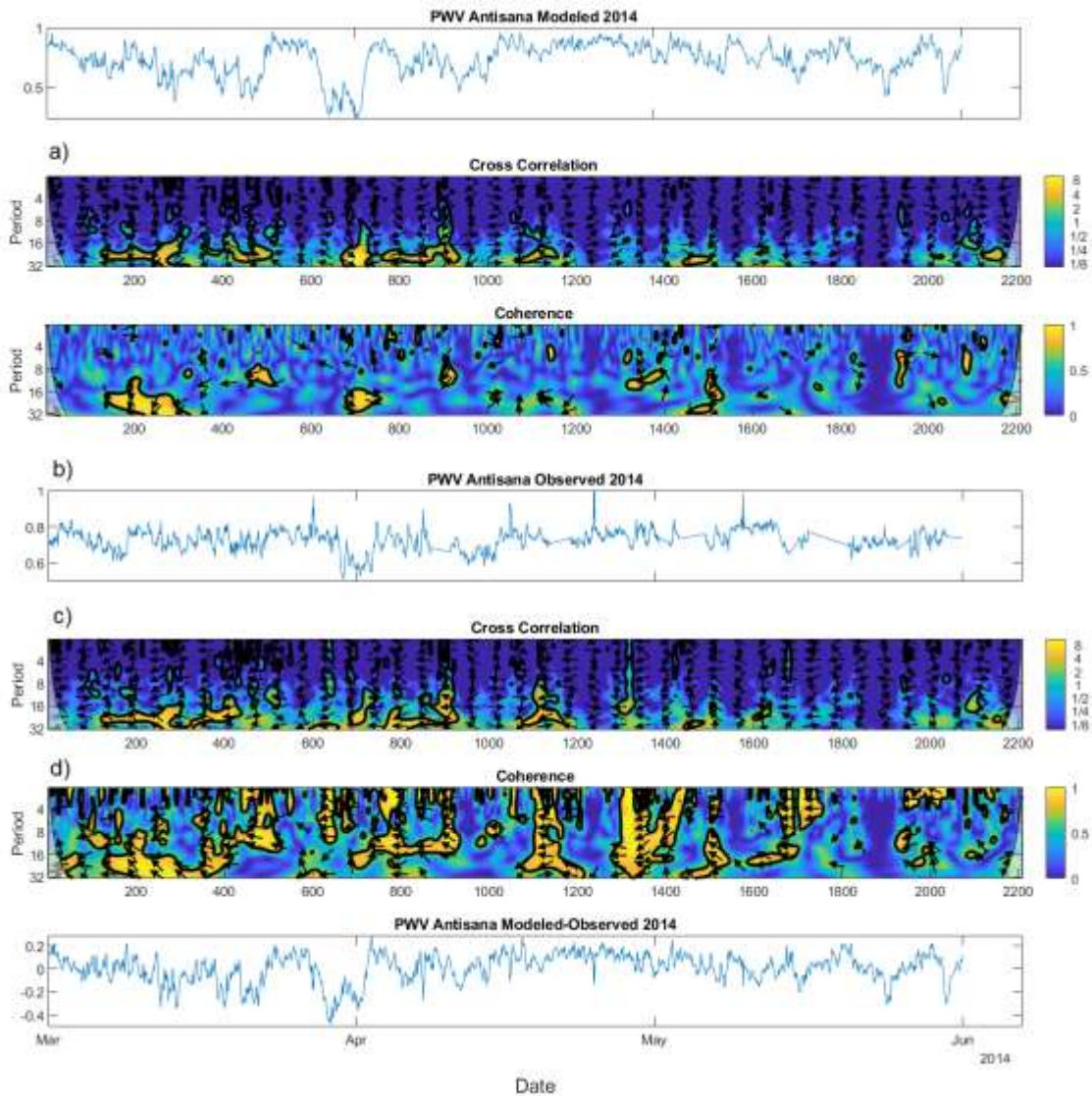

Appendix E

Continuous wavelet cross-spectrum XWT and wavelet coherence scalogram of PWV in HV1 for entire year 2014. (a) (b) between Modeled and Observed
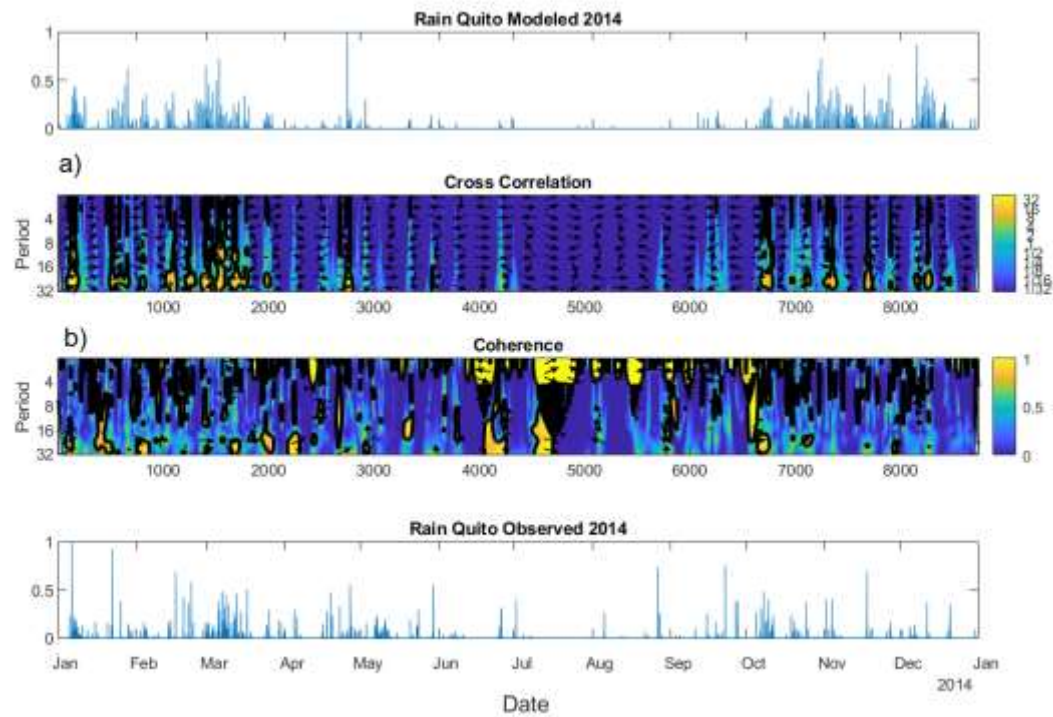
Appendix F

Continuous wavelet cross-spectrum XWT and wavelet coherence scalogram of PWV in HMA for entire year 2014. (a) (b) between Modeled and Observed.
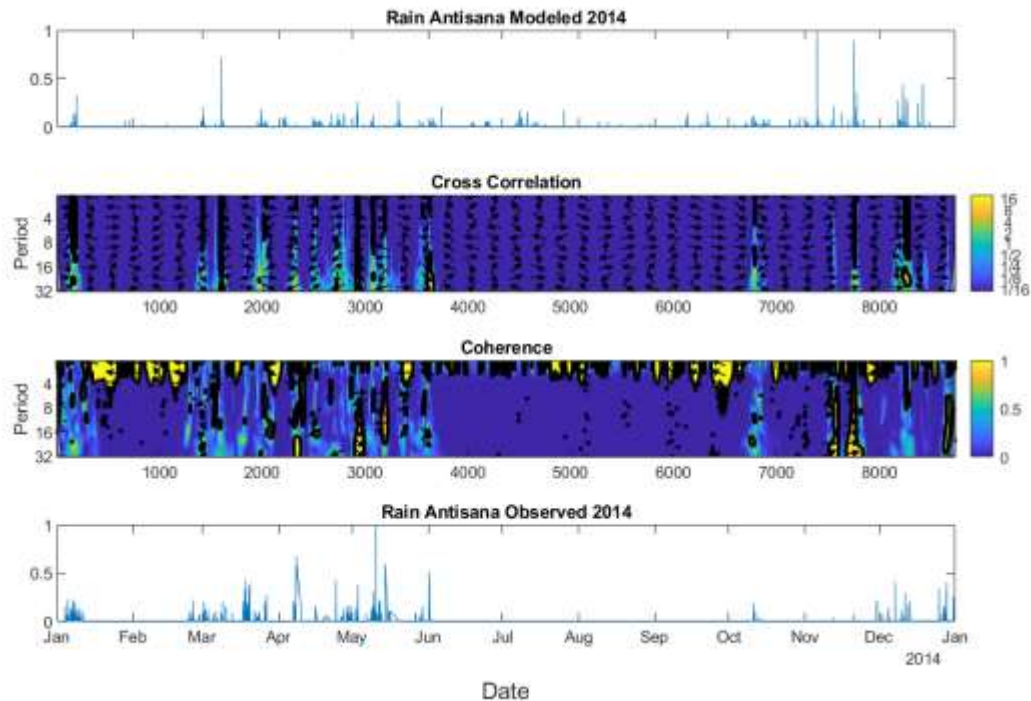
Appendix G

Continuous wavelet cross-spectrum XWT and wavelet coherence scalogram of Precipitation in HV1 for entire year 2014. (a) (b) between Modeled and Observed.

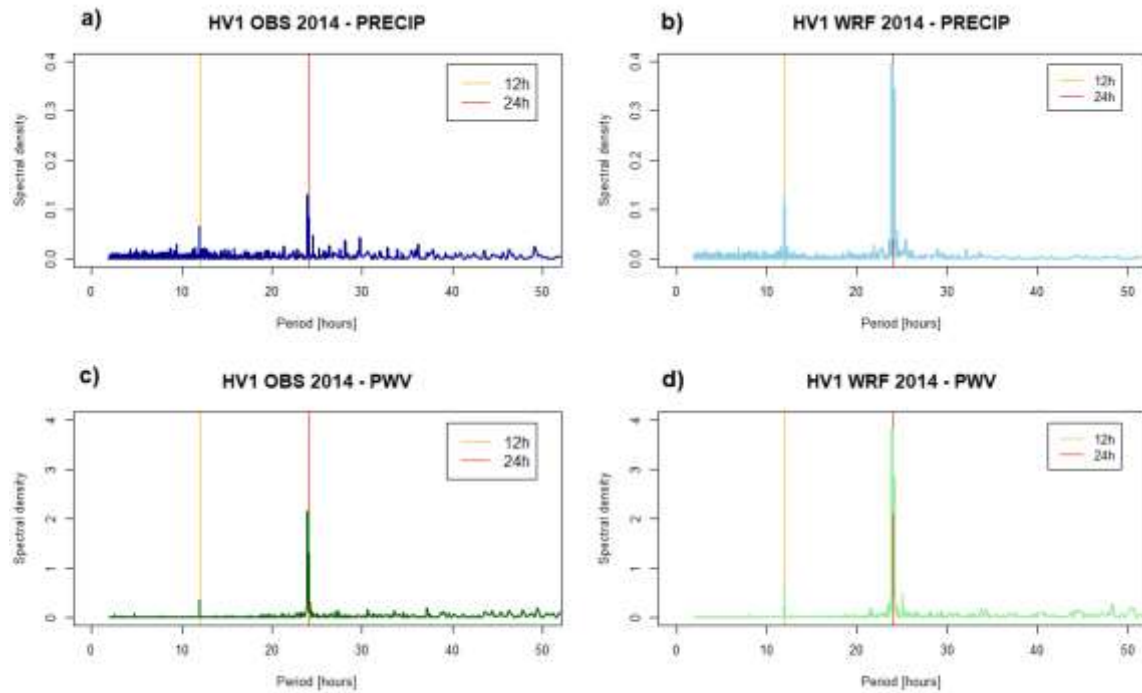María Sheila Fabiola Serrano Vincenti

Appendix H

Continuous wavelet cross-spectrum XWT and wavelet coherence scalogram of Precipitation in HMA for entire year 2014. (a) (b) between Modeled and Observed.
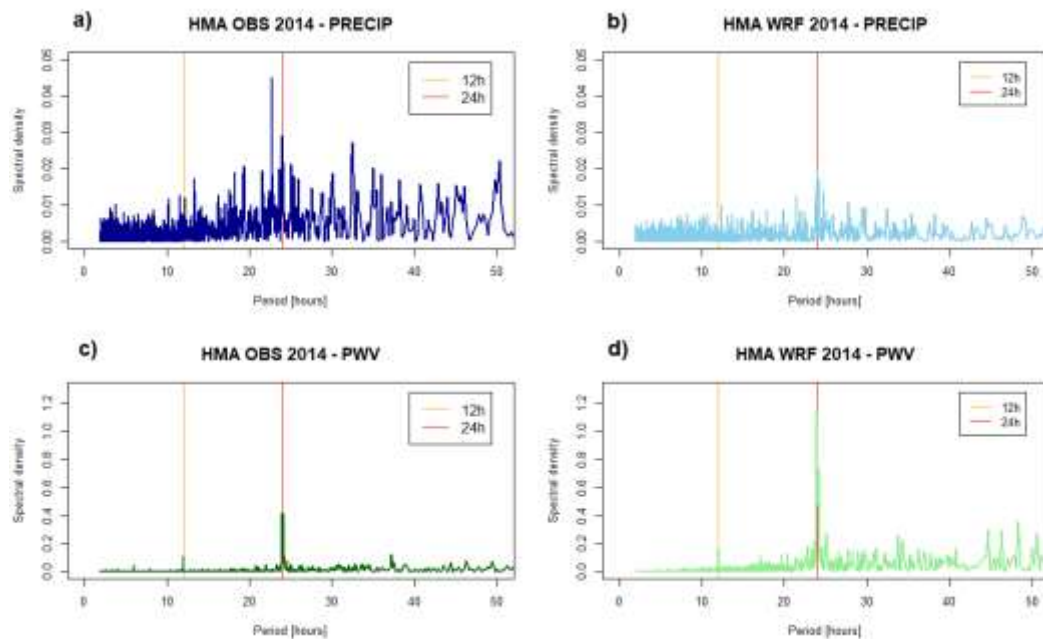

Appendix I

Fourier series characteristic periods of PWV a) Observed b) Modeled, and rainfall for d) Observed and d) Modeled series for Quito HV1.

Appendix J

Fourier series characteristic periods of PWV a) Observed b) Modeled, and rainfall for d) Observed and d) Modeled series for Antisana HMA.

About the author

María Sheila Fabiola Serrano Vincenti was born in Cochabamba, Bolivia on April 28th, 1978. She graduated as Physicist and as Master in Physics, from the Physics Department at National Polytechnic School in Ecuador. Editor in Chief of the Life Sciences Journal, La Granja. Director of the Environmental Modeling Research Center CIMA-UPS. Associate Professor at the Salesian Polytechnic University since 2002. The main scientific interests are intense rainfall, extreme atmospheric precipitation events and Climate Change.

Peer Review Publications:

Serrano-Vincenti, S., Condom, T., Campozano, L., Escobar, L. A., Walpersdorf, A., Carchipulla-Morales, D., & Villacís, M. (2022). Harmonic Analysis of the Relationship between GNSS Precipitable Water Vapor and Heavy Rainfall over the Northwest Equatorial Coast, Andes, and Amazon Regions. 13, 1809. https://doi.org/10.3390/atmos13111809

Serrano-Vincenti, S., Condom, T., Campozano, L., Guamán, J., & Villacís, M. (2020). An Empirical Model for Rainfall Maximums Conditioned to Tropospheric Water Vapor Over the Eastern Pacific Ocean. Frontiers in Earth Science, 8(July), 1–12. https://doi.org/10.3389/feart.2020.00198

Vargas D.; Pucha D.; Serrano S.; Burneo A.; Carlosama L.; Herrera M.; Cerna M.; Molnár N.; Jull T.; Temovski M.; László E.; Futó I.; Horváth A.; Palcsu L. (2022). ITCZ precipitation and cloud cover excursions control Cedrela nebulosa tree-ring oxygen and carbon isotopes in the northwestern Amazon, Global and Planetary Change. Vol (211), 103791. ISSN 0921-8181 https://www.sciencedirect.com/science/article/pii/S0921818122000583

Montenegro M.; Campozano L.; Urdiales D.; Maisincho L.; Serrano S.; Borbor M. (2021). Assessment of the impact of higher temperatures in the mortality risk indexes in Ecuador until 2070 due to climate change, Frontiers in Earth Science, section Atmospheric Science 10.3389/feart.2021.794602

Serrano S.; Reisancho A.; Borbor M.; Stewart A. (2016). Análisis de inundaciones costeras por precipitaciones intensas, cambio climático y Fenómeno de el Niño. Caso de

estudio: Machala.   La   Granja   Vol.   24   (2).
https://revistas.ups.edu.ec/index.php/granja/article/view/24.2016.05

Serrano S.; Vásquez N.; Jácome P.; Basile L. (2014). Critical Phenomena of Rainfall in
Ecuador. Sun and Geosphere. 9(1):73-76.
http://newserver.stil.bas.bg/SUNGEO/00SGArhiv/SG_v9_No1_2014-pp-75-78.pdf

Serrano, S.; Ruiz J.; Bersosa F. (2017). Heavy rainfall and temperature proyections in a
climate change scenario over Quito, Ecuador. La Granja: Revista de Ciencias de la
Vida. Vol. 25(1):16- 32. pISSN:1390-3799; eISSN:1390-8596.
https://dspace.ups.edu.ec/bitstream/123456789/13875/1/Lgr_n25_Serrano_Ru%c3%
adz_B ersosa.pdf

Serrano, S.; Zuleta D.; Moscoso V.; Jácome P.; Palacios E.; Villacís M. (2012). Análisis
estadístico de datos meteorológicos mensuales y diarios para la determinación de
variabilidad climática y cambio climático en el Distrito Metropolitano de Quito. La
Granja.   Vol.   16(2):   23-47.   ISSN:1390-3799.
https://www.researchgate.net/publication/236235036_Statistical_analysis_of_daily_a
nd_mo
nthly_meteorological_data_of_the_Metropolitan_District_of_Quito_for_weather_varia
bility_and _climate_change_studies

Serrano, S.; Basile L. (2012). La precipitación intensa vista desde la criticalidad
autoorganizada y las transiciones de fase continuas: un nuevo enfoque de estudio.
La Granja. Vol. 15(1): 5-18. ISSN: 1390-3799.
https://www.researchgate.net/publication/236235061_Intense_precipitation_since_se
lforganized_criticality_and_continuous_phase_transitionsa_new_aproach

Doctoral courses

1. Introduction to Database Management for climate analysis, 2017.

2. Numerical methods, 2017.

3. Scientific writing, 2017.

4. Water Resources Planning and Management, 2017.

5. Scientific oral and poster presentations, 2017.

6. Mathematics for Data Analysis and Processing with Applications in Environmental Science