



**UNIVERSIDAD TÉCNICA DEL NORTE**  
**FACULTAD DE POSTGRADO**

**Maestría en Ciencias de la Ingeniería para la Gestión de los Recursos  
Hídricos**

“Modelos basados en datos para estimar humedad del suelo y precipitación en los Andes  
tropicales del norte de Ecuador”

Trabajo de investigación previo a la obtención del título de Magister en Ciencias de la  
Ingeniería para la Gestión de Recursos Hídricos

**AUTOR:** Mat. Diego Polivio Escobar González

**TUTOR:** Ing. Marcos Joshua Villacis Erazo PhD.  
**CO-TUTOR:** Ing. Gabriel Alexis Jácome Aguirre MSc.

**Ibarra-ECUADOR**

**2024**

## **ParamoSUS**

The research presented in this thesis is part of the ParamoSUS project, the international research collaboration funded by ARES CCD (2017-2022). It is a partnership between researchers from Belgium (UCLouvain and UNamur) and Ecuador (Escuela Politécnica Nacional and Universidad de Cuenca). It also involves close collaboration with Ecuadorian stakeholders including Fondo para la Protección del Agua (FONAG), Empresa Publica Metropolitana de Agua Potable y Saneamiento de Quito (EPMAPS), Empresa Pública Municipal de Agua Potable, Alcantarillado, Telecomunicaciones de Cuenca (ETAPA) and Electro Generadora del Austro (ELECAUSTRO). The ParamoSUS project aims to contribute to sustainable ecosystem management by producing evidence-based research on the status, conservation, and restoration of páramo soils, soil-water-vegetation interactions, and soil hydrological regulation in the high Andes. For its implementation, the members of the project selected two soil observatories: (i) the Jatunhuayco headwater basin located in the northern Andes near Ecuador's capital, Quito, and (ii) the Machangara and Yanuncay headwater basins located in the southern Ecuadorian Andes, close to the city of Cuenca. The main study object of this thesis is the soil observatory located in northern Ecuador.

## ***DEDICATORIA***

*A mi esposa por su entrega, comprensión y cariño, fuente inagotable de amor hacia nuestros hijos: Dieguito, Sarita y Sebitas, me inspiran y me ayudan a seguir adelante en todo momento.*

*A mis padres, quienes contra todo pronóstico siguen mejorando de salud después de permanecer varios días en UCI.*

## AGRADECIMIENTOS

A Marcos y Veerle: Mi tutor y asesora, quienes desde la dirección del proyecto ParamoSUS, gestionaron el financiamiento de la beca para mis estudios de maestría, además estuvieron presentes en cada una de las etapas de mi investigación, sus conocimientos y guía han sido cruciales; la palabra gracias se queda pequeña en comparación a su vital aporte.

A Gabriel y Sebastián: Mi co-tutor y revisor respectivamente, quienes brindaron orientación experta, críticas constructivas y valiosas sugerencias que enriquecieron mi investigación. Su dedicación, preocupación y compromiso con mi crecimiento académico son invaluableles.

A Teresa, César, Rafael y Bert: Su dirección institucional al frente de EPMAPS y FONAG, han sido fundamentales, su constante apoyo y preocupación me ofreció la tranquilidad y confianza necesaria para la completa dedicación a mis estudios de maestría.

A mis compañeros Juan, Vivi y Clau, por su amistad y continuo apoyo académico y laboral, quienes desinteresadamente compartieron sus conocimientos y me involucraron en algunos aspectos técnicos, que eran desconocidos para mí al inicio de esta investigación.

A mis suegros Rodrigo y Elina: Por su continuo apoyo en innumerables situaciones familiares, laborales y de estudio, su inmenso amor hacia sus nietos, me brindaron la motivación y tranquilidad necesaria para salir adelante en este proyecto, también a mi cuñada Tania por su constante apoyo.

A mi esposa Eli: Por su incalculable amor y paciencia con nuestros hijos: Dieguito, Sarita y Sebitas, además de su cariño y apoyo incondicional en cada uno de los buenos y malos momentos que hemos vivido.

A mis padres y hermanos que de una u otra forma siempre han estado presentes apoyándome incondicionalmente.

A mis compañeros de la maestría, en especial a Érick, Álvaro, Daya y Henry: Por el inmenso valor de compartir experiencias, conocimientos y desafíos juntos. Su compañía hizo que esta etapa sea más llevadera.

Con gratitud y aprecio,  
*Diego*



# UNIVERSIDAD TÉCNICA DEL NORTE

## DIRECCIÓN DE BIBLIOTECA

### 1. IDENTIFICACIÓN DE LA OBRA

En cumplimiento del Art. 144 de la Ley de Educación Superior, hago la entrega del presente trabajo a la Universidad Técnica del Norte para que sea publicado en el Repositorio Digital Institucional, para lo cual pongo a disposición la siguiente información:

DATOS DE CONTACTO			
CÉDULA DE IDENTIDAD:	1719039131		
APELLIDOS Y NOMBRES:	Escobar González Diego Polivio		
DIRECCIÓN:	Av. Las Casas Oe9B y Antonio Herrera		
EMAIL:	diego.p.escobar.g@gmail.com		
TELÉFONO FIJO:	3203567	TELÉFONO MÓVIL:	0984647302

DATOS DE LA OBRA			
TÍTULO:	MODELOS BASADOS EN DATOS PARA ESTIMAR HUMEDAD DEL SUELO Y PRECIPITACIÓN EN LOS ANDES TROPICALES DEL NORTE DE ECUADOR		
AUTOR (ES):	Diego Escobar González, Marcos Villacís Erazo, Gabriel Jácome Aguirre		
FECHA: DD/MM/AAAA	10/12/2024		
PROGRAMA:	<input type="checkbox"/> GRADO	<input checked="" type="checkbox"/> POSGRADO	
TITULO POR EL QUE OPTA:	Máster en Ciencias de la Ingeniería para la Gestión de Recursos Hídricos		
ASESOR /DIRECTOR:	Ing. Marcos Villacís Erazo PhD.		

### 2. CONSTANCIAS

El autor manifiesta que la obra objeto de la presente autorización es original y se la desarrolló, sin violar derechos de autor de terceros, por lo tanto, la obra es original y que es el titular de los derechos patrimoniales, por lo que asume la responsabilidad sobre el contenido de la misma y saldrá en defensa de la Universidad en caso de reclamación por parte de terceros.

Ibarra, a los 10 días del mes de diciembre de 2024

#### EL AUTOR:

Firma



Firmado electrónicamente por:  
DIEGO POLIVIO  
ESCOBAR GONZALEZ

Nombre: Diego Escobar González



Ibarra, 25 de noviembre de 2024



Dra.  
Lucía Yépez  
**DECANA FACULTAD DE POSGRADO**

**ASUNTO:** Conformidad con el documento final

Señora Decana:

Nos permitimos informar a usted que revisado el Trabajo final de Grado “MODELOS BASADOS EN DATOS PARA ESTIMAR HUMEDAD DEL SUELO Y PRECIPITACIÓN EN LOS ANDES TROPICALES DEL NORTE DE ECUADOR” del maestrante Diego Polivio Escobar González, de la Maestría en Ciencias de la Ingeniería para la Gestión de los Recursos Hídricos. certificamos que han sido acogidas y satisfechas todas las observaciones realizadas.

Atentamente,

	<b>Apellidos y Nombres</b>	<b>Firma</b>
Director	Ing. Marcos Joshua Villacis Erazo PhD.	 <p>Firmado electrónicamente por: <b>MARCOS JOSHUA VILLACIS ERAZO</b></p>
Asesor	Ing. Gabriel Alexis Jácome Aguirre MSc.	 <p>Firmado electrónicamente por: <b>GABRIEL ALEXIS JACOME AGUIRRE</b></p>

## CONTENIDO

<b>ParamoSUS</b>	<b>I</b>
<b>DEDICATORIA</b>	<b>II</b>
<b>AGRADECIMIENTOS</b>	<b>III</b>
<b>LISTA DE TABLAS</b>	<b>VI</b>
<b>LISTA DE FIGURAS</b>	<b>VII</b>
<b>LISTA DE SÍMBOLOS Y SIGLAS</b>	<b>X</b>
<b>RESUMEN</b>	<b>1</b>
<b>ABSTRACT</b>	<b>3</b>
<b>CAPÍTULO I: EL PROBLEMA</b>	<b>4</b>
1.1 Planteamiento del Problema . . . . .	4
1.2 Antecedentes . . . . .	6
1.3 Objetivos . . . . .	11
1.3.1 Objetivo General . . . . .	11
1.3.2 Objetivos Específicos . . . . .	11
1.4 Justificación . . . . .	11
<b>CAPÍTULO II: MARCO REFERENCIAL</b>	<b>13</b>
2.1 Marco Teórico . . . . .	13
2.1.1 Contenido de agua en el suelo y métodos de medición . . . . .	13
2.1.2 Parámetros hidrometeorológicos y sus métodos de medición . . . . .	14
2.1.3 Análisis estadístico cuantitativo . . . . .	18
2.1.4 Inteligencia artificial para pronóstico de series temporales . . . . .	20
2.1.5 Transfer Learning . . . . .	24
2.2 Marco Legal . . . . .	27
<b>CAPÍTULO III:METODOLOGÍA</b>	<b>29</b>

3.1	Área de estudio . . . . .	30
3.2	Análisis y evaluación de datos recopilados . . . . .	32
3.2.1	Inventario de información eco-hidrometeorológica . . . . .	32
3.2.2	Preprocesamiento y validación de la información . . . . .	37
3.3	Descripción de la climatología y humedad del suelo . . . . .	39
3.3.1	Análisis exploratorio de datos en la microcuenca JTU_01 y en la estación JTU_AWS . . . . .	39
3.3.2	Distribución geográfica de la precipitación multianual mensual en el DMQ y zonas de influencia . . . . .	40
3.3.3	Análisis de la intensidad duración frecuencia de eventos de lluvia . . . . .	42
3.3.4	Distribución temporal de la información meteorológica . . . . .	45
3.4	Descripción de la relación entre precipitación y humedad del suelo . . . . .	45
3.5	Diagrama de flujo de la metodología del uso de redes neuronales para pronóstico de humedad del suelo . . . . .	46
3.5.1	Arquitectura de la Red Neuronal . . . . .	46
3.5.2	Hiperparámetros . . . . .	47
3.6	Evaluación . . . . .	51
<b>CAPÍTULO IV: RESULTADOS Y DISCUSIÓN</b>		<b>53</b>
4.1	Análisis de datos . . . . .	53
4.1.1	Análisis exploratorio de datos climatológicos en JTU_AWS . . . . .	53
4.1.2	Análisis exploratorio de datos de temperatura y humedad del suelo en JTU_01 . . . . .	55
4.1.3	Distribución geográfica estacional de la lluvia . . . . .	59
4.1.4	Distribución de la precipitación horaria en la estación JTU_AWS . . . . .	67
4.1.5	Curvas intensidad duración frecuencia en JTU_AWS . . . . .	70
4.1.6	Distribución temporal de la información climatológica . . . . .	72
4.2	Descripción de la relación entre precipitación y humedad del suelo . . . . .	75
4.2.1	Ciclo diurno de la humedad del suelo . . . . .	79
4.2.2	Tiempo de respuesta de la humedad del suelo frente a un evento de lluvia intensa . . . . .	83



4.2.3	Efecto de la precisión del reflectómetro <i>CS650</i> en las series de humedad del suelo . . . . .	84
4.3	Red Neuronal Profunda . . . . .	87
4.3.1	Optimización de la función de pérdida . . . . .	87
4.3.2	Red Base para pronóstico de humedad del suelo de almohadilla en la zona baja (LO) del horizonte A . . . . .	88
4.3.3	Aplicación de la red base para pronóstico de humedad del suelo en pajonal usando transfer-learning . . . . .	91
4.3.4	Aplicación de la red base para pronóstico de humedad del suelo en almohadilla usando transfer-learning . . . . .	94
<b>CONCLUSIONES</b>		<b>98</b>
<b>RECOMENDACIONES Y TRABAJO FUTURO</b>		<b>101</b>
<b>REFERENCIAS</b>		<b>104</b>

#### LISTA DE TABLAS

Tabla N°1.1	Métodos de AI-ML usados en el estado del arte, para pronosticar humedad del suelo, organizados de acuerdo al modelo. . . . .	10
Tabla N°2.1	Marco normativo relacionado al Cambio Climático . . . . .	27
Tabla N°2.2	Normativa Legal relacionada con Hidrología y Estudio del Suelo	28
Tabla N°3.1	Descripción perfiles del suelo . . . . .	32
Tabla N°3.2	Descripción perfiles del suelo, ubicación de reflectómetros. . . . .	33
Tabla N°3.3	Ubicación de estaciones hidrometeorológicas usadas para la caracterización de la estacionalidad de la precipitación. . . . .	35
Tabla N°3.4	Ubicación de estaciones usadas para la caracterización de la estacionalidad de la precipitación. . . . .	36
Tabla N°3.5	Arquitectura propuesta para el pronóstico de 48 horas de $\theta_{CU\_LOA}$ dada información horaria de siete días de: $P, T, HR, RS, \theta_{CU\_LOA}$ y $T_{CU\_LOA}$ . . . . .	48
Tabla N°4.1	Resumen estadístico de la lluvia diaria y de las variables meteorológicas con frecuencia de 5 minutos en JTU_AWD Jatunhuayco.	54
Tabla N°4.2	Descripción estadística de la temperatura de suelo ( $^{\circ}C$ ) en los perfiles A y 2A, en almohadilla. . . . .	55
Tabla N°4.3	Descripción estadística de la temperatura de suelo ( $^{\circ}C$ ) en los perfiles A y 2A, en pajonal. . . . .	56

Tabla N°4.4	Descripción estadística de la humedad del suelo (%), en los perfiles A y 2A, en almohadilla. . . . .	57
Tabla N°4.5	Descripción estadística de la humedad del suelo (%), en los perfiles A y 2A, en pajonal. . . . .	57
Tabla N°4.6	Cantidad de precipitación máxima estimada (mm) por duración y período de retorno. . . . .	71
Tabla N°4.7	Métricas obtenidas para la red base que pronostica 48 horas de $\theta_{CU\_LOA}$ , dada la información de las variables: $P, T, HR, RS, \theta_{CU\_LOA}$ y $T_{CU\_LOA}$ de siete días. . . . .	91
Tabla N°4.8	Métricas obtenidas para las redes profundas que pronostican 48 horas de $\theta_{TU}$ , dada la información de las variables: $P, T, HR, RS, \theta_{TU}$ y $T_{TU}$ . . . . .	94
Tabla N°4.9	Métricas obtenidas para las redes profundas que pronostican 48 horas de $\theta_{CU}$ , dada la información de las variables: $P, T, HR, RS, \theta_{CU}$ y $T_{CU}$ . . . . .	97

## LISTA DE FIGURAS

Figura N° 1.1	Generalización taxonómica de los modelos de IA empleados en la Tabla 1.1. . . . .	9
Figura N° 2.1	El ML visto como un subcampo de la inteligencia artificial; el DL como un subcampo del ML; y las redes neuronales como núcleo del DL. . . . .	20
Figura N° 2.2	Representación básica de una Red Neuronal . . . . .	21
Figura N° 2.3	Representación en 2-D de una convolución. El deslizamiento o <i>stride</i> del kernel $k$ es de 1 paso sobre la matriz de entrada $I$ . El resultado $I*K$ es el <i>feature map</i> , es decir el de productos de elementos a elemento (element wise). . . . .	23
Figura N° 2.4	Estructura básica de una RNN, esta red recurrente sólo procesa información de la entrada $x$ incorporándola al estado $h$ que se transmite hacia adelante a través del tiempo. . . . .	23
Figura N° 2.5	Estructura de la unidad que compone una una red LSTM. Donde $\odot$ es el producto de Hadamard. . . . .	25
Figura N° 3.1	Diagrama de flujo de la metodología propuesta. . . . .	29
Figura N° 3.2	Área de estudio . . . . .	31
Figura N° 3.3	Disponibilidad de información de la estación JTU_AWD y de los perfiles y horizontes de los dos tipos de suelo en Jatunhuayco JTU_01, desde octubre 2018 a octubre 2022. . . . .	33

Figura N° 3.4	Metodología para construir la distribución geográfica de la precipitación multianual mensual en el DMQ y su área de influencia, y caracterización de las condiciones meteorológicas en la estación JTU_AWS y de suelo en la microcuenca Jatunhuayco JTU_01. . . . .	39
Figura N° 3.5	Diagrama de flujo para pronóstico de humedad del suelo con redes neuronales. . . . .	46
Figura N° 3.6	Red base: pronóstico de humedad del suelo en almohadillas en el perfil A ( $\theta_{CU\_LOA}$ ) y transfer-learning para dominos similares en las 15 combinaciones. . . . .	49
Figura N° 4.1	Violin-plots de las variables meteorológicas en JTU_AWS y de suelo en la microcuenca Jatunhuayco JTU_01. . . . .	58
Figura N° 4.2	Visualización de la correlación entre de las series de precipitación mensual de las estaciones seleccionadas. . . . .	60
Figura N° 4.3	Visualización de las componentes principales de precipitación multianual mensual. . . . .	61
Figura N° 4.4	Dendograma circular de precipitación multianual mensual . . . . .	62
Figura N° 4.5	Comparación de estacionalidades en el Distrito Metropolitano de Quito y sus principales sistemas de abastecimiento hídrico. . . . .	64
Figura N° 4.6	Tipos de estacionalidad en el DMQ y sus principales sistemas de abastecimiento hídrico . . . . .	66
Figura N° 4.7	Precipitación horaria en Jatunhuayco JTU_AWS, enero 2019 - septiembre 2022 . . . . .	69
Figura N° 4.8	Ecuación IDF, escala logarítmica, estación JTU_AWS . . . . .	71
Figura N° 4.9	Distribución de la temperatura del aire media diaria ( $^{\circ}C$ ) en Jatunhuayco de enero 2019 a octubre 2022. . . . .	73
Figura N° 4.10	Media mensual del número de días secos en Jatunhuayco. . . . .	74
Figura N° 4.11	Distribución de la humedad relativa media diaria (%) en Jatunhuayco 2019 - octubre 2022. . . . .	75
Figura N° 4.12	Precipitación y humedad del suelo en almohadilla en los horizontes A y 2A a diferentes escalas temporales, $\theta_{CU\_UPA}$ y $\theta_{CU\_UP2A}$ , Jatunhuayco enero 2019 - septiembre 2022. . . . .	77
Figura N° 4.13	Precipitación y humedad del suelo en pajonal en los horizontes A y 2A a diferentes escalas temporales, $\theta_{TU\_UPA}$ y $\theta_{TU\_UP2A}$ , Jatunhuayco enero 2019 - septiembre 2022. . . . .	78
Figura N° 4.14	Serie de $\Delta\theta_{CU\_UPA}$ en Jatunhuayco, enero 2019 - septiembre 2022	81
Figura N° 4.15	Serie de $\Delta\theta_{TU\_UPA}$ en Jatunhuayco, enero 2019 - septiembre 2022	82

Figura N° 4.16	(a) Heatmap precipitación horaria acumulada Jatunhuayco, enero-2020. (b) Heatmap de la serie de $\Delta\theta_{CU\_UPA}$ en Jatunhuayco enero-2020, (c) Heatmap de la serie de $\Delta\theta_{TU\_UPA}$ en Jatunhuayco enero-2020. . . . .	84
Figura N° 4.17	Box-plot de las series de diferencias consecutivas: $\Delta\theta_{CU\_UPA}$ y $\Delta\theta_{TU\_UPA}$ . . . . .	86
Figura N° 4.18	Función de pérdida de la red que pronostica los valores horarios de $\theta_{CU\_LOA}$ , para las siguientes 48 horas, dada información horaria de: $P, T, HR, \theta_{CU\_LOA}$ y $T_{CU\_LOA}$ , de siete días. . . . .	88
Figura N° 4.19	Pronóstico de 48 horas de $\theta_{CU\_LOA}$ , para el conjunto de datos de evaluación horarios de 7 días de: $P, T, HR, RS, \theta_{CU\_LOA}$ y $T_{CU\_LOA}$ . . . . .	89
Figura N° 4.20	Pronóstico para diferentes momentos de la serie de evaluación $\theta_{CU\_LOA}$ . . . . .	90
Figura N° 4.21	Pronósticos en diferentes momentos de humedad del suelo en pajonal $\theta_{TU}$ , en todos los perfiles y horizontes analizados. . . . .	93
Figura N° 4.22	Pronósticos en diferentes momentos de humedad del suelo en almohadilla $\theta_{CU}$ , en todos los perfiles y horizontes analizados. . . . .	96

## LISTA DE SÍMBOLOS Y SIGLAS

### SÍMBOLOS

- $I$  : Intensidad de Precipitación  
 $P$  : Precipitación  
 $t$  : Tiempo  
 $\Theta$  : Humedad del suelo

### SIGLAS

- ANN : Artificial Neural Networks  
ARIMA : Autoregressive Integrated Moving Average  
BRF : Boruta Random Forest  
CNN : Convolution Neural Networks  
DMQ : Distrito Metropolitano de Quito  
ELM : Extreme Learning Machine  
EPMAPS : Empresa Pública Metropolitana de Agua Potable y Saneamiento  
FONAG : Fondo para la Protección del Agua  
GAN : Generative Adversarial Networks  
GNN : Graph Neural Networks  
GRU : Gated Recurrent Units  
IA : Inteligencia Artificial  
IDF : Intensidad-Duración-Frecuencia  
INAMHI : Instituto Nacional de Meteorología e Hidrología  
LASSO : Least Absolute Shrinkage and Selection Operator,  
LSTM : Long Short-Term Memory  
MLP : Multi-layers Perceptrons  
RNN : Recurrent Neural Networks  
SSP : Trayectorias Socioeconómicas Compartidas  
SVR : Support vector regression

UNIVERSIDAD TÉCNICA DEL NORTE  
INSTITUTO DE POSGRADO  
PROGRAMA DE MAESTRÍA EN  
CIENCIAS DE LA INGENIERÍA PARA LA GESTIÓN DE LOS RECURSOS HÍDRICOS

**MODELOS BASADOS EN DATOS PARA ESTIMAR HUMEDAD DEL SUELO Y  
PRECIPITACIÓN EN LOS ANDES TROPICALES DEL NORTE DE ECUADOR**

**Autor:** Mat. Diego Polivio Escobar González

**Tutor:** Ing. Marcos Joshua Villacis Erazo PhD.

**Co-Tutor:** Ing. Gabriel Alexis Jácome Aguirre MSc.

**Año:** 2023

**RESUMEN**

El problema abordado en esta investigación se enfoca en la necesidad de comprender el comportamiento de diferentes parámetros hidrometeorológicos y de suelo, en particular la interacción suelo-lluvia, considerando los dos tipos predominantes de vegetación: almohadilla y pajonal, en la microcuenca de Jatunhuayco (JTU\_01), ubicada en los páramos de los Andes tropicales del Norte de Ecuador. El trabajo tiene como objetivo desarrollar un modelo basado en datos para predecir la humedad del suelo mediante variables hidrometeorológicas, con la finalidad de mejorar la planificación futura de los recursos hídricos. Para este acometido, se utilizaron técnicas de clusterización para caracterizar la estacionalidad de la lluvia en la microcuenca de estudio y sus zonas de influencia, se analizaron los eventos lluvia considerando la intensidad, duración y frecuencia, por otro lado, se estudió la distribución de frecuencias de todas las variables utilizadas, así como, un análisis exploratorio de datos, lo que permitió comprender el comportamiento de las mismas. Se analizó la relación entre precipitación y humedad del suelo a diferentes escalas temporales, considerando también las diferencias consecutivas de los registros de humedad; se utilizaron modelos de redes neuronales, para construir una red base para pronosticar humedad del suelo, luego se utilizaron

técnicas de transfer-learning, para realizar pronósticos de humedad del suelo en dominios similares, finalmente se evaluó la eficiencia. Los resultados proporcionaron una base sólida para el pronóstico preciso de la humedad del suelo. Estos hallazgos también resaltan la utilidad de las técnicas de transfer-learning en la mejora de los pronósticos en dominios similares. Este trabajo contribuye al avance de la comprensión y predicción de la humedad del suelo, lo cual es fundamental para la gestión eficiente de los recursos hídricos y la toma de decisiones en el ámbito hidrológico y ambiental.

**Palabras clave:** Humedad del suelo, redes neuronales, transfer-learning, meteorología, climatología.

# **DATA-DRIVEN MODELS FOR ESTIMATING SOIL MOISTURE AND PRECIPITATION IN THE TROPICAL ANDES OF NORTHERN ECUADOR**

## **ABSTRACT**

The problem addressed in this research focuses on the need to understand the behavior of different hydrometeorological and soil parameters. In particular, the soil-rainfall interaction of the two predominant vegetation types, almohadilla and pajonal, in the Jatunhuayco micro-watershed (JTU\_01), located in the moorlands of the tropical Andes of northern Ecuador are studied. In order to improve future water resource planning, this work aims to develop a data-based model to predict soil moisture using hydrometeorological variables. In this context, clustering techniques were used to characterize the microbasin rainfall seasonality and rainfall events were analyzed considering intensity, duration and frequency. The frequency distribution of all variables used was studied, and exploratory data analysis was performed, which allowed to better understand their behavior. The relationship between precipitation and soil moisture at different time scales was analyzed by considering the consecutive differences of moisture records. On the other hand, neural network models were used to build a base network to forecast soil moisture, then transfer-learning techniques were used to perform soil moisture forecasts in similar domains. Each model was evaluated by efficiency metrics. The results provided a solid basis for accurate soil moisture forecasting. These findings also highlight the usefulness of transfer-learning techniques in terms of improving forecasts in similar contexts. This work contributes to the advancement of soil moisture understanding and prediction, which is critical for efficient water resource management and hydrological and environmental decision making.

**Key words:** Soil moisture, neural networks, transfer-learning, hydrometeorology, climatology.



## CAPÍTULO I: EL PROBLEMA

### 1.1 Planteamiento del Problema

La humedad del suelo es una variable clave en el sistema climático, así como en los ciclos hídricos, energéticos y biogeoquímicos (Oki et al., 2006; Trenberth et al., 2009), además su mejor comprensión y cuantificación ayudan significativamente a reducir la incertidumbre en los escenarios climáticos futuros, en particular con respecto a eventos extremos (sequías, períodos largos continuos de lluvia), así como impactos en los ecosistemas y la agricultura, como lo describe Seneviratne et al. (2010). La precipitación, como principal vía por la cual el agua desciende a la Tierra y alimenta a las plantas, cuerpos de agua y acuíferos subterráneos (Kuttippurath et al., 2021), se mide y registra utilizando diversos equipos, como pluviómetros, pluviógrafos y disdrómetros (González-Fernández, 2017). Estos instrumentos proporcionan datos precisos sobre la cantidad de precipitación, permitiendo comprender su impacto en el ciclo hidrológico.

En relación al páramo andino, se reconoce la importancia del suelo y su cobertura vegetal en la hidrología de esta región, como se menciona en Bonnesoeur et al. (2019). La dinámica de la humedad del suelo durante eventos de lluvia es de particular interés para la planificación de los recursos hídricos futuros (Patiño et al., 2018). Además, en Buytaert et al. (2006) se manifiesta que: “las condiciones atmosféricas previas y la humedad del suelo juegan un papel crucial en la conectividad de las laderas y el funcionamiento hidrológico”; estos aspectos resaltan la estrecha relación entre la precipitación, la humedad del suelo y la hidrología del páramo andino.

De esta manera, comprendiendo la influencia de la precipitación en el ciclo hidrológico y reconociendo el papel esencial del suelo y su humedad en el contexto del páramo andino, es posible desarrollar un enfoque integral para la gestión sostenible de los recursos hídricos en esta importante región.

Iniciativas de cooperación científica como el proyecto ParamoSUS y sus primeros resultados presentados en (Páez-Bimos, Molina et al., 2022; Páez-Bimos, Villacís et al., 2022),

muestran como influye la lluvia en el comportamiento de la humedad del suelo frente a dos tipos de cobertura vegetal *Azorella pedunculata* (Almohadilla) y *Calamagrostis intermedia* (Pajonal) presentes en Jatunhuayco (páramo del norte de Ecuador), por otro lado, conocer el comportamiento histórico de la humedad del suelo y su relación con la precipitación, permite comprender de mejor manera, fenómenos como por ejemplo: sequías y eventos intensos de precipitación. Se debe considerar también que, la precipitación en los Andes tropicales del norte de Ecuador es muy heterogénea, aunque se consideren pequeñas distancias, como se puede ver en Ruiz-Hernández et al. (2021), además, la humedad del suelo también es muy variable, ya que tenemos la influencia de lagunas como La Mica, glaciares como el Antisana, así como, humedales y diferentes tipos de cobertura vegetal y suelo.

Resulta interesante plantear un modelo que tenga como entrada precipitación, junto con otras variables hidrometeorológicas y cuya salida sea el pronóstico de humedad del suelo. La literatura muestra que existen algunas investigaciones para casos de estudio muy particulares y con una frecuencia de datos de entrada y salida semanales o en el mejor de los casos diaria (Gao, Qiu et al., 2021; Han et al., 2021). Puesto que, normalmente son escasos los registros de contenido volumétrico del agua en el suelo, sería de suma importancia contar con un modelo basado en datos para el pronóstico de humedad del suelo, el cual permitiría mejorar el monitoreo ambiental y la gestión del recurso hídrico (Prasad et al., 2019); por otro lado, los modelos basados en procesos físicos y/o deterministas, requieren gran cantidad de variables ambientales observadas, lo cual no siempre es factible, por ejemplo, en Prasad et al. (2019) se utiliza el modelo físico WaterDyn, para simular los parámetros hidrológicos, incluido el nivel de humedad del suelo, utilizando datos meteorológicos registrados previamente, características del suelo, tipo vegetación, radiación solar y albedo, de la misma forma en Páez-Bimos, Molina et al. (2022), se utilizó el modelo HYDRUS-1D para simular los procesos hidrológicos del suelo en intervalos diarios, incluyendo la evapotranspiración, la drenaje profundo y almacenamiento de agua en el suelo.

Si se consigue un modelo basado en datos a baja frecuencia de lluvia y otras variables hidrometeorológicas, para predecir la humedad del suelo, sería un buen aporte, ya que con esa base se podría también predecir eventos extremos, como potenciales incendios en suelos

secos, la influencia en la calidad del agua en épocas de lluvias continuas, principalmente por la lluvia antecedente (Rostami et al., 2018); igualmente, pronosticar la humedad del suelo podría ser un importante insumo para identificar zonas de posible riesgo de derrumbes y aluviones.

Se propone describir el régimen climático de la zona de estudio, realizando un análisis exploratorio de datos utilizando diferentes medidas estadísticas descriptivas, análisis espacial con clusters, identificando el régimen estacional de pluviosidad y el tipo de eventos de precipitación de la zona de estudio, construyendo y estudiando las curvas intensidad-duración-frecuencia de lluvia (Escobar-González et al., 2022), luego se plantea describir la relación entre precipitación - humedad de suelo, para posteriormente aplicar modelos basados en datos como redes neuronales, para pronosticar humedad utilizando datos de variables hidrometeorológicas registradas en campo; posteriormente se plantea utilizar técnicas de transfer-learning, para transferir el conocimiento de un modelo ya entrenado y pronosticar humedad del suelo en otros puntos de monitoreo que cuenten con un dominio similar. Finalmente se espera evaluar la eficacia de los modelos usados minimizando en la función de pérdida el error cuadrático medio (RMSE), optimizando los estadísticos coeficiente de eficiencia de Nash–Sutcliffe (NSE) y el coeficiente de eficiencia Kling-Gupta (KGE).

## **1.2 Antecedentes**

Es importante conocer las investigaciones previas que se han realizado y serán la base de la presente investigación, por tanto, se presenta una revisión de literatura preliminar relacionada con la línea del tema planteado. Es conocido que la interacción suelo-lluvia es fundamental en la hidrología del páramo Andino (Patiño et al., 2018), donde también se manifiesta que: “la dinámica de la humedad del suelo durante eventos de lluvia se ve afectada y por tanto, es posible realizar una planificación futura”. Por otro lado, Buytaert et al. (2006) y Páez-Bimos (2022) indican que las condiciones previas de humedad junto con las condiciones atmosféricas presentes, pueden indicar el tipo de conexión actual que afectará la conectividad de las laderas y el funcionamiento hidrológico. En Seneviratne et al. (2010), uno de sus resultados indica que: “una mejor comprensión y cuantificación de la humedad

del suelo ayuda significativamente a reducir la incertidumbre en los escenarios climáticos futuros, en particular con respecto a eventos extremos”.

Existen pocos estudios que analizan la relación entre precipitación y humedad del suelo en los Andes tropicales en Ecuador, sin embargo y dado que existen pocas mediciones que caractericen los procesos de generación de escorrentía en los Andes del sur de Ecuador, en Molina et al. (2007) se utilizó un simulador de lluvias (con intensidades cercanas a 40 mm/h) y así se cuantificó la respuesta hidrológica de distinta cobertura o uso de suelo y se comprobó que en suelos degradados se genera escorrentía superficial a los pocos minutos del inicio del evento de lluvia, siendo más aguda la respuesta hidrológica, si además hay presencia de laderas; por otro lado, comportamientos más variados se presentan en coberturas vegetales como pastizales o tierras de cultivo, caracterizados por su alta capacidad de infiltración, así mismo, en Patiño et al. (2018), se indica que conocer el comportamiento histórico de precipitación y humedad del suelo permite comprender de mejor manera eventos severos como: sequías y eventos extremos de precipitación.

Si se consideran las investigaciones realizadas en los Andes del norte de Ecuador, se encuentra una limitada línea de investigación histórica, sin embargo, en el proyecto de cooperación ”Linking Global Change with Soil and Water Conservation in the High Andes”(ParamoSus), se han desarrollado investigaciones como (Páez-Bimos, Molina et al., 2022; Páez-Bimos, Villacís et al., 2022), donde se estudia la humedad del suelo en la microcuenca de Jatunhuayco de  $0.7 \text{ km}^2$ , en la cual predominan dos tipos de cobertura vegetal: almohadilla y pajonal, sobre las cuales se está monitoreando variables hidrometeorológicas y de suelo en cuatro diferentes perfiles desde 2018, considerando los dos tipos de vegetación. Además, en estos estudios se consideran los ciclos hidrológicos y biogeoquímicos, donde se verifica que influyen en los flujos y el transporte de agua del suelo, que son fundamentales para la meteorización química. Para ampliar los conocimientos sobre la infraestructura natural para mitigar la erosión en los Andes, en Vanacker et al. (2022) se analizó sistemáticamente literatura gris y revisada por pares, donde se verifica la efectividad de las intervenciones y medidas de conservación del suelo que reducen la erosión del suelo por el agua; sin embargo, la mitigación de la erosión es mayor cuando se mantiene la vegetación natural, por otro

lado, se necesita más investigación para evaluar si la efectividad se mantiene durante eventos extremos, como por ejemplo: El Niño-Oscilación del Sur.

El análisis de la relación entre precipitación y humedad del suelo ha sido estudiado con muchos enfoques, considerando entre ellos: esquemas de sequías (Berg & Sheffield, 2018; Iñiguez et al., 2016; Quiring et al., 2015; Yuan et al., 2020), con el fin de estimar la probabilidad de propagación de incendios forestales, así como, la gestión y la distribución del recurso hídrico en presencia de caudales bajos; y períodos de lluvia intensa (Nicholson et al., 1994; Noguchi et al., 1997), donde se analiza la relación de suelos saturados consecuencia de periodos de lluvia intensos o lluvia sostenida, con zonas de inundación, deslizamientos, así como, operatividad de sistemas de abastecimiento de agua potable.

Es importante señalar que en nuestro país existe un vacío en la investigación sobre suelos profundos, es así que, en estudios como Patiño et al. (2018), se analiza la respuesta de la humedad del suelo de suelos poco profundos frente a eventos de lluvia intensos, considerando cobertura vegetal prístina y se establecen umbrales en el cambio máximo de la humedad del suelo, en la misma se considera también la humedad y precipitación antecedente.

Por otro lado, en Ruiz-Hernández et al. (2021) se manifiesta que: “la precipitación en los Andes tropicales del norte de Ecuador es muy variable, incluso en distancias muy cortas, además se indica una posible espacialización de la precipitación y distribución del ciclo diurno, lo cual puede ser un insumo para estimar y pronosticar la humedad del suelo”. En González-Vergara et al. (2020), se realizó una clasificación de eventos extremos de precipitación en la ciudad de Quito utilizando redes neuronales, análisis de componentes principales, regresión de relieve, utilizando información de diferentes parámetros hidrometeorológicos y evaluando estos resultados con registros in-situ, estos resultados muestran las ventajas de utilizar modelos basados en datos y su posible extensión y/o adaptación al pronóstico de variables como a humedad del suelo. En Páliz Larrea et al. (2021), se realizaron pronósticos de niveles de agua con modelos de redes neuronales (NN) y modelos tipo adaptive neuro-fuzzy inference systems (ANFIS), en el embalse de Salve Faccha en los Andes tropicales del norte de Ecuador, utilizando como datos de entrada información de precipitación y de niveles en el embalse, con sensores instalados en la zona de estudio, lo cual muestra los beneficios y

versatilidad de utilizar modelos basados en datos, como redes neuronales, en el pronóstico de variables y fenómenos hidrometeorológicos en los Andes del norte de Ecuador.

Una tendencia en los últimos años es la utilización de modelos basados en datos, particularmente redes neuronales y sus aplicaciones han sido estudiadas ampliamente en ciencia de la tierra, como en (Sun et al., 2022), así como pronósticos del tiempo utilizando modelos globales como (Irrgang et al., 2021) e incluso en hidrología (Ibrahim et al., 2022; Kim et al., 2021; Lange, 2020). En particular, la Tabla 1.1, describe los trabajos más actuales con enfoques de AI (Artificial Intelligence), ML(Machine Learning) y DL (Deep Learning) que se dedican al pronóstico de la humedad del suelo, considerando como inputs diferentes tipos de variables eco-hidrometeorológicas, como: precipitación, temperatura, humedad relativa, radiación solar, presión, velocidad y dirección de viento, temperatura y humedad del suelo, entre otras.

Desde un punto de vista más amplio, usando la información de la Tabla 1.1, se puede organizar los modelos de AI que sirven para el pronóstico de humedad del suelo como se indica en la Figura 1.1.

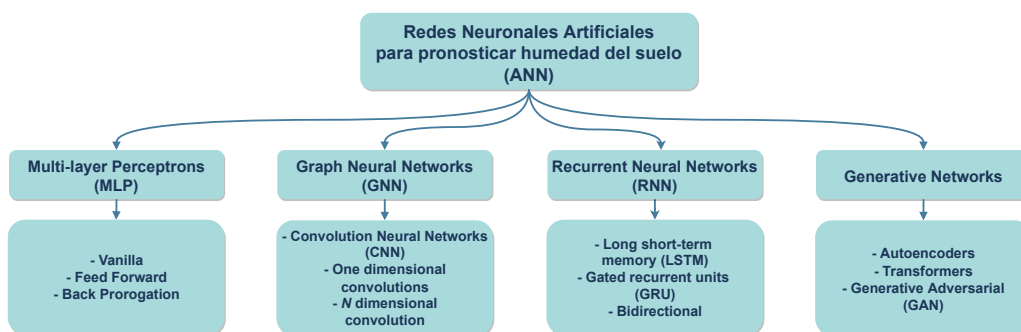


Figura N° 1.1: Generalización taxonómica de los modelos de IA empleados en la Tabla 1.1.

**Tabla N° 1.1:** Métodos de AI-ML usados en el estado del arte, para pronosticar humedad del suelo, organizados de acuerdo al modelo.

Modelo	Objetivo	Referencia
LSTM	Proporcionar un pronóstico de la humedad del suelo, casi en tiempo real, basado en humedad del suelo usando LSTM y kernels.	Fang y Shen (2020)
	Pronosticar la humedad media diaria del suelo de un huerto de <i>litchi</i> durante cuatro temporadas de crecimiento diferentes a partir de información meteorológica y datos de humedad del suelo.	Gao, Qiu et al. (2021)
	Usar un modelo basado en LSTM para pronosticar los valores de humedad del suelo en función de los parámetros recopilados de varios sensores inteligentes utilizados para el riego de cultivos.	Suebsombut et al. (2021)
ANN	Comparar los resultados del pronóstico de la humedad del suelo obtenida mediante ANN entrenadas usando optimización por enjambre de partículas, búsqueda de Cuckoo y <i>Flower Pollination Algorithm</i> .	Chatterjee et al. (2020)
	Investigar modelos regionales de humedad del suelo para capas superficiales (0–10 cm) y subsuperficiales (10–20 cm) mientras se evalúan los principales parámetros que afectan los resultados en un contexto de cambio climático.	Bartels et al. (2021)
	Comparar una arquitectura de red neuronal que utiliza un algoritmo <i>longicorn beetle search</i> y optimizadores GA para pronosticar la humedad del suelo a partir de datos meteorológicos.	An y Zhao (2021)
	Construir una ANN de regresión usando optimización de retropropagación con GA para la predicción de la humedad del suelo en pendientes laterales altas.	Liu et al. (2022)
CNN	Un algoritmo CNN para pronosticar la humedad del suelo en zonas agrícolas utilizando imágenes Sentinel-1.	Hegazi et al. (2021)
GAN	Predicción de la humedad disponible en el suelo para lograr una agricultura sostenible de la planta de “cacahuete” ( <i>Apios americana</i> ) utilizando una GAN con el algoritmo de optimización Sailfish.	Sivasankaran et al. (2021)
ANN,LSTM	Mejorar las predicciones de la humedad y la conductividad eléctrica del suelo utilizando información meteorológica.	Gao, Xie et al. (2021)
	Pronosticar con 1-6 días de antelación la humedad del suelo a profundidades de 100, 200, 500 y 1000 mm desde la superficie, basándose en los datos meteorológicos de las características del suelo.	Han et al. (2021)
ANN, ELM	Proponer una red neuronal de regresión generalizada mejorada (EGRNN) para la estimación de la humedad del suelo.	Zhang et al. (2022)
CNN,GRU	Diseñar un modelo híbrido de DL para pronosticar humedad superficial del suelo diario utilizando imágenes de satélite.	A. M. Ahmed et al. (2021)
CNN, LSTM, ConvLstm	Aplicar modelos de aprendizaje de transferencia (transfer learning) para la predicción de la humedad del suelo activa y pasiva de 3, 5 y 7 días.	Li et al. (2021)
LSTM,SVR	Comparar un LSTM basado en Boruta-RF con un SVR para emular la humedad del suelo bajo escenarios de calentamiento global.	A. Ahmed et al. (2021)
Graph Neural Network, SVR, ARIMA, RNN	Proponer una nueva red neuronal gráfica para aprender estructuras temporales y predecir la humedad del suelo.	Vyas y Bandyopadhyay (2020)
SVR, LASSO, RF,ANN	Predicción de la humedad del suelo para el cultivo de papa aplicando selección de características.	Dubois et al. (2021)

## **1.3 Objetivos**

### ***1.3.1 Objetivo General***

Pronosticar la variabilidad espacio - temporal de la humedad de suelo con base en modelos basados en datos hidrometeorológicos recopilados en campo.

### ***1.3.2 Objetivos Específicos***

- Caracterizar la estacionalidad y las condiciones ambientales de la zona de estudio, con el fin de establecer tendencias en los pronósticos de humedad del suelo.
- Describir la relación entre humedad del suelo, precipitación y los datos hidrometeorológicos recopilados en el campo.
- Pronosticar valores de humedad del suelo usando redes neuronales y técnicas de transfer-learning, para transferir el conocimiento de un modelo ya entrenado en un modelo con dominio similar y evaluar la eficiencia de los modelos.

## **1.4 Justificación**

Nuestro país no dispone de una red hidrometeorológica que esté bien distribuida espacialmente, ni tampoco robusta en el tiempo, como se indica en Cadilhac et al. (2017) y Vanacker et al. (2018), el estudio estricto de las relaciones existentes entre diferentes variables hidrometeorológicas, así como el Cambio Climático exige contar con registros de varias décadas de información (Arnbjerg-Nielsen et al., 2013).

Dentro de la limitada información hidrometeorológica recopilada en nuestro país, en particular el parámetro precipitación, es del cual se dispone mayor cantidad de registros, no así la humedad del suelo. Un conocimiento más amplio de la interacción suelo-lluvia, nos permitirá entender de mejor manera la hidrología del suelo e hidrología de las cuencas, durante eventos de lluvia y de esta manera planificar de mejor manera la gestión de los recursos hídricos, por otro lado, el análisis de información histórica de precipitación y humedad, permitirá comprender fenómenos, como: sequías y eventos de intensa precipitación (Patiño et



al., 2018).

En la subcuenca de Jatunhuayco JTU\_01 en El Antisana, ubicado en los Andes tropicales del Norte de Ecuador (Ochoa-Tocachi et al., 2016), se dispone de información hidrometeorológica, incluida humedad del suelo desde octubre 2018, que puede ser usada para analizar la interacción entre estas variables a diferentes escalas temporales, como por ejemplo en Páez-Bimos, Molina et al. (2022), por otro lado, se pueden utilizar modelos basados en datos, que permitan obtener humedad del suelo en función de otras variables hidrometeorológicas, como en Páez-Bimos, Villacís et al. (2022), lo cual podría ser un buen aporte para predecir eventos extremos como incendios en suelos secos, derrumbes, aluviones y otros fenómenos en suelos muy húmedos (Prasad et al., 2018).

## CAPÍTULO II: MARCO REFERENCIAL

### 2.1 Marco Teórico

#### 2.1.1 Contenido de agua en el suelo y métodos de medición

Como se describe en Silva Robledo et al. (2015), el suelo está compuesto de: “partículas orgánicas y minerales, así como de zonas porosas donde se encuentra aire y agua”, es decir el volumen total

$$V_t = V_s + V_p, \text{ donde } V_p = V_a + V_w \quad (2.1)$$

con  $V_s$  como el volumen de las partículas del suelo,  $V_p$  volumen de poros,  $V_a$  volumen de aire y  $V_w$  volumen de agua. Existen dos formas directas de expresar el contenido de agua en el suelo:

- **Contenido gravimétrico de agua ( $w$ ).**- corresponde a la masa de agua  $M_w$  por una unidad de masa de suelo secado a  $105^\circ C$ , hasta un peso constante denominado  $M_s$ , es decir:

$$w = \frac{M_w}{M_s},$$

donde  $M_w = M_t - M_s$ , con  $M_t$  como la masa de total o de suelo húmedo, como se detalla en Schmutge et al. (1980).

- **Contenido volumétrico de agua ( $\Theta$ ).**- se define como la relación entre el volumen de agua  $V_w$  y el volumen del suelo  $V_t$  y es equivalente al contenido gravimétrico del agua por la densidad aparente del suelo  $D_a$ , como se muestra en Silva Robledo et al. (2015) y Andrades et al. (2015).

$$\Theta = \frac{V_w}{V_t} = W D_a.$$

Por otro lado, existen sensores para estimar el contenido volumétrico de agua, conocidos como: Time Domain Reflectometer (TDR) y Water Content Reflectometer (WCR), como se describe en Topp et al. (1980).

Aquí, se reporta la retención, como el contenido volumétrico de agua ( $\theta, cm^3 cm^{-3}$ ), para un potencial mátrico dado  $\Psi_m(kPa)$ , es así que, para la saturación de agua ( $\Theta_{SAT}$ ) el potencial será  $0 kPa$ , capacidad de campo ( $\Theta_{FC}$ ), será  $-10 kPa$  y el punto de marchitez ( $\Theta_{WP}$ ) en  $-1500 kPa$ . Finalmente el agua total disponible, ( $\Theta_{TAW}$ ) se calcula como la diferencia entre la retención de agua de la capacidad de campo y el punto de marchitez, como se indica en Páez-Bimos (2022).

## ***2.1.2 Parámetros hidrometeorológicos y sus métodos de medición***

### **2.1.2.1 Hidrometeoro**

Según González-Fernández (2017) un hidrometeoro es: “un conjunto de partículas de agua líquida o sólida en suspensión en la atmósfera o cayendo a través de ella” y se clasifican en cinco grupos de partículas:

- En suspensión, como por ejemplo: nubes, niebla neblina.
- En precipitación, como por ejemplo: lluvia, llovizna, nieve, granizo.
- En depósito, como por ejemplo: rocío, escarcha, hielo liso.
- Levantadas por el viento desde la superficie terrestre, como por ejemplo: ventisca o columna de nieve.
- Con giro en vórtice, como por ejemplo: tornado, nube embudo.

### **Precipitación**

Según Glickman (2000) “la precipitación es un hidrometeoro liberado desde la atmósfera y que intercepta la superficie terrestre, se puede presentar en formas como: lluvia, llovizna, nieve, agua nieve, manga de agua o tromba, monzón, chubasco, tormentas que puede incluir granizo”. Se debe indicar también que la lluvia corresponde a la precipitación de partículas líquidas de agua. Por otro lado, “la virga, neblina, escarcha y rocío no son formas de precipitación sino de condensación” (Gultepe, 2008). La lluvia se mide en milímetros de agua (mm), lo cual corresponde a la cantidad de litros precipitados por  $m^2$ .

Existen diferentes equipos que miden la lluvia de forma manual o automática, como se describe en Sheppard y Joe (2008), como por ejemplo:

- Pluviómetro totalizador: corresponde a un recipiente de forma cilíndrica con un embudo en su interior, cuya descarga se da en un envase con escala graduada.
- Pluviómetro de intensidades o de sifón: es de similar estructura externa al pluviómetro totalizador, pero que internamente cuenta con un tambor que gira a velocidad constante y registra con una pluma la altura de agua en una faja de papel graduado, gracias a un sistema de boya o flotador, según la cantidad de lluvia recolectada, para más detalles se puede ver (Chen, 1974).
- Pluviómetro de balancín: similar a un pluviómetro totalizador, pero que internamente cuenta con un sistema tipo sube y baja, en cuyos extremos se encuentra un recipiente con un volumen de usualmente 0.1 mm de lluvia, donde cada basculación es registrada en un datalogger, que genera automáticamente una serie de temporal de datos, como se describe en (Chen, 1974).

#### 2.1.2.2 Curvas Intensidad Duración y Frecuencia

Una curva Intensidad-Duración-Frecuencia (IDF), según Mohyont et al. (2004) “es una relación matemática, usualmente empírica, entre la intensidad de precipitación ( $i$ ), su duración ( $t$ ) y el período de retorno ( $T$ )”. Existe un gran número de aproximaciones empíricas que relacionan:  $I$ ,  $t$  y  $T$ , además, el periodo de retorno de un evento, indica la rareza o frecuencia de este evento y se define por la inversa de la probabilidad de ocurrencia (Mohyont et al., 2004; Willems, 2000).

#### 2.1.2.3 Evento de lluvia

Un evento de lluvia constituye la precipitación de agua líquida, que resulta de un proceso atmosférico independiente, es decir, se encuentra separado temporalmente de otro proceso atmosférico que resulte también en lluvia.

Un hietograma según García (2006) es: “la representación gráfica cronológica de la

intensidad de precipitación, cantidad precipitación instantánea o acumulada”.

#### 2.1.2.4 Evento de lluvia extremo

Es un evento de lluvia fuera de lo común en términos de cantidad y duración, y se lo puede definir considerando la lluvia acumulada, correspondiente a los valores más altos de períodos de retorno.

Un hietograma según García (2006) es: “la representación gráfica cronológica de la intensidad de precipitación, cantidad precipitación instantánea o acumulada”.

#### 2.1.2.5 Evento de lluvia extremo

Es un evento de lluvia fuera de lo común en términos de cantidad y/o duración, y se lo puede definir considerando la lluvia acumulada, correspondiente a los valores más altos de períodos de retorno.

#### 2.1.2.6 Temperatura del aire

Según Rodríguez et al. (2004) “la temperatura mide la cantidad de energía cinética de las partículas que constituyen la materia, mientras más rápido se muevan mayor será la temperatura”. Por lo tanto, la temperatura del aire se puede entender como la medida de la energía cinética promedio de las partículas analizadas en el aire.

El termómetro es el instrumento que mide la temperatura y existen diferentes tipos, como los descritos en Rodríguez et al. (2004), como:

- Termómetros digitales: “utilizan dispositivos transductores, que a través de circuitos electrónicos transforman tensión en una unidad de temperatura” (Chen, 1974).
- Termómetro de máximas: en el interior dispone cerca del depósito una estrangulación, que cuando la dilatación del mercurio empuja con suficiente fuerza, vence la resistencia de la estrangulación, quedando en el extremo libre la posición medida más avanzada.
- Termómetro de mínimas: normalmente es de alcohol y posee sumergido en su interior un esmalte como un índice, que cuando la temperatura disminuye arrastra éste índice

registrando la temperatura más baja alcanzada.

#### 2.1.2.7 Humedad relativa

La humedad relativa ( $\phi$ ) en una mezcla de aire y agua, se “define como la razón entre la cantidad de agua presente en una determinada porción de atmósfera ( $p_{H_2O}$ ) y la cantidad total de vapor de agua que la porción de atmósfera puede soportar ( $p_{H_2O}^*$ ) a una determinada temperatura” (Brutsaert, 2013), por tanto se puede expresar en (%) y se calcula como:

$$\phi = \frac{p_{H_2O}}{p_{H_2O}^*} * 100$$

La humedad relativa se mide con un instrumento llamado higrómetro, existen higrómetros automáticos que registran en un datalogger la información como una serie temporal a una frecuencia configurable (Rodríguez et al., 2004).

#### 2.1.2.8 Radiación solar

Como se indica en Constante-Segura y Palacios-Chacón (2014), “la energía producida por la radiación solar en la superficie terrestre es diferente a la radiación que incide sobre la parte exterior de la atmósfera y las dos se componen de un conjunto diferente de frecuencias y longitudes de onda, en su conjunto representan el espectro de la radiación solar incidente en la Tierra”. La energía emitida por un cuerpo negro ideal puede ser calculada por la ecuación de Plank:

$$E = h\nu$$

Donde  $\nu$  la frecuencia de oscilación de los fotones y  $h$  es la constante de Plank, aproximadamente igual a  $6.625 \times 10^{-34} Js$ ; además, en Constante-Segura y Palacios-Chacón (2014), se manifiesta que: “los fotones con poca energía no pueden ingresar a la atmósfera terrestre, lo que muestra la principal diferencia entre la radiación solar terrestre y extraterrestre”.

La radiación solar que incide sobre la parte exterior de la atmósfera de la superficie terrestre, se la llama “constante solar  $G_{sc}$  y representa la energía perpendicular a la dirección de propagación de la radiación por unidad de tiempo en un área de un  $m^2$ ” (Duffie et al., 2020)

y corresponde a  $1366.1 \text{ W/m}^2$ .

Para medir la radiación solar se disponen de una amplia gama de métodos y equipos, como se indica en Duffie et al. (2020), entre directos (sensores de radiación) e indirectos (los cuales requieren ser modelados y estimados), como por ejemplo:

- Directos: piranómetro, pirheliómetro, radiómetros.
- Indirectos: detectores de horas sol, detectores fotovoltaicos, detectores termoelectrónicos.

Destacamos entre ellos al piranómetro, el cual “es un sensor electrónico que mide con gran exactitud la radiación incidente en una superficie, con  $180^\circ$  de campo de visión y que puede ser conectado a un datalogger para construir una serie temporal con una frecuencia configurable” (Rodríguez et al., 2004).

### ***2.1.3 Análisis estadístico cuantitativo***

#### ***2.1.3.1 Análisis exploratorio de datos***

El análisis exploratorio de datos se refiere a diferentes procedimientos implementados a series temporales o conjunto de datos, con el objetivo de comprender y describir patrones que indiquen posibles tendencias o concordancias entre variables, como se describe en Andrienko et al. (2006). Además, incluye la verificación del conjunto inicial de datos, así como, el uso de técnicas cuantitativas representadas en gráficos que permitan visualizar la estructura de la información.

#### ***2.1.3.2 Estacionalidad***

La estacionalidad es la presencia de variaciones que se producen a intervalos regulares de tiempo, usualmente inferiores a un año; en el caso del clima corresponde a patrones periódicos de lluvia o sequía, repetitivos que pueden ser predecibles en una serie temporal, como se menciona en Andrienko et al. (2006).

### 2.1.3.3 Series temporales

Se define a una serie temporal  $X$  como un conjunto de  $n$  observaciones ordenadas cronológicamente:

$$X = (x_0, x_1, \dots, x_t, \dots, x_n), \text{ donde } x_t \in \mathbb{R}. \quad (2.2)$$

### 2.1.3.4 Clasificación no supervisada

La clasificación no supervisada, más conocida como clustering, consiste en un análisis de conglomerados homogéneos, los cuales son métodos utilizados en varias ramas de la ciencia e ingeniería para agrupar grandes cantidades de información, en Hidrología por ejemplo, se utiliza para encontrar patrones climáticos; la métrica más utilizada para comparar dos conjuntos de información es la *distancia* (Ramos, 2001; Santos et al., 2019)

### 2.1.3.5 Distancia

Dadas las series de registros de lluvia de tres diferentes estaciones  $P_i, P_j, P_k$ , una distancia  $d$  satisface:

- $d(P_i, P_j) \geq 0$ , positividad.
- $d(P_i, P_j) = 0$ , si y solo si  $P_i = P_j$ .
- $d(P_i, P_j) = d(P_j, P_i)$ , simetría.
- $d(P_i, P_j) \leq d(P_i, P_k) + d(P_k, P_j)$ , desigualdad triangular.

Las distancias que usamos son las de la familia de Minkowski (Kreyszig, 1991), la cual se define como:

$$L_p(P_i, P_j) = \left( \sum_{k=1}^m |P_{k,i} - P_{k,j}|^p \right)^{1/p}; \quad \forall p \geq 1, \quad p \in \mathbb{Z}^+. \quad (2.3)$$

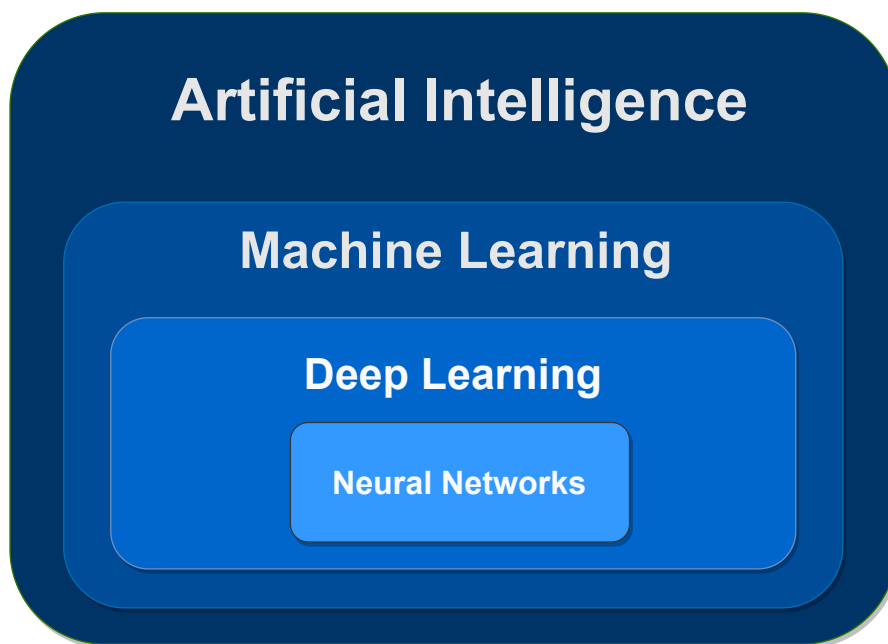
Se puede notar que las distancias: Manhattan, Euclidiana y del máximo, son casos especiales de la distancia de Minkowski cuando  $p = 1$ ,  $p = 2$  y  $p \rightarrow \infty$  respectivamente, como se puede ver en la *Enciclopedia de Distancias* (Deza et al., 2009).



#### 2.1.4 Inteligencia artificial para pronóstico de series temporales

La inteligencia artificial *Artificial intelligence (AI)*, es un concepto macro, usado para referirse a máquinas o “robots” capaces de realizar tareas propias de los seres humanos, (Joshi, 2020). De tal manera que, “la AI es una teoría general para el desarrollo de sistemas informáticos que permitan replicar la inteligencia humana” (Joshi, 2020). Sin embargo, no hay una definición universal, por ejemplo Russell (2010) “define la AI como el estudio de agentes que reciben percepciones del entorno y realizan acciones”.

Como se menciona en Hurwitz et al. (2018) y Murphy (2012), el aprendizaje de máquina *Machine Learning (ML)*, es una rama de la AI, que proporciona métodos automatizados para el análisis de datos. A pesar de que el ML es un área de estudio muy relacionado con la estadística y minería de datos, “se particulariza por proveer métodos para detectar automáticamente patrones en los datos y luego mediante los patrones descubiertos tomar de decisiones” (Murphy, 2012). Otra rama de la AI es el aprendizaje profundo, *Deep Learning (DL)*, que no es mas que una familia de modelos de redes neuronales con múltiples capas de unidades en la red (Hurwitz et al., 2018; Kelleher et al., 2018). Esta conceptualización es explicada en la Figura 2.1.



**Figura N° 2.1:** El ML visto como un subcampo de la inteligencia artificial; el DL como un subcampo del ML; y las redes neuronales como núcleo del DL.

De acuerdo a Gamboa (2017), “existe una gran predisposición de usar DL para pronosticar series temporales, siendo los modelos más relevantes las Redes Neuronales Artificiales, Redes Neuronales Convoluciones, Redes Neuronales Recursivas y representaciones híbridas o derivadas de estas”.

#### 2.1.4.1 Red Neuronal Artificial

Conocida también como Multilayer Perceptron, Vanilla Neural Network, Artificial Neural Network o Feedforward Neural Networks. “Es una abstracción matemática para simular las neuronas biológicas como dispositivos computacionales” (Alpaydin, 2016). De acuerdo a Kelleher et al. (2018) “la unidad más simple de una red neuronal se llama *neurona*, y el conjunto de neuronas alineados secuencialmente se lo conoce como *capa*”. Las neuronas y las capas se pueden unir asignándoles un *peso*  $w$  y una *función de activación*  $f$  que aplica un mapeo no lineal, lo que según Kelleher et al. (2018), convierte a la red neuronal una función de regresión lineal de múltiples entradas. Sin embargo, según Goodfellow et al. (2016), el problema de las redes neuronales se puede plantear la siguiente manera: dado un conjunto de entradas  $X = \{x_0, \dots, x_n\}$ , optimizar los hiperparámetros  $\theta =$ [número de neuronas, número capas escondidas, funciones de activación, pesos] para encontrar la representación más cercana la variable objetivo  $Y = y_0, \dots, y_j$ . Es decir, desde un punto de vista analítico, la red neuronal se puede ver como el mapeo  $\phi(X, \theta) \mapsto \hat{y}$ . De manera gráfica, esto se puede representar como en la Figura 2.2, donde las entradas  $x_0, \dots, x_n$  se unen con los pesos  $w_{j0}, \dots, w_{jn}$  y forman el valor  $H_j$  de una neurona de la capa escondida. La función de activación  $f$  genera un nuevos valores  $H_j$ . La capa de salida genera una nuevo valor  $\hat{y}$  que aproxima al valor real  $y$ .

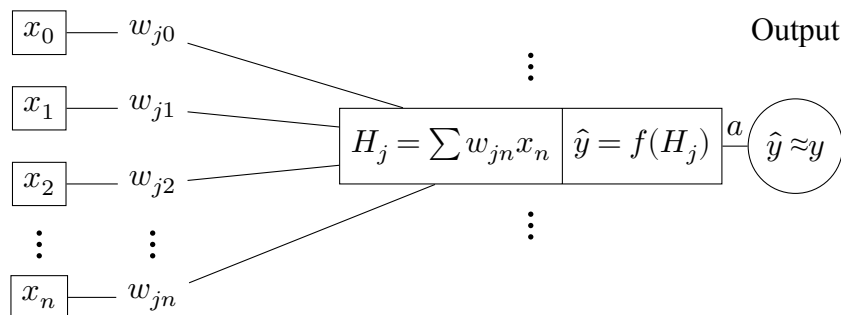


Figura N° 2.2: Representación básica de una Red Neuronal

Como se describe en Goodfellow et al. (2016), uno de los hitos importantes en las redes neuronales fue incorporar correctamente el algoritmo de retropropagación (back-propagation) al entrenamiento de la red. Este algoritmo, utiliza la derivada de una función compuesta, a través de la regla de la cadena y permite encontrar el gradiente de una función objetivo, también llamada función de costo, función de pérdida o función de error.

#### 2.1.4.2 Red Neuronal Convolutiva

De acuerdo a Goodfellow et al. (2016), son un tipo de red neuronal especializadas para procesar datos con topología de mallas, que realiza el aprendizaje directamente de sus datos; por tal razón, son muy comunes en aplicaciones de visión por computador, reconocimiento de imágenes, entre otras. Su nombre se deriva de la operación de convolución que se lleva a cabo en el proceso de aprendizaje. La operación de convolución, según Ketkar y Santana (2017), en entradas de 1-Dimensión se puede definir como:

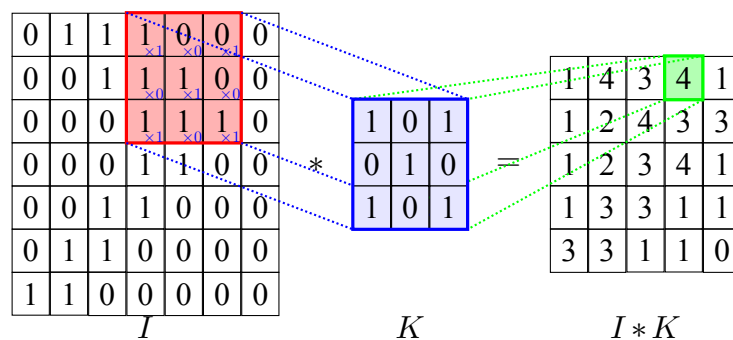
$$s(t) = \sum_a I(a) * \kappa(t - a), \quad (2.4)$$

donde  $I(t)$  es una entrada y  $\kappa(a)$  el kernel. Mientras que para entradas de 2-Dimensiones se define como:

$$s(t) = \sum_a \sum_b I(a, b) * \kappa(m - a, n - b), \quad (2.5)$$

donde  $I(m, n)$  representan los datos de entrada y  $\kappa(a, b)$  es el kernel. De tal manera, para realizar una convolución el kernel (también llamado filtro) “se representa como una matriz multidimensional de parámetros, que se desliza sobre los valores de entrada, efectuando una operación matemática en ambas” (Alpaydin, 2016). La salida de aplicar una convolución entre los datos de entrada y un filtro se conoce como *feature map*, Figura 2.3. Finalmente, Goodfellow et al. (2016) indica que la red neuronal convolutiva completa su estructura usando una capa de detección o *pooling* que reemplaza los valores de las representaciones internas para entradas específicas, usando una función estadística como la media, el máximo o el mínimo.

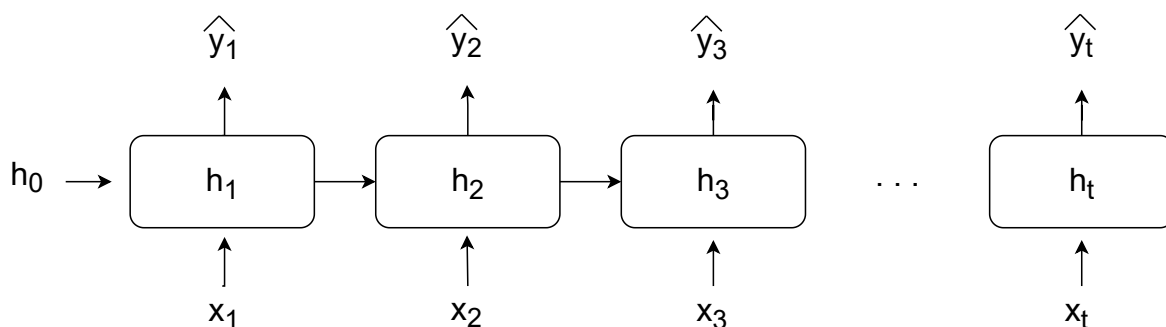
En Goodfellow et al. (2016), se aclara que el uso de estas redes tienen beneficios sobre



**Figura N° 2.3:** Representación en 2-D de una convolución. El deslizamiento o *stride* del kernel  $k$  es de 1 paso sobre la matriz de entrada  $I$ . El resultado  $I*K$  es el *feature map*, es decir el de productos de elementos a elemento (element wise).

el resto, dado que dependiendo el tipo de filtro a usar se puede generar una disminución en el número de conexiones y por ende menor número de hiperparámetros a optimizar; o en su defecto aumentar el tamaño de los datos para generar una representación nueva de los datos.

Según Goodfellow et al. (2016) “son una familia de redes neuronales conceptualizadas para de retroalimentarse de los resultados previos entre sus conexiones”. Es decir, que no solamente utilizan los datos de entrada  $x_i$ , sino también los resultados de entradas anteriores, *e.g.* :  $x_{i-1}$ , para generar la predicción actual  $\hat{y}_i$ . Son comúnmente usadas procesar datos secuenciales y se utilizan a menudo para pronosticar series temporales (Petneházi, 2019). De manera general, su estructura se muestra en la Figura 2.4.



**Figura N° 2.4:** Estructura básica de una RNN, esta red recurrente sólo procesa información de la entrada  $x$  incorporándola al estado  $h$  que se transmite hacia adelante a través del tiempo.

Sin embargo, una implementación simple de RNN carece de practicidad en secuencias largas debido al desvanecimiento del gradiente: “a medida que el gradiente recorre las conexiones recurrentes de la red, se vuelve *desvanecidamente* pequeño, lo que impide que el peso actualice su valor” (Sit et al., 2020). Una solución es usar las redes Long Short-Term Me-

mory, que según Sit et al. (2020) resuelven el problema de desvanecimiento de gradiente. Las redes LSTM están compuestas por un estado  $c_t$  y tres “compuertas”, olvido-entrada-salida. Antes de pasar por las compuertas, se combina en un vector  $v$  el valor de entrada  $x_t$  con el valor de la neurona anterior  $h_{t-1}$ . La compuerta de olvido se conforma por el vector  $v$  y una función sigmoideal  $\sigma$ , el resultado es  $f_t$ . La compuerta de entrada se conforma por el vector  $v$ , una función sigmoideal  $\sigma$ , una función  $\tanh$ , el resultado es  $\tilde{c}_t$ . El estado  $c_t$  se obtiene usando  $C_{t-1}, f_t, i_t, \tilde{c}_t$  el resultado es  $\tilde{c}_t$ . Finalmente, la compuerta de salida se obtiene usando el vector  $v, o_t, \tilde{c}_t$ , el resultado es  $h_t$ . Así, la Figura 2.5, en conjunto con las ecuaciones (2.6) a (2.11), resumen este proceso.

$$i_t = \sigma(w_{1,i} \cdot h_{t-1} + w_{2,i} \cdot x_t + b_i) \quad (2.6)$$

$$f_t = \sigma(w_{1,f} \cdot h_{t-1} + w_{2,f} \cdot x_t + b_f) \quad (2.7)$$

$$o_t = \sigma(w_{1,o} \cdot h_{t-1} + w_{2,o} \cdot x_t + b_o) \quad (2.8)$$

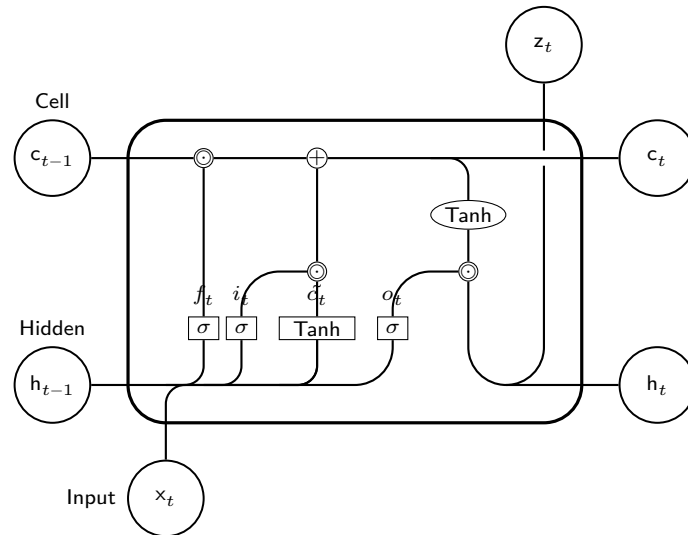
$$\tilde{c}_t = \tanh(w_{1,c} \cdot h_{t-1} + w_{2,c} \cdot x_t + b_c) \quad (2.9)$$

$$c_t = i_t \odot \tilde{c}_t + f_t \odot C_{t-1} \quad (2.10)$$

$$z_t = h_t = o_t \odot \tanh(C_t) \quad (2.11)$$

### 2.1.5 Transfer Learning

La idea detrás del aprendizaje por transferencia es que los modelos preentrenados ya han aprendido características generales a partir de grandes conjuntos de datos, lo que los hace



**Figura N° 2.5:** Estructura de la unidad que compone una red LSTM. Donde  $\odot$  es el producto de Hadamard.

altamente eficaces para tareas similares. Al utilizar un modelo preentrenado como punto de partida, se puede reducir significativamente el tiempo y los recursos necesarios para entrenar un modelo desde cero. La metodología, se centra en transferir el conocimiento a través de diferentes dominios (Zhuang et al., 2020). Es decir, Transfer Learning “tiene como objetivo mejorar el performance de los modelos de aprendizaje mediante el traspaso del conocimiento de un dominio de particular hacia otro relacionado” (Zhuang et al., 2020). Además de ahorrar tiempo y recursos, el aprendizaje por transferencia también puede mejorar la precisión del modelo, especialmente cuando el conjunto de datos de entrenamiento es pequeño. “Al utilizar un modelo preentrenado, se pueden evitar problemas como el sobreajuste” (Tan et al., 2018), ya que el modelo ha aprendido a generalizar características útiles. De igual manera, “el aprendizaje de transferencia profunda (*Deep transfer learning*) estudia cómo utilizar el conocimiento de otros campos mediante redes neuronales profundas” (Tan et al., 2018). Es así que, “existen cuatro categorías de Deep transfer learning: instances-based, mapping-based, network-based y adversarial-based” (Tan et al., 2018). En particular, este trabajo usa el network-based transfer-learning, que reutiliza una red neuronal pre-entrenada en un dominio de origen, para aprovechar ese conocimiento en otro dominio.

En resumen, el aprendizaje por transferencia es una técnica eficaz de aprendizaje automático que permite aprovechar los modelos preentrenados para resolver tareas similares,

reducir el tiempo y los recursos necesarios para entrenar un modelo desde cero y mejorar la precisión del modelo.

## 2.2 Marco Legal

En este contexto se indica la relevancia del presente proyecto de investigación, mismos que servirán como base para la gestión de recursos hídricos, difusión e investigación de conocimiento; sobre todo en el ámbito del cambio climático.

**Tabla N° 2.1: Marco normativo relacionado al Cambio Climático.**

Cuerpo Legal	Fecha y Alcance	Sección	Art.	Extracto/Resumen
Acuerdo de París	2015 - Global	Artículo 1		<i>"b) Aumentar la capacidad de adaptación a los efectos adversos del cambio (...)"</i>
		Artículo 7		<i>"(...) aumentar la capacidad de adaptación, fortalecer la resiliencia y reducir la vulnerabilidad al cambio climático con miras a contribuir al desarrollo sostenible (...)"</i>
Constitución de la República del Ecuador	2008 - Nacional	TITULO VII- Régimen del Buen Vivir Capítulo Segundo- Biodiversidad y Recursos Naturales Sección Sexta-Agua	Art. 414	<i>"El Estado adoptará medidas adecuadas y transversales para la mitigación del cambio climático (...)"</i> .
Código Orgánico del Ambiente	2017 - Nacional	TITULO I-Objeto, Ámbito y Fines	Art. 3	<i>"Establecer medidas eficaces, eficientes y transversales para enfrentar los efectos del cambio climático a través de acciones de mitigación y adaptación (...)"</i>
		TITULO II-De la Adaptación y Mitigación del Cambio Climático	Art. 257	<i>"En las acciones de adaptación se crearán y fortalecerán las capacidades del país para afrontar los impactos del cambio climático, con énfasis en la reducción de la vulnerabilidad (...)"</i>
		Capítulo I- Disposiciones Generales	Art. 258	<i>"2. Considerar los escenarios actuales y futuros del cambio climático en los instrumentos de planificación territorial, el desarrollo de infraestructura, el desarrollo de actividades productivas y de servicios, los asentamientos humanos".</i>
		Capítulo II-Medidas Mínimas para Adaptación y Mitigación	Art. 261.	<i>"(...) protección de las zonas vulnerables a inundaciones, sequías, heladas, y degradación del suelo (...)"</i>



**Tabla N° 2.2: Normativa Legal relacionada con Hidrología y Estudio del Suelo.**

<b>Cuerpo Legal</b>	<b>Fecha y Alcance</b>	<b>Sección</b>	<b>Art.</b>	<b>Extracto/Resumen</b>
Ley Orgánica de Recursos Hídricos Usos y Aprovechamiento del Agua	2014 - Nacional	TITULO I- Disposiciones Generales	Art. 5	<i>"La planificación y gestión de los recursos hídricos debe considerar el estudio de la calidad del agua y su impacto en el suelo y la biodiversidad."</i>
Ley Orgánica de Ordenamiento Territorial, Uso y Gestión de Suelo	2016 - Nacional	TITULO II- Clasificación de Suelos	Art. 12	<i>"Se establecen categorías de suelos en función de su capacidad de retención de agua para fines de planificación hidrológica."</i>
Ley Orgánica de Ordenamiento Territorial, Uso y Gestión de Suelo	2016 - Nacional	TITULO III- Estrategias de Mitigación	Art. 8	<i>"Esta ley establece estrategias para la mitigación de la degradación de suelos y la conservación de su calidad."</i>

### CAPÍTULO III: METODOLOGÍA

En este capítulo, se describe la metodología empleada para abordar los objetivos propuestos. En primer lugar, se presenta la zona de estudio y se realiza un inventario de la información hidrometeorológica y de suelo, incluyendo su preprocesamiento y validación. Posteriormente, se abordan los objetivos del proyecto siguiendo el diagrama de flujo de la Figura 3.1, el cual muestra tres bloques principales. En el primer bloque, se caracterizan las condiciones ambientales, identificando la estacionalidad de la lluvia en el DMQ y los principales sistemas de abastecimiento hídrico. Además, se realiza la caracterización de las condiciones meteorológicas, temperatura y humedad del suelo en Jatunhuayco, se presenta una descripción más detallada de este bloque en la Figura 3.4. En el segundo bloque, se describe la relación entre la lluvia y la humedad del suelo a diferentes escalas temporales, analizando el comportamiento de los ciclos diarios a través de las diferencias consecutivas de los registros de humedad del suelo en contraste con la lluvia. Finalmente, en el tercer bloque se presenta el pronóstico de la humedad del suelo. Se identifica la arquitectura de la red base y se aplican técnicas de transfer-learning para estimar la humedad del suelo en sitios con dominios similares. Los resultados se evalúan utilizando estadísticos adecuados. Se encuentran más detalles sobre la metodología aplicada en este bloque en el diagrama de flujo de la Figura 3.5.

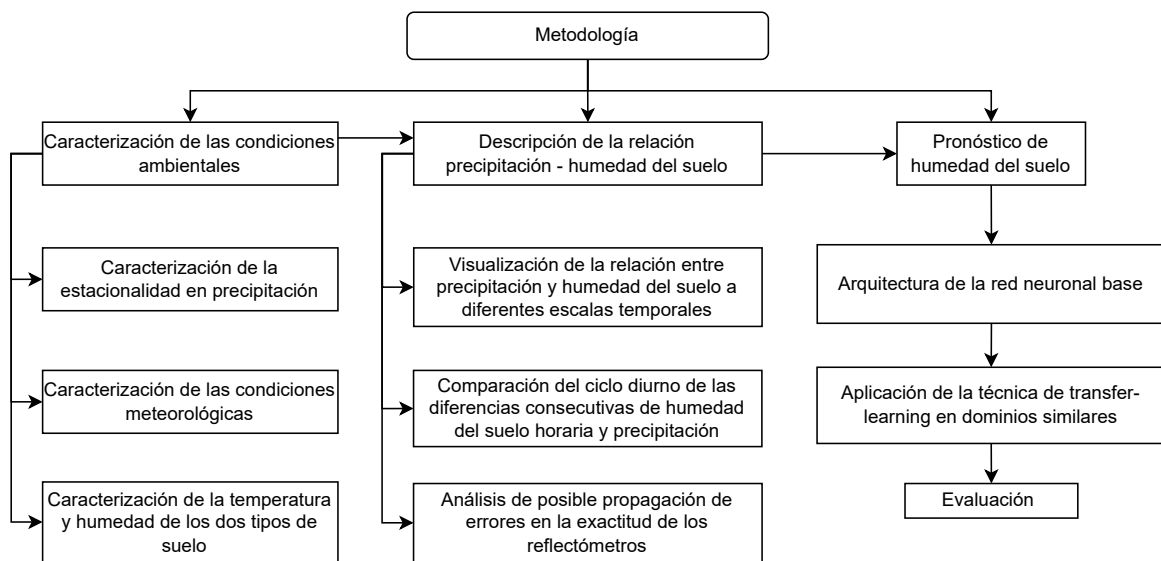


Figura N° 3.1: Diagrama de flujo de la metodología propuesta.

### 3.1 Área de estudio

La microcuenca de estudio es Jatunhuayco (JTU\_01) con una superficie de  $0.7 \text{ km}^2$  se encuentra ubicado en las laderas occidentales del volcán Antisana, en los Andes del Norte de Ecuador, dentro del Área de Conservación de Agua del Antisana (ACA del Antisana); cuya elevación oscila entre 4083 y 4279 m.s.n.m y sus pendientes entre el 10 y 22%, como se describe en (Páez-Bimos, Molina et al., 2022; Páez-Bimos, Villacís et al., 2022) Figura 3.2(B). Su importancia radica en que es una microcuenca de estudio ecohidrológicamente bien monitoreada, que pertenece al ecosistema páramo y se encuentra dentro del sistema La Mica Quito Sur, el cual proporciona agua a cerca de medio millón de habitantes del Distrito Metropolitano de Quito, capital de Ecuador (EPMAPS, 2018). El área está gestionada por el Fondo de Protección del Agua (FONAG, 2022) y desde 2011 están prohibidas las actividades antropogénicas y el pastoreo extensivo. La vegetación desempeña un papel fundamental en el ciclo hidrológico a diferentes escalas espaciales y temporales, ya que extrae agua del suelo e influye en las vías y flujos del agua (Kelly et al., 1998; Lahuatte et al., 2022; Molina et al., 2019; Páez-Bimos, Molina et al., 2022). Se excavaron cuatro perfiles de suelo en la cumbre (upper - UP), una réplica en la cumbre (UR), ladera media (middle - MI), ladera inferior (lower - LO), para dos tipos de vegetación, *Azorella pedunculata* (CU-Cushion) y *Calamagrostis intermedia* (TU-Tussock), es decir un total de ocho perfiles y por el tipo de suelo se han identificado cuatro horizontes (A, 2A, 2BC y 3BC); en este estudio se analizarán solamente los horizontes A y 2A, además las toposecuencias se puede ver en la Figura 3.2(B) y en Páez-Bimos, Villacís et al. (2022).

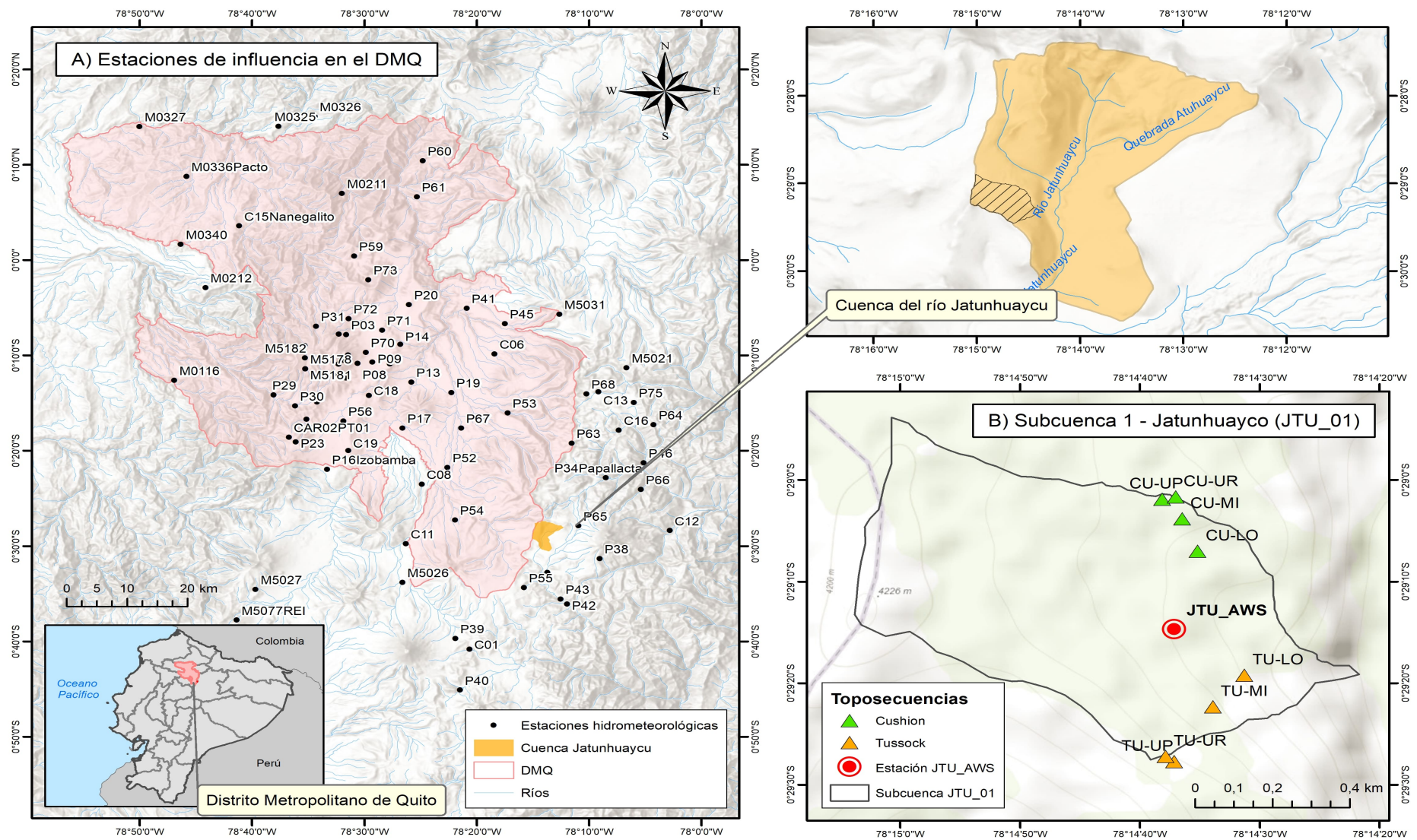


Figura N° 3.2: (A) Estaciones que miden precipitación en el Distrito Metropolitano de Quito y sus principales sistemas de abastecimiento hídrico, (B) Área de estudio microcuenca de Jatunhuayco JTU\_01, estación meteorológica JTU\_AWS y toposecuencias.

## 3.2 Análisis y evaluación de datos recopilados

### 3.2.1 Inventario de información eco-hidrometeorológica

La Tabla N°3.1 muestra la descripción de los perfiles de suelo analizados dentro de la subcuenca Jatunhuayco, considerando las dos coberturas vegetales predominantes *Azorella pedunculata* (CU-Cushion) y *Calamagrostis intermedia* (TU-Tussock), su ubicación, la pendiente media, las variables que se registran: temperatura del suelo y contenido volumétrico del agua en cada perfil analizado; así como, la cobertura vegetal dominante. Además en cada perfil del suelo se encuentran instalados tres reflectómetros para medir el contenido del agua (CS650, Cambell Scientific), como se indica en (Páez-Bimos, Molina et al., 2022), para este estudio se utilizaron los datos de los reflectómetros de los perfiles A y 2A instalados en las profundidades indicadas en la Tabla N°3.2.

Tabla N° 3.1:

#### Descripción perfiles del suelo.

Perfil del suelo	Coordenadas (WGS84)	Altitud (m.s.n.m)	Slope (%)	Variables	Cobertura vegetal dominante (%)
CU_UP	0°29'1.90" S / 78°14'38.15" W	4197	2.5	$T_A, T_{2A},$ $\theta_A, \theta_{2A}$	51.4 ±23.6, <i>Azorella pedunculata</i>
CU_UR	0°29'1.69" S / 78°14'37.69" W	4196	2	$T_A, T_{2A},$ $\theta_A, \theta_{2A}$	78.4 ±6.9, <i>Azorella pedunculata.</i>
CU_MI	0°29'4.22" S / 78°14'36.51" W	4185	12	$T_A, T_{2A},$ $\theta_A, \theta_{2A}$	41.6 ±34.8, <i>Azorella pedunculata</i>
CU_LO	0°29'6.89" S / 78°14'35.08" W	4174	10	$T_A, T_{2A},$ $\theta_A, \theta_{2A}$	54.2 ±38.4, <i>Azorella pedunculata</i>
TU_UP	0°29'27.94" S / 78°14'37.07" W	4225	6.5	$T_A, T_{2A},$ $\theta_A, \theta_{2A}$	55.8 ±21.6, <i>Calamagrostis intermedia</i>
TU_UR	0°29'26.99" S / 78°14'38.14" W	4227	10.5	$T_A, T_{2A},$ $\theta_A, \theta_{2A}$	15.3 ±5.5, <i>Calamagrostis intermedia</i>
TU_MI	0°29'22.36" S / 78°14'34.01" W	4186	22	$T_A, T_{2A},$ $\theta_A, \theta_{2A}$	3.6 ±9.9, <i>Calamagrostis intermedia</i>
TU_LO	0°29'19.08" S / 78°14'31.42" W	4161	20	$T_A, T_{2A},$ $\theta_A, \theta_{2A}$	83.9 ±9.8, <i>Calamagrostis intermedia</i>

Nota 1:  $T_A, T_{2A}, \theta_A, \theta_{2A}$ : temperaturas y contenido volumétrico de agua, perfiles A y 2A.

Nota 2: Mean value ± standard deviation.

Fuente: Adaptado de Páez-Bimos, Villacís et al. (2022).

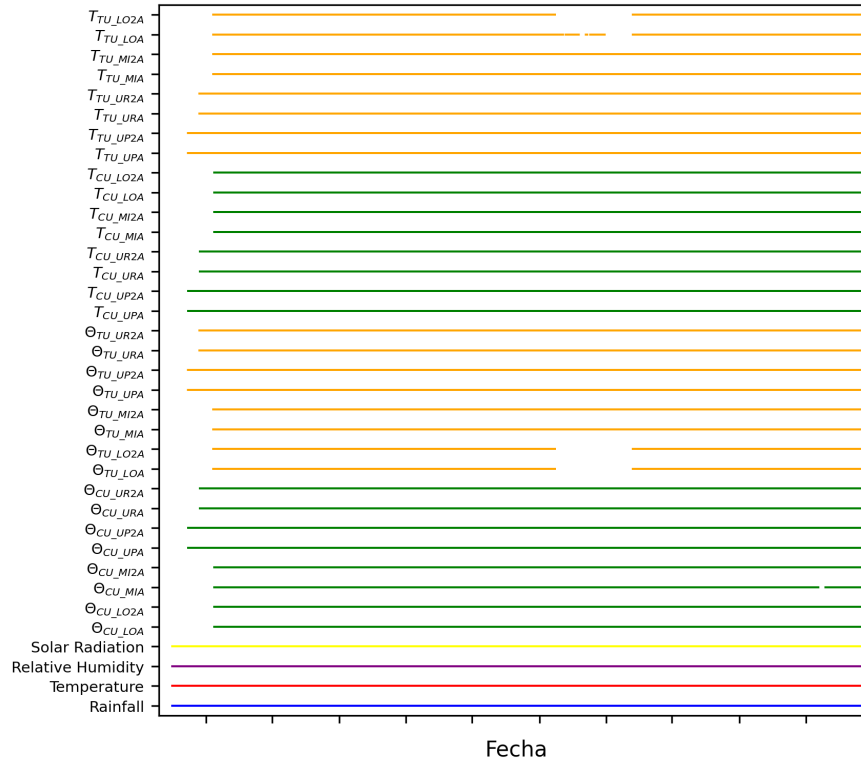
Tabla N° 3.2:

**Descripción perfiles del suelo, ubicación de reflectómetros.**

Soil profile	CU_UP	TU_UP
Mean slope(%)	2	6.5
Soil horizon depth (cm)		
A	8-30	5-30
2A	30-55	30-60
Suction cups and soil water	20	25
Reflectometers installation depth (cm)	40	50

Fuente: Adaptado de Páez-Bimos, Villacís et al. (2022)

En la Figura 3.3, se presenta la disponibilidad de información desde octubre de 2018 a septiembre 2022, con frecuencia de 5 minutos de la estación JTU\_AWD, así como, la información de temperatura y humedad de suelo de los horizontes A y 2A, para los dos tipos de suelo: almohadilla y pajonal; en los 4 perfiles analizados: UP, UR, MI y LO; en todas las variables analizadas, se tiene un porcentaje superior al 97% de información útil en promedio.



**Figura N° 3.3:** Disponibilidad de información de la estación JTU\_AWD y de los perfiles y horizontes de los dos tipos de suelo en Jatunhuayco JTU\_01, desde octubre 2018 a octubre 2022.

En las Tablas N°3.3 y 3.4 se muestran las variables hidrometeorológicas a analizar, que se registran en la estación automática JTU\_AWS, ubicada dentro de la microcuenca de estudio y que recopila información con frecuencia de cinco minutos, desde el 16 de octubre de 2018. Además, se analizó la información mensual de las estaciones que registran lluvia pertenecientes a las redes de monitoreo hidrometeorológico de la Empresa Pública Metropolitana de Agua Potable y Saneamiento (EPMAPS, 2022), Fondo para la Protección del Agua (FONAG, 2022) y del Instituto Nacional de Meteorología e Hidrología (INAMHI, 2022), como se aprecia en la Figura 3.2(A), además se consideró el mismo espacio temporal que el tiempo de registro de JTU\_AWS; con el propósito de caracterizar el régimen pluviométrico dominante de la zona de estudio, considerando el comportamiento estacional presente en el Distrito Metropolitano de Quito y sus principales sistemas de abastecimiento hídrico.

**Tabla N° 3.3:**

**Ubicación de estaciones y variables hidrometeorológicas usadas para la caracterización de la estacionalidad de la precipitación.**

Código	Nombre	X WGS84 TMQ	Y WGS84 TMQ	Altura (m)	Variables
JTU_AWS	Jatunhuayco	806821.00	9946064.00	4142	P, T, HR, RS
C01	Maucatambo	517419.37	9924715.44	3840	P
C02	Rumihurco	495826.29	9985646.41	3586	P
C04	Rumipamba	496515.57	9980594.73	3355	P
C05	Bellavista	504265.47	9979906.2	2960	P
C06	Yaruquí	521558.16	9981817.92	2680	P
C07	San Antonio	504966.01	9997963.62	2467	P
C08	IASA	509533.41	9956640.29	2727	P
C10	La Mica Presa	530267.2	9939548.97	3922	P
C11	Pita Campamento	506872.94	9945112.99	3387	P
C12	Quijos Campamento	550490.3	9947685.48	2596	P
C13	Salve Faccha	538701.84	9974481.34	3888	P
C14	Mindo Captación	490186.09	9983470.18	3600	P
C15	Nanegalito	479360.46	10006618.39	1767	P
C16	Guaytaloma	542097.78	9967076.3	3782	P
C18	Puengasí	500769.56	9973818.84	2970	P
C19	El Troje	497372.93	9963120.85	3122	P
CAR02PT01	Carachas	9965739	765827	3575	P
M0116	Chiriboga	746874.48	9976709.72	1857	P
M0211	La Vinna	774581.45	10012884.73	1611	P
M0212	Mindo INECEL	752059.08	9994653.47	1973	P
M0325	García Moreno	764135.52	10025874.09	1456	P
M0326	Selva Alegre	770105.97	10027780.4	1682	P
M0327	Chontal	741151.63	10025808.92	720	P
M0336	Pacto	748916.92	10016192.79	1208	P
M0340	Mindo INAMHI	747954.43	10003019.9	1518	P
M0361	Nono	770170.21	9991887.14	2870	P
M5021	Yurafaccha Oyacachi	543361.34	9979162.32	3710	P
M5026	Cotopaxi Control Norte	506309.31	9937626.99	3670	P
M5027	Lomahurco	482055.36	9936278.54	3727	P
M5031	Chumillos	532237.24	9989513.39	3750	P
M5077	REI	478914.97	9930361.89	3983	P
M5178	Tayango Guagua Pichincha	490243.86	9978969.7	4077	P
M5181	Yuracyacu Guagua Pichincha	490244.32	9978969.62	4220	P
M5182	Mindo Guagua Pichincha	9964050	817507	4334	P
P03	Rumihurco Machángara	497026.08	9985543.76	3245	P
P08	Rumipamba Bodegas	498892.65	9979986.62	3005	P
P09	Iñaquito INAMHI	501372.08	9980269.6	2804	P
P11	Antenas	497326.67	9981633.83	3800	P
P12	Toctiuco	496962.95	9977180.18	3232	P
P13	Cumbaya	507800.34	9976387.82	2339	P
P14	Zámbiza	505970.78	9983673.19	2677	P

*Nota.* P: precipitación, T: temperatura aire, HR: humedad relativa, RS: radiación solar.



**Tabla N° 3.4:**

**Ubicación de estaciones usadas para la caracterización de la estacionalidad de la precipitación.**

Código	Nombre	X WGS84 TMQ	Y WGS84 TMQ	Altura (m)	VARIABLES
P15	El Cinto	492222.74	9972532.36	3273	P
P16	Izobamba	493859.16	9959517.34	3046	P
P17	El Tingo	506337.01	9967507.34	2483	P
P19	La Tola	514419.41	9974348.23	2480	P
P20	Calderón	507421.06	9991397.36	2771	P
P22	Chillogallo	490453.88	9969226.21	3202	P
P23	Atacazo	488666.66	9964786.15	3877	P
P27	San Francisco	495580.29	9977632.63	3551	P
P28	Cruz Loma	495770.84	9979839.05	4012	P
P29	Palmira	485032.59	9973924.91	2673	P
P30	Santa Rosa	488572.27	9971789.71	2913	P
P31	Pichán	492032.35	9987187.85	3516	P
P34	Papallacta	539936.28	9957892.78	3156	P
P38	San Simón	538914.71	9942255.9	4303	P
P39	Yangahuagra	515068.11	9926804.23	3972	P
P40	Tambo 2A	515838.44	9916833.32	3880	P
P41	Guayllabamba	516939.9	9990666.1	2373	P
P42	Ramón Huañuna	533533.37	9933428.27	3691	P
P43	Limboasi	532466.92	9934401.31	3734	P
P45	El Quinche	523254.72	9987683.71	2686	P
P46	Chalpi Grande	546170.21	9960778.58	2905	P
P52	Pintag	513777.27	9959882.66	2814	P
P53	Paluguillo	523709.79	9970416.81	2969	P
P54	El Carmen	515012.27	9949730.5	3283	P
P55	Diguchi	526392.68	9936650.63	3958	P
P56	Tanque Solanda	496576.51	9968833.19	2916	P
P59	Calacalí	498369.85	10000783.43	2894	P
P60	San José de Minas	509664.02	10019182.05	2466	P
P61	Perucho	508705.65	10012229.77	1858	P
P62	Blanco Chico Alto	539591.68	9949890.69	4320	P
P63	La Virgen Papallacta	534308.65	9964564.66	4365	P
P64	Laguna Encantada	547840.09	9968126.57	4000	P
P65	Laguna Santa Lucía	535420.41	9948588.36	4379	P
P66	Blanco Grande	545738.02	9955637.48	3208	P
P67	Relleno El Inga	516046.18	9967465.44	2658	P
P68	Salve Faccha alto	536741.69	9974088.22	3919	P
P70	CC El Bosque	500293.65	9982112.95	2903	P
P71	Collaloma Medio	502983.59	9986436.34	2968	P
P72	Colinas Alto	497432.65	9988624.02	3088	P
P73	Rayocucho	500713.8	9996168.7	2901	P
P75	Laguna los Patos	544539.22	9972457.12	4028	P

*Nota.* P: precipitación.

### **3.2.2 Preprocesamiento y validación de la información**

La información hidrometeorológica de la estación JTU\_AWS, fue analizada considerando las escalas temporales: 5 minutos, 1 hora y 1 día; utilizando el protocolo de validación, desarrollado y aplicado en Páez-Bimos (2020) y sus resultados se utilizaron en la presente investigación, a continuación se presenta un resumen de la metodología aplicada en dicho documento.

#### **3.2.2.1 Validación de pluviometría**

El pluviómetro instalado en la estación JTU\_AWS en Jatunhuayco es de balancín marca y modelo (Texas Electronic, TE252MM), su frecuencia de registro es de 5 minutos; para el control de calidad, se consideró que los registros negativos sean eliminados y reemplazaron por NA, se eliminaron los valores duplicados, se acumuló la precipitación registrada y se comparó con registros de estaciones cercanas detalladas en (Páez-Bimos, 2020), identificando posibles errores en el funcionamiento del pluviómetro, se calculó la matriz de correlación entre pluviómetros cercanos, para “identificar errores en la dependencia lineal” (Ochoa-Tocachi et al., 2018), posteriormente se realizaron las correcciones necesarias a la serie temporal de precipitación en JTU\_AWS.

#### **3.2.2.2 Validación de las variables meteorológicas**

La información de las variable meteorológicas registradas en la estación JTU\_AWS, fueron validadas en Páez-Bimos (2020), donde se compararon con los registros de las estaciones meteorológicas del proyecto ParamoSus AN-CU-UP y AN-TU-UP y las estaciones meteorológicas M5126 y Eddy Covariance (Administrado por la Escuela Politécnica Nacional), la información de M5126 se descargó del sitio web administrado por el FONAG: <http://sedc.fonag.org.ec/>.

La información con frecuencia de registro a 5 minutos fue tratada y validada en Páez-Bimos (2020), donde se siguieron los siguientes pasos:

- Eliminar valores duplicados en la misma ubicación temporal.

- Identificación y cuantificación de vacíos.
- Se identifican y eliminan valores que sobrepasan extremos meteorológicos o registros que no podrían ser medidos por los sensores y los que sean mayores o menores a tres veces el rango inter-cuantil.
- Eliminación de de valores donde las diferencias consecutivas entre registros en valor absoluto, sobrepasen o sean menores a los rangos de valores establecidos, según los criterios de Zahumenský (2004).
- Se realizaron gráficos de dispersión entre las mismas variables de distintas estaciones, donde se incluyen los coeficiente de determinación  $r^2$ .

Por otro lado, la información de las estaciones de las Tablas 3.3 y 3.4, ha sido validada por las instituciones a las cuales pertenecen, en este caso EPMAPS, FONAG e INAMHI.

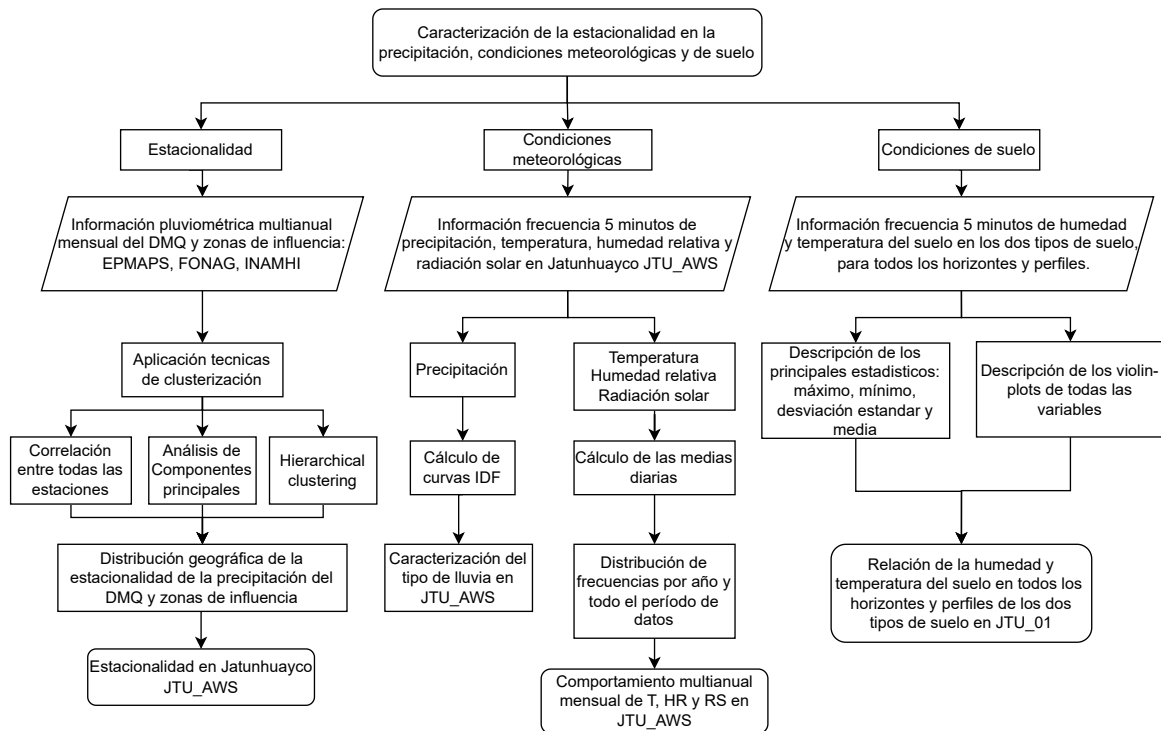
### 3.2.2.3 Validación de la humedad del suelo

Los instrumentos utilizados para medir el contenido volumétrico del agua fueron los reflectómetros (CS650, Campbell Scientific), cuyo rango de medida es de 0 a 100 %, donde la exactitud del contenido volumétrico del agua es  $\pm 1\%$  (con calibración específica del suelo) y  $\pm 3\%$  (típico con el modelo de fábrica), con una precisión  $<0,05\%$ , según los manuales Scientific (2017).

La información fue validada y calibrada utilizando muestras de suelo no alteradas y recolectadas en campo, mismas que fueron procesadas en el laboratorio LEMSUR de la Escuela Politécnica Nacional en Quito, donde las mismas se dejaron secar mientras se extraían periódicamente tres submuestras y los resultados se registraron en un datalogger (CR300, Campbell Scientific), todo este trabajo se lo realizó en (Páez-Bimos, 2022), de donde se obtubieron 6 ecuaciones de calibración para los diferentes horizontes analizados, gracias a este último trabajo se puede manifestar que la exactitud del contenido volumétrico del agua corresponde al  $\pm 1\%$ .

### 3.3 Descripción de la climatología y humedad del suelo

En la Figura 3.4, se describe la metodología para esta sección, donde se utilizaron herramientas estadísticas que permiten ubicar temporal y espacialmente, la distribución de la lluvia en el DMQ y sus principales sistemas de abastecimiento hídrico, además, se identificó la estacionalidad con respecto a la pluviosidad, así como, la forma usual de la lluvia con respecto a la intensidad y duración de diferentes eventos de lluvia. Por otro lado, se muestra también la metodología para analizar la distribución temporal de los parámetros hidrometeorológicos: temperatura, humedad, radiación solar, así como, humedad y temperatura del suelo de JTU\_01 (Jatunhuayco), es decir, establecer el comportamiento de las variables analizadas considerando la distribución de las frecuencias de cada variable de estudio.



**Figura N° 3.4:** Metodología para construir la distribución geográfica de la precipitación multianual mensual en el DMQ y su área de influencia, y caracterización de las condiciones meteorológicas en la estación JTU\_AWS y de suelo en la microcuenca Jatunhuayco JTU\_01.

#### 3.3.1 Análisis exploratorio de datos en la microcuenca JTU\_01 y en la estación JTU\_AWS

En esta etapa, se utilizarán estadísticos que resuman el comportamiento de la información que previamente ha sido validada, así mismo, se emplearán herramientas gráficas que

ayuden a interpretar y comprender patrones, relaciones o plantear preguntas que con métodos estadísticos más formales puedan ser comprobados, como se indica en Jones (1987), los estadísticos más usuales y que fueron utilizados en las variables, son los siguientes: media, máximo, mínimo, desviación estándar y los tres principales cuartiles, en todas las variables registradas. Además, se emplearon herramientas visuales como Violín-Plots que faciliten la descripción de la distribución de la información recopilada.

### ***3.3.2 Distribución geográfica de la precipitación multianual mensual en el DMQ y zonas de influencia***

Para análisis espacial de la información de precipitación se utilizaron técnicas de clusterización, considerando los datos mensuales de la estación JTU\_AWS y la información de lluvia de las redes de monitoreo hidrometeorológico de instituciones como EPMAPS, FONAG e INAMHI, las cuales se encuentran distribuidas dentro de las zonas urbanas, zonas de conservación hídrica y principales sistemas de abastecimiento hídrico del Distrito Metropolitano de Quito, en la Figura 3.2(A), se puede ver la distribución geográfica de las estaciones analizadas. La idea es identificar grupos homogéneos de lluvia, es decir, áreas donde el comportamiento estacional de la precipitación sea similar, en trabajos como (Ruiz-Hernández et al., 2021), se evidencia que la precipitación en los valles altoandinos de Ecuador puede ser muy variable en distancias cortas, es por eso que, utilizamos métodos de homogeneización, aplicando distintas métricas (Kreyszig, 1991), las cuales ayudan a comparar las series temporales analizadas.

#### **3.3.2.1 Hierarchical clustering aplicado a precipitación**

Las técnicas para dividir objetos en grupos homogéneos, como el “agrupamiento jerárquico y aglomerativo, están bien establecidas dentro de la clasificación no supervisada, teniendo un gran crecimiento en muchas áreas de la ciencia e ingeniería” (Johnson, 1967), estos esquemas surgen de la “partición de conjuntos de datos, fusionando paso a paso los más cercanos entre sí, obteniendo grupos homogéneos” (Müllner, 2011), para lo cual se usan diferentes métricas. Algunos de los métodos más conocidos son: complete, average - UPGMA,

weighted - WPGMA, Ward, centroid - UPGMC and median - WPGMC, como se puede ver en Müllner (2011), a continuación describimos los dos métodos implementados, aplicados a la precipitación.

- **Método de Ward.**- Según Murtagh y Legendre (2014), “es un método aglomerativo jerárquico en el que se minimiza la varianza para cada agrupamiento, utilizando la distancia euclídea”. La distancia inicial para este método se calcula como la distancia euclídea al cuadrado. Para clusters disjuntos  $C_i, C_j$  y  $C_k$  con tamaños  $n_i, n_j$  y  $n_k$  respectivamente se calcula la distancia como:

$$d(C_i \cup C_j, C_k) = \frac{n_i + n_k}{n_i + n_j + n_k} d(C_i, C_k) + \frac{n_j + n_k}{n_i + n_j + n_k} d(C_j, C_k) - \frac{n_k}{n_i + n_j + n_k} d(C_i, C_j).$$

La función a minimizar en el método de Ward es:

$$\delta(C_i, C_j) = \frac{|C_i||C_j|}{|C_i + C_j|} \|C_i - C_j\|^2. \quad (3.1)$$

- **Método Unweighted Pair-Group Method with Arithmetic (UPGMA).**-

Según Gronau y Moran (2007) es un método aglomerativo jerárquico, en el que se toma como la distancia mínima entre elementos:  $d_m = \min d(P_i, P_j)$ , donde  $P_i$  e  $P_j$  son los registros de precipitación correspondientes a dos estaciones. Se calcula la profundidad de la rama (del dendograma), como  $\frac{1}{2} d_m$ . La distancia entre clusters se calcula como:

$$d(C_i, C_j) = \frac{1}{|C_i||C_j|} \sum_{P_i \in C_i} \sum_{P_j \in C_j} d(P_i, P_j),$$

con esta distancia se genera una nueva matriz en la cual se busca una nueva distancia mínima, volviendo a repetir el ciclo.

### 3.3.2.2 Análisis de componentes principales aplicado a precipitación

El método de componentes principales (PCA, por sus siglas en inglés) es una técnica de análisis estadístico que se utiliza para reducir la complejidad de un conjunto de datos multivariados. Este método se utiliza para “encontrar las características más importantes o representativas de un conjunto de datos y representarlos en un espacio de menor dimensión” (Rueda Bayona et al., 2016), aquí se aplica este método a la variable precipitación.

El proceso de PCA comienza con la normalización de los datos, para que todas las variables tengan la misma escala:

$$Z_{ij} = (I_{ij} - \mu_i) / \sigma_i, \quad (3.2)$$

donde  $Z$  es el valor estandarizado (o  $z$  - score), para el índice  $i$  en el tiempo  $j$ .

Luego, se calculan los valores propios y los vectores propios de la matriz de covarianza de los datos. Los vectores propios representan las direcciones principales de la variabilidad en los datos, y los valores propios representan la cantidad de variabilidad explicada por cada vector propio. Estos vectores propios se utilizan para construir una matriz de transformación que se utiliza para proyectar los datos originales en un nuevo espacio de menor dimensión.

Este método se centra principalmente en la descomposición de autovalores de la matriz de covarianza.

$$COV(X) = \frac{X^T X}{n - 1},$$

Donde los valores propios miden la cantidad de varianza capturada:  $COV(X)P_a = \lambda_a P_a$ ,

$$\sum_{a=1}^m \lambda_a = 1.$$

### 3.3.3 Análisis de la intensidad duración frecuencia de eventos de lluvia

En esta subsección analizamos el comportamiento de la lluvia con respecto a la duración de los eventos más intensos registrados en la zona de estudio, así, se puede inferir cual es el tipo de precipitación más usual que se presenta, es decir, si son más frecuentes por ejemplo, eventos muy intensos de corta duración o eventos de baja intensidad y larga dura-

ción, de esta manera se podría asociar a la respuesta de la humedad del suelo considerando el comportamiento de los eventos de lluvia.

Con los registros a más baja frecuencia de la estación *JTU\_AWS*, se calculan los eventos de duración de 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70 y 75 minutos y 2, 4, 6, 9, 12, 14, 16, 18, 20, 22, 24 y 30 horas, con un algoritmo implementado en el lenguaje de programación R, considerando cada evento como un episodio de lluvia separado temporalmente del resto por una distancia mayor que el intervalo de tiempo que se esté considerando ( $t$ ), siguiendo la metodología de Escobar-González et al. (2022).

### 3.3.3.1 Funciones de Distribución

Se analizan las series de cantidad de precipitación máxima, siguiendo las recomendaciones dadas en Casas-Castillo (2005), para las respectivas duraciones  $t$ , en períodos de 5 minutos a 1800 horas. En cada una de las series se obtiene un ajuste mediante la función de distribución Gamma ya que en Escobar-González et al. (2022), se comprobó mediante los criterios BIC y AIC (Criterio de Información de Akaike) y BIC (Criterio de Información Bayesiano, que es la función de densidad que presenta la mejor bondad de ajuste, además para este estudio resulta beneficiosa, pues está limitada a valores positivos, con asimetría hacia la derecha y su función de densidad es:

$$f(x) = \begin{cases} \frac{\lambda(\lambda x)^{k-1}}{\Gamma(k)} e^{-\lambda x} & \text{if } x \geq 0 \\ 0 & \text{otherwise,} \end{cases} \quad (3.3)$$

donde la función Gamma, según Rincón (2014), se define por:

$$\Gamma(k) = \int_0^{\infty} t^{k-1} e^{-t} dt,$$

siendo  $\lambda$  y  $k$  los parámetros de escala y forma respectivamente de la distribución, que se obtiene a partir de la media  $m$  y la desviación estándar  $\sigma$ , según las ecuaciones:



$$m = \frac{k}{\lambda}, \quad \sigma^2 = \frac{k}{\lambda^2}, \quad (3.4)$$

se puede notar que, la distribución exponencial es un caso particular de la distribución Gamma, si tomamos  $k = 1$ .

“Existe un gran número de aproximaciones empíricas que relacionan la intensidad  $I$ , duración  $t$  y el período de retorno  $T$ ” (Jove et al., 2020), “el periodo de retorno de un evento, indica la rareza o frecuencia de este evento y se define por la inversa de la probabilidad de ocurrencia ” (Mohymont et al., 2004); donde  $N$  es el número de años de información y  $D$  el número total de eventos utilizados para la calibración de la función de distribución.

$$T(\text{years}) = \frac{N}{D[1 - f(x)]}. \quad (3.5)$$

Para el presente estudio, se consideraron los datos de precipitación y duraciones  $t$ , fijando el período de retorno  $T$  en (1, 3, 5, 10, 15, 20, 25) años se determina la ecuación de Sherman:

$$I(t) = \frac{a}{(t+b)^c}. \quad (3.6)$$

Se conoce que el parámetro  $a$  tiene un comportamiento logarítmico, con respecto al período de retorno  $T$ , como se puede ver en Casas-Castillo (2005), entonces se puede escribir  $a = \ln(T) + B$ .

Luego se analizan las series de cantidad de precipitación máxima, siguiendo las recomendaciones descritas en Escobar-González et al. (2022), en períodos de 5 a 1800 minutos. Para cada una de las series se obtiene un ajuste mediante la función de distribución Gamma, donde para calcular la precipitación  $x$ , para una duración  $t$  dada, se utilizó el método de mínimos cuadrados, para resolver ecuaciones no-lineales, de la forma:

$$f(x) = ax^b e^{cx}. \quad (3.7)$$

De esta forma, se pueden caracterizar los eventos de precipitación dentro de la zona de estudio con una ecuación que relaciona la intensidad, duración y frecuencia, para de esta manera, relacionarlos con la humedad del suelo.

### ***3.3.4 Distribución temporal de la información meteorológica***

En este apartado se realizó un análisis gráfico de la información climatológica registrada en la estación JTU\_AWD de la temperatura y humedad relativa, para este propósito se calcularon las medias diarias de las variables analizadas y se construyeron los diagramas de la distribución de frecuencias para cada mes, en todos los años analizados. El objetivo fue describir cual es la distribución de las variables en el tiempo, posteriormente se realiza el mismo procedimiento pero considerando todo el set de datos y se realiza el mismo gráfico de la distribución de frecuencias de cada uno de los doce meses del año, es decir graficamos la distribución de frecuencias de todos los eneros, febreros, etc. De esta manera se puede describir el comportamiento multianual mensual de la temperatura y humedad relativa.

Adicionalmente, en esta sección se calcula el número de días secos mensuales, donde “un día seco se define como un día cuya precipitación acumulada sea menor o igual a 0.2 mm” (Alexander & Herold, 2016), posteriormente se analiza la relación con temperatura y humedad relativa.

## **3.4 Descripción de la relación entre precipitación y humedad del suelo**

Se realizó una descripción visual exploratoria de la relación que existe entre humedad del suelo media y precipitación acumulada, para almohadilla y pajonal en el perfil de la cumbre, para el horizonte A; en las siguientes escalas temporales: horaria, diaria, quincenal y mensual.

Por otro lado, se calcularon y graficaron las diferencias consecutivas horarias de la humedad del suelo del perfil A, para las dos coberturas vegetales estudiadas ( $\Delta\Theta_{CU\_UPA}$  Y  $\Delta\Theta_{TU\_UPA}$ ). Además, se realizó un minucioso análisis de los efectos de la exactitud de los reflectómetros, que podrían tener en las diferencias consecutivas de las series de humedad de suelo, para esto se incluyó un box-plot de series de diferencias consecutivas de humedad del

suelo  $\Delta\Theta_{CU\_UPA}$  y  $\Delta\Theta_{TU\_UPA}$ , y se comparó con los valores de exactitud correspondientes al  $\pm 1\%$ , que según los manuales de Scientific (2017), es la exactitud cuando se realiza una calibración de los sensores.

### 3.5 Diagrama de flujo de la metodología del uso de redes neuronales para pronóstico de humedad del suelo

En la Figura 3.5, se presenta la metodología para pronosticar humedad del suelo utilizando como datos de entrada variables ambientales, utilizando una arquitectura de redes neuronales, que se ajuste a los datos validados e implementar técnicas de transfer-learning, para transferir el aprendizaje de un dominio particular, hacia otro similar (Zhuang et al., 2020). Es decir, se toma el conocimiento aprendido de la Red Base, para el pronóstico de la humedad del suelo en almohadilla, en la parte baja del horizonte A ( $\theta_{CU\_LOA}$ ), para llevarlo al pronóstico de la humedad del suelo en pajonal o almohadilla, para diferentes perfiles y horizontes con dominio similar.

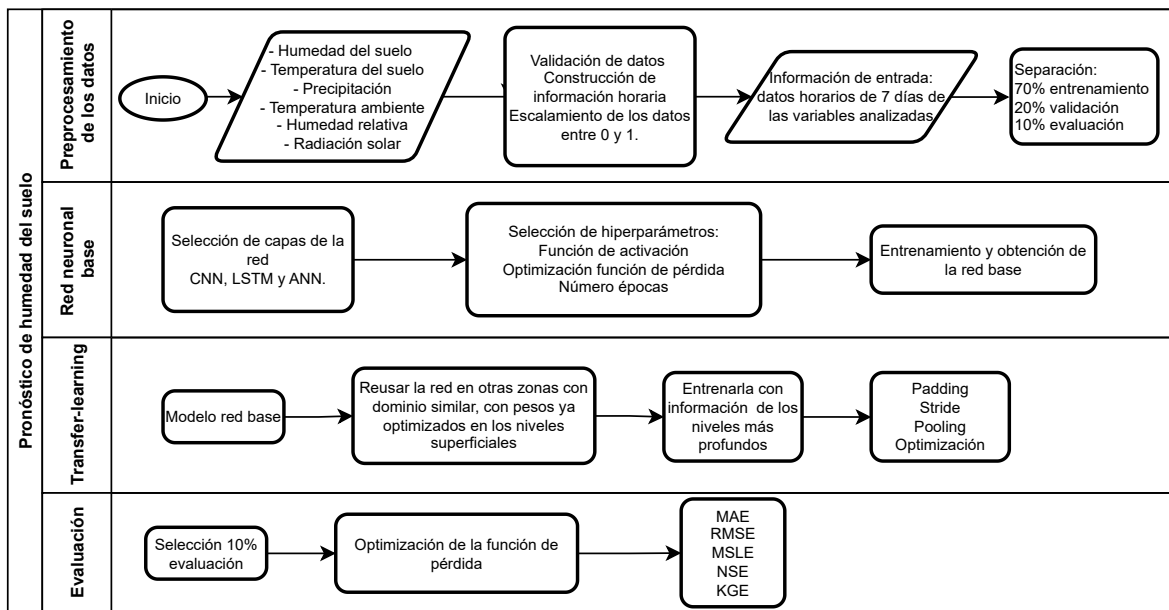


Figura N° 3.5: Diagrama de flujo para pronóstico de humedad del suelo con redes neuronales.

#### 3.5.1 Arquitectura de la Red Neuronal

La red neuronal propuesta para el pronóstico de humedad del suelo, consiste de una arquitectura profunda que aprovecha las características de redes CNN, LSTM y ANN. La

selección de estas redes permite aprovechar las nuevas representaciones generadas por las capas convolucionales, aprender de la temporalidad de los datos capturados por las LSTM y aproximar los valores de humedad de suelo a través del mapeo generalizado de las ANN. Además, para realizar operaciones solo en el segmento temporal de los datos y no tener una interacción aleatoria entre diferentes pasos de tiempo y canales, se utiliza un *envoltorio temporalmente distribuido*, que consiste en una capa que se aplica sobre otra capa de una red neuronal, que permite compartir los parámetros de la capa LSTM en todos los pasos de tiempo. Esto ayuda a la red neuronal a procesar secuencias de diferentes longitudes de manera más eficiente y efectiva.

Para el pronóstico de 48 horas de humedad del suelo, dada los parámetros ecohidro-meteorológicos horarios de siete días, registrados en la estación JTU\_AWS:  $P$ ,  $T$ ,  $HR$ ,  $RS$ ,  $\theta_{CU\_LOA}$ ,  $T_{CU\_LOA}$ , la Tabla 3.5 resume la arquitectura propuesta.

Para el pronóstico de 48 horas de humedad del suelo, dado registros horarios de siete días de las variables:  $P$ ,  $T$ ,  $HR$ ,  $RS$ ,  $\theta_{CU\_LOA}$ ,  $T_{CU\_LOA}$ , la arquitectura propuesta se encuentra en Tabla 3.5. Estas arquitecturas son las redes neuronales base, desde las cuales se obtiene el pronóstico de humedad del suelo en el primer nivel.

Una vez ya entrenado el modelo se obtiene la red base para pronosticar  $\theta_{CU\_LOA}$ , ver Figura 3.6, posteriormente se usa network-based deep transfer learning, es decir, se reusa la red con los pesos ya optimizados en los niveles superficiales para entrenarla con información de los niveles más profundos y de esta manera pronosticar humedad del suelo en las coberturas vegetales: CU, TU; de los perfiles: LO, MI, UP, UR; en los horizontes A y 2A; así por ejemplo, se puede pronosticar humedad del suelo en pajonal en el perfil UR del horizonte 2A ( $\theta_{TU\_UR2A}$ ) y las otras 15 combinaciones posibles.

### 3.5.2 Hiperparámetros

En el entrenamiento, los hiperparámetros fueron ajustados de la siguiente manera:

- **Función de activación**

Tabla N° 3.5:

Arquitectura propuesta para el pronóstico de 48 horas de  $\theta_{CU\_LOA}$  dada información horaria de siete días de:  $P, T, HR, RS, \theta_{CU\_LOA}$  y  $T_{CU\_LOA}$ .

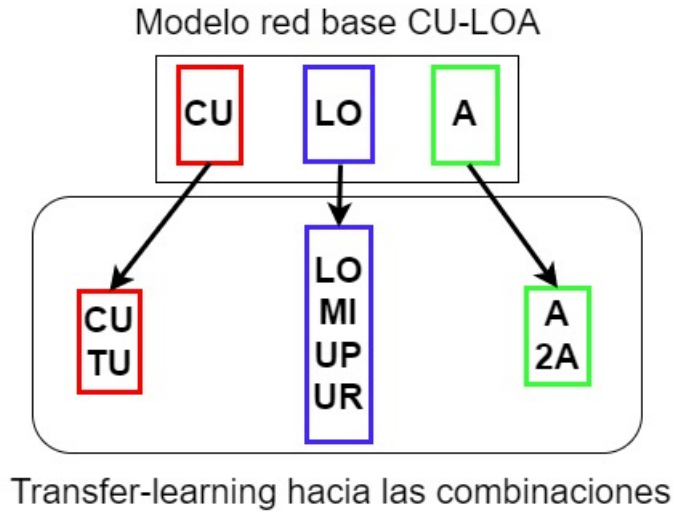
Tipo de capa	Detalles	Salida	# Param
Input Layer	Ventanas de 7 días (168 timesteps) con 5 variables	(168,5,1)	0
Conv2D	32 filtros de (3, 3), activation = 'relu'	(168,5,32)	320
MaxPooling 2D	Pool size = (2, 2)	(84,3,32)	0
Conv2D	32 filtros de (3, 3), activation = 'relu'	(84,3,32)	9248
MaxPooling 2D	Pool size = (2, 2)	(42,2,32)	0
Conv2D	32 filtros de (3, 3), activation = 'relu'	(42,2,32)	9248
TimeDistributedLSTM	32, activation='tanh', recurrent activation='sigmoid'	(42,2,32)	8320
TimeDistributedLSTM	64, activation='tanh', recurrent activation='sigmoid'	(42,2,64)	24 832
Fully Conected ANN	units=100, activation='relu'	(42,2,100)	6500
Fully Conected ANN	units=100, activation='relu'	(42,2,100)	10 100
Flatten	Vectorization	(8400)	0
Fully Conected ANN	units=2048, activation='relu'	(2048)	17 205 248
Fully Conected ANN	units=1024, activation='relu'	(1024)	2 098 176
Fully Conected ANN	units=512, activation='relu'	(512)	524 800
Fully Conected ANN	units=256, activation='relu'	(256)	131 328
Fully Conected ANN	units=48, activation='relu'	48	12 336

- Rectifier linear unit (ReLU) que devuelve un valor mayor a cero:

$$f(x) = \text{máx}(x, 0). \quad (3.8)$$

- Tangente Hiperbólica con valores entre -1 y 1 como:

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \quad (3.9)$$



**Figura N° 3.6:** Red base: pronóstico de humedad del suelo en almohadillas en el perfil A ( $\theta_{CU\_LOA}$ ) y transfer-learning para dominios similares en las 15 combinaciones.

- Función logística o sigmoïdal que resulta valores entre 0 y 1:

$$\sigma(x) = \frac{1}{1 + e^{-x}}. \quad (3.10)$$

- **Función de pérdida**

- La función elegida para ser optimizada es el error cuadrático medio (MSE) definido como:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2. \quad (3.11)$$

- **Optimización**

- Se optimiza los pesos de la red a través de Adaptive learning rate optimization algorithm (ADAM):

$$\begin{aligned} \text{Learning Rate} &= 0.005, \\ \beta_1 &= 0.9, \\ \beta_2 &= 0.999, \\ \epsilon &= 1 \times 10^{-7}, \end{aligned} \quad (3.12)$$

donde  $\beta_1$  y  $\beta_2$  son tasas de decaimiento exponencial para las estimaciones del 1<sup>er</sup> y 2<sup>do</sup> momento; y  $\epsilon$  es una pequeña constante para la estabilidad numérica.

- **Padding**

- Las entradas se ajustan uniformemente con ceros alrededor de modo que la salida tenga la misma dimensión de alto/ancho que la entrada.

- **Stride**

- Se toman pasos de (3,3).

- **Pooling**

- Se toman pools de tamaño (2,2).

### 3.6 Evaluación

La evaluación de la eficiencia del ajuste se realiza utilizando la librería hydroGOF en R, con la función `gof()` y los coeficientes RMSE, NSE, PBIAS y Pearson, según las recomendaciones de Moriasi et al. (2007), que se describen a continuación:

Error cuadrático medio (RMSE), siempre es positivo y mientras esté más cercano a cero, mejor es el ajuste y se calcula como:

$$RMSE = \sqrt{\frac{1}{n_{samples}} \sum_{i=1}^{n_{samples}} (p_{max}(i) - \hat{p}_{max}(i))^2}, \quad (3.13)$$

El coeficiente de eficiencia de Nash-Sutcliffe (NSE) como se indica en Moriasi et al. (2007), se utiliza para evaluar la habilidad predictiva de un modelo e indica el ajuste de datos observados versus los simulados y se define como:

$$NSE = 1 - \frac{\sum_{i=1}^{n_{samples}} (p_{max}(i) - \hat{p}_{max}(i))^2}{\sum_{i=1}^{n_{samples}} (p_{max}(i) - \bar{p}_{max})^2}, \quad (3.14)$$

donde  $\bar{p}_{max}$  es la media de los registros. Se tiene que  $-\infty < NSE \leq 1$  y mientras el  $NSE$  es positivo y más cercano a 1 el modelo es más eficiente, si el  $NSE \leq 0$  se entiende que la media de los valores observados son un mejor predictor que los valores simulados.

El sesgo porcentual *PBIAS* según Yapo et al. (1996): mide la tendencia del promedio de los valores simulados a ser mayores o menores que los observados y se definen por:



$$PBIAS = 100 \frac{\sum_{i=1}^{n_{samples}} (\hat{p}_{max}(i) - p_{max}(i))}{\sum_{i=1}^{n_{samples}} p_{max}(i)}, \quad (3.15)$$

el valor óptimo de  $PBIAS$  es 0, mientras que valores negativos de  $PBIAS$  indican subestimación de la precipitación acumulada máxima simulada y los valores positivos de  $PBIAS$  indican sobrestimación de la precipitación acumulada máxima simulada.

El coeficiente de determinación  $R^2$  según Legates y McCabe Jr (1999), se define por:

$$R^2 = \frac{\sum_{i=1}^{n_{samples}} (p_{max}(i) - \bar{p}_{max}) (\hat{p}_{max}(i) - \bar{\hat{p}}_{max}(i))}{\sum_{i=1}^{n_{samples}} (p_{max}(i) - \bar{p}_{max})^2}, \quad (3.16)$$

y describen el grado de colinealidad entre los datos simulados y registrados e indica la proporción de la varianza de los datos registrados explicada por el modelo, además  $0 \leq R^2 \leq 1$  y los valores de  $R^2$  más cercano a 1 indican una menor varianza de error.

## CAPÍTULO IV: RESULTADOS Y DISCUSIÓN

### 4.1 Análisis de datos

Para el análisis exploratorio de los datos en JTU\_AWD, se consideraron los estadísticos más usuales como: media, máximo, mínimo, desviación estándar y los tres principales cuartiles, que están resumidos en las Tablas 4.1, 4.2 y 4.3. Además, se emplearon herramientas visuales como Violín-Plots Figura 4.1, donde es fácil visualizar si una variable sigue una distribución normal, de esta forma se explica de mejor manera la información recopilada.

Se utilizaron técnicas de clusterización, para encontrar la distribución espacial de la precipitación multianual mensual, en la zona de estudio JTU\_AWD, con respecto al Distrito Metropolitano de Quito y sus principales sistemas de abastecimiento hídrico; mientras que para el ciclo diurno de lluvia, se plotearon los registros horarios desde 2018 hasta octubre 2022 y también se plotó la media horaria anual, para encontrar el comportamiento usual de la lluvia durante el día; además, para analizar el comportamiento de los eventos de lluvia intensa en JTU\_AWD y su posible relación con respecto a humedad del suelo, se construyeron las curvas de intensidad duración y frecuencia. Por otro lado, se analizó también el comportamiento de temperatura y humedad relativa analizando la distribución de frecuencias diarias, para entender la climatología de la zona de estudio, adicionalmente se incluyó el comportamiento de los días secos consecutivos.

#### 4.1.1 Análisis exploratorio de datos climatológicos en JTU\_AWS

En la Tabla 4.1, se describen la precipitación diaria ( $I$ ), de la estación JTU\_AWS y de las otras variables meteorológicas con una frecuencia de 5 minutos: temperatura ambiente ( $T$ ), humedad relativa ( $RH$ ) y radiación solar ( $SR$ ). Si bien es cierto, la variable precipitación es muy compleja de analizar solamente con los estadísticos utilizados en la Tabla 4.1, sin embargo, que el valor del percentil  $P_{25}$ , sea cero y que el valor del percentil  $P_{75}$  sea 2.6, muestra por simple inspección que la intensidad es baja, así como se ve en el Violín-Plot de  $I$ , de la Figura 4.1, además, se puede apreciar que el registro máximo de lluvia diaria es 40

mm, esto indica que la precipitación en la zona de estudio no es muy intensa en comparación a los registros de otras zonas del DMQ, como se puede ver en Escobar-González et al. (2022), en las siguientes secciones se estudia más a detalle el comportamiento de la lluvia a baja frecuencia, diaria y mensual.

Por otro lado, los registros de temperatura del aire muestran que el mínimo registrado es  $-5.48\text{ }^{\circ}\text{C}$ , además los percentiles encontrados  $P_{25}$ ,  $P_{50}$  y  $P_{75}$ , indican que en general la temperatura es inferior a los  $6.09\text{ }^{\circ}\text{C}$ , mientras que la temperatura máxima es  $14.1\text{ }^{\circ}\text{C}$ , es decir, coincide con la temperatura fría alto andina usual de páramo, como se puede ver en Sklenar et al. (2015), además, esto se reafirma observando que la mayoría de información de  $T$  graficada el violin-plot de la Figura 4.1, se encuentra distribuida entre los 0 y  $10\text{ }^{\circ}\text{C}$ .

**Tabla N° 4.1:**

**Resumen estadístico de la lluvia diaria y de las variables meteorológicas con frecuencia de 5 minutos en JTU\_AWD Jatunhuayco.**

	$I(\text{mmd}^{-1})$	$T\text{ (}^{\circ}\text{C)}$	$RH\text{ (}\%\text{)}$	$SR\text{ (Wm}^{-2}\text{)}$
count	1409	418168	414135	404179
mean	2.1	4.12	93.72	138.01
std	3.45	2.72	8.52	213.09
min	0	-5.48	58	0
25%	0	2.13	89.9	0
50%	0.5	3.59	97.8	0.76
75%	2.6	6.09	100	224
max	40	14.1	100	960

De los resultados de la Tabla 4.1, se puede observar que la media de los registros de humedad relativa es  $93.72\%$  y todos los percentiles analizados superan el  $90\%$ , es decir, predomina un ambiente húmedo usual del páramo de los Andes del norte de Ecuador, lo cual concuerda con estudios como los de Bendix et al. (2001) y Sklenar et al. (2015), además, estos resultados se confirman con la información de humedad relativa, graficada en el violin-plot de la Figura 4.1, donde los valores predominantes son superiores al  $90\%$ .

Los resultados de la Tabla 4.1, para la variable radiación solar, pueden conducir a conclusiones no muy acertadas y esto se debe a que en horas de la noche los registros son cero,

evidentemente por la ausencia de luz solar, por otro lado, la información de  $SR$ , graficada el violin-plot de la Figura 4.1, muestra que la radiación solar se encuentra distribuida la mayor parte del tiempo, entre 0 y  $500 \text{ Wm}^{-2}$ , sin embargo el máximo registro es de  $960 \text{ Wm}^{-2}$ .

#### 4.1.2 Análisis exploratorio de datos de temperatura y humedad del suelo en JTU\_01

##### 4.1.2.1 Temperatura del suelo en JTU\_01

Según la Tabla 4.2, se puede apreciar que para todos los perfiles en almohadilla, la temperatura media del horizonte  $A$  oscila entre  $[6.55 < T_{CU\_A} < 7.06]^\circ C$ , mientras que para el horizonte  $2A$ , lo hace entre  $[6.44 < T_{CU\_2A} < 7.21]^\circ C$ . Por otro lado, para pajonal según la Tabla 4.3, la temperatura media en el horizonte  $A$  oscila entre  $[6.48 < T_{TU\_A} < 7.54]^\circ C$ , mientras que para el horizonte  $2A$ , lo hace entre  $[6.1 < T_{TU\_2A} < 6.86]^\circ C$ .

Un primer resultado que se encuentra de las Tablas 4.2 y 4.3, es que en todos los perfiles y dos tipos de coberturas vegetales analizadas, la temperatura media en el horizonte  $A$  es mayor que la del horizonte  $2A$ , excepto en la almohadilla para el horizonte upper-UP, donde la temperatura media en el horizonte  $A$  es  $7^\circ C$ , mientras que en  $2A$  es  $7.21^\circ C$ , esto quizá podría justificarse por su ubicación en la cumbre y la influencia del viento en esa zona.

**Tabla N° 4.2:**

**Descripción estadística de la temperatura de suelo ( $^\circ C$ ) en los perfiles A y 2A, en almohadilla.**

	$T_{CU\_UPA}$	$T_{CU\_UP2A}$	$T_{CU\_URA}$	$T_{CU\_UR2A}$	$T_{CU\_MIA}$	$T_{CU\_MI2A}$	$T_{CU\_LOA}$	$T_{CU\_LO2A}$
count	400761	400759	349770	349890	257577	257563	306167	305165
mean	7	7.21	6.55	6.44	7.06	6.62	6.96	6.6
std	1.37	0.86	1.94	1.61	1.77	1.46	2.47	1.83
min	0.76	2.9	0	0.01	0.6	1.83	0.01	0.22
25%	6.15	6.65	5.27	5.4	5.93	5.67	5.3	5.39
50%	6.95	7.21	6.32	6.34	6.89	6.54	6.64	6.52
75%	7.82	7.72	7.65	7.41	8.03	7.51	8.4	7.76
max	18.76	16.9	23.6	17.1	23.4	20.9	19.4	14.3
$d(P_{25}, P_{75})$	1.67	1.07	2.38	2.01	2.1	1.84	3.1	2.37

*Nota.*  $d(P_{25}, P_{75})$ : diferencia entre los valores de temperatura del percentil 25 y su correspondiente percentil 75.

**Tabla N° 4.3:**

**Descripción estadística de la temperatura de suelo ( $^{\circ}C$ ) en los perfiles A y 2A, en pajonal.**

	$T_{TU\_UPA}$	$T_{TU\_UP2A}$	$T_{TU\_URA}$	$T_{TU\_UR2A}$	$T_{TU\_MIA}$	$T_{TU\_MI2A}$	$T_{TU\_LOA}$	$T_{TU\_LO2A}$
count	405498	405496	399437	399437	392222	392221	351712	350776
mean	6.48	6.1	6.86	6.41	7.54	6.86	7.27	6.59
std	1.82	1.51	1.68	1.31	1.89	1.22	2.56	1.7
min	0.79	0.74	1.79	1.81	0.39	1.11	0.34	1.21
25%	5.27	5.06	5.75	5.54	6.33	6.08	5.57	5.44
50%	6.21	5.96	6.6	6.33	7.21	6.8	6.66	6.39
75%	7.37	6.94	7.71	7.18	8.43	7.6	8.51	7.59
max	20.9	18.5	15.57	15.56	16.91	14.45	20.9	19
$d(P_{25}, P_{75})$	2.1	1.88	1.96	1.64	2.1	1.52	2.94	2.15

*Nota.*  $d(P_{25}, P_{75})$ : diferencia entre los valores de temperatura del percentil 25 y su correspondiente percentil 75.

Un segundo resultado se tiene al utilizar en todos los perfiles  $d(P_{25}, P_{75})$ , que es la diferencia entre los valores de temperatura del percentil 25 y su correspondiente percentil 75 y comparar los resultados entre los horizontes A y 2A, por ejemplo, para  $T_{CU\_UPA}$  se tiene  $1.67^{\circ}C$  y para  $T_{CU\_UP2A}$   $1.07^{\circ}C$ , lo cual indica que existe una mayor variación en los valores de temperatura en el perfil A que en el 2A y esto se verifica para todos los casos y resulta coherente, ya que a mayor profundidad existen menos factores externos que influyan sobre la variación de la temperatura, además la desviación estándar (*std*) sigue un comportamiento similar, lo cual ratifica este resultado.

#### 4.1.2.2 Humedad del suelo en JTU\_01

Para este análisis, en primer lugar se observa los violin-plots de las filas 2 a 5, de la Figura 4.1, donde en todos los casos no se puede asociar directamente una distribución de los datos, es decir, se necesita hacer un análisis más detallado para describir la variable humedad del suelo o quizá en un futuro con mayor cantidad de información se pueda hacer un mejor análisis. Después se analizó el comportamiento de la humedad del suelo, considerando la variación entre registros consecutivos y la posible relación que existe con la precipitación registrada.

Resulta interesante observar que en la Tabla 4.4, los valores medios de humedad del suelo para almohadilla en el horizonte A siempre son mayores que en el horizonte 2A, además, también son mayores que los valores medios de humedad del suelo del pajonal en los

dos horizontes A y 2A; de esta última observación, se tiene que no existe una relación directa entre la humedad del suelo del pajonal en los dos horizontes A y 2A.

**Tabla N° 4.4:**

**Descripción estadística de la humedad del suelo (%), en los perfiles A y 2A, en almohadilla.**

	$\Theta_{CU\_LOA}$	$\Theta_{CU\_LO2A}$	$\Theta_{CU\_MIA}$	$\Theta_{CU\_MI2A}$	$\Theta_{CU\_UPA}$	$\Theta_{CU\_UP2A}$	$\Theta_{CU\_URA}$	$\Theta_{CU\_UR2A}$
count	306018	305020	253007	257375	400587	400525	349564	349377
mean	0.63	0.5	0.67	0.55	0.63	0.57	0.64	0.51
std	0.03	0.03	0.02	0.03	0.02	0.02	0.02	0.02
min	0.57	0.42	0.59	0.46	0.58	0.51	0.58	0.45
25%	0.61	0.48	0.66	0.53	0.61	0.56	0.63	0.5
50%	0.63	0.5	0.67	0.55	0.63	0.57	0.64	0.51
75%	0.65	0.52	0.68	0.56	0.64	0.59	0.65	0.52
max	0.67	0.56	0.71	0.6	0.67	0.63	0.67	0.56

**Tabla N° 4.5:**

**Descripción estadística de la humedad del suelo (%), en los perfiles A y 2A, en pajonal.**

	$\Theta_{TU\_LOA}$	$\Theta_{TU\_LO2A}$	$\Theta_{TU\_MIA}$	$\Theta_{TU\_MI2A}$	$\Theta_{TU\_UPA}$	$\Theta_{TU\_UP2A}$	$\Theta_{TU\_RA}$	$\Theta_{TU\_UR2A}$
count	350589	350623	391963	391983	405277	405154	399231	399258
mean	0.59	0.61	0.59	0.62	0.6	0.64	0.59	0.5
std	0.02	0.02	0.01	0.01	0.01	0.02	0.02	0.02
min	0.54	0.56	0.56	0.57	0.55	0.16	0.54	0.16
25%	0.58	0.6	0.59	0.61	0.59	0.62	0.58	0.49
50%	0.59	0.62	0.6	0.62	0.6	0.64	0.59	0.51
75%	0.6	0.63	0.6	0.63	0.6	0.65	0.6	0.52
max	0.63	0.66	0.62	0.66	0.63	0.68	0.63	0.56

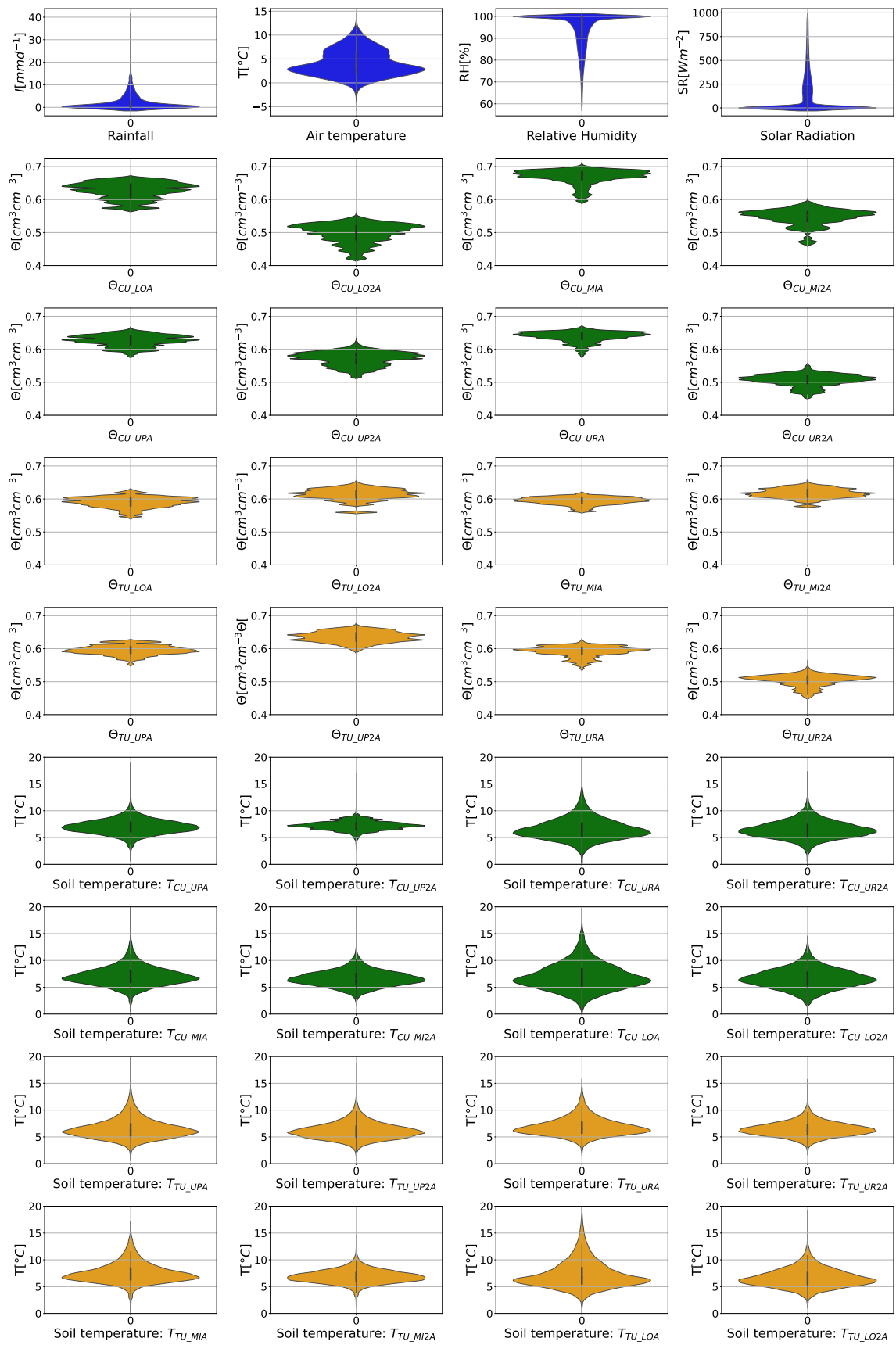


Figura N° 4.1: Violin-plots de las variables meteorológicas en JTU\_AWS y de suelo en la microcuenca Jatunhuayco JTU\_01.

### 4.1.3 Distribución geográfica estacional de la lluvia

Para estudiar la distribución espacial de la precipitación de la microcuenca JTU\_01 Jatunhuayco, se consideró la información mensual de las estaciones pertenecientes a las redes de monitoreo de instituciones como EPMAPS, FONAG e INAMHI, descritas en las Tablas 3.3 y 3.4, las cuales se encuentran distribuidas dentro del Distrito Metropolitano de Quito y sus principales sistemas de abastecimiento hídrico, donde se aplicaron diferentes técnicas de homogeneización entre las series de precipitación mensual de cada estación.

En la Figura ??(a), se presenta un heat-map de la matriz de correlaciones entre las 83 estaciones analizadas, se seleccionó manualmente 2 como el número de clusters a clasificar, principalmente porque la lluvia responde a influencia: amazónica, ZCIT de la costa del pacífico norte, y una mixtura entre las dos que se forma principalmente en la zona interandina, para identificar las estaciones que se encuentran distribuidas dentro de un grupo de similares características, y de esta manera encontrar una primera distribución de grupos cuyas estaciones tienen información homogénea; se debe considerar que en este método de agrupamiento las relaciones entre estaciones, se encuentran determinados por valores que están dentro del intervalo  $[-1, 1]$ , donde se describe el grado de colinealidad entre estaciones y mientras su valor sea más cercano a 1, se presenta una menor varianza, es decir, el comportamiento es más parecido.

Si bien es cierto el número de grupos la Figura 4.2 fue seleccionado manualmente, esto se justifica por la presencia de dos comportamientos predominantes: uno asociado al régimen estacional amazónico y que corresponde principalmente a las estaciones ubicadas dentro de la zona de páramos altoandinos cercanos a zonas como Papallacta y Antisana (estaciones: P34 Papallacta, C13 Salve Faccha, etc), con un comportamiento unimodal, con máximos entre los meses de mayo a agosto, mientras que el segundo comportamiento se lo asocia a un régimen interandino bimodal, usual dentro de la ciudad de Quito, cuyos valores máximos se encuentran en los meses de abril y octubre, además se presenta una época bien definida de bajas precipitaciones entre los meses de julio a agosto, como se puede ver en Ruiz-Hernández et al. (2021), y una pequeña recesión entre los meses de diciembre a enero.



Llama la atención que dentro del mismo clúster que se encuentra la estación JTU\_AWS Jatunhuayco, están además las estaciones: P38 San Simón, P40 Tambo2A y P65 Laguna Santa Lucía, más adelante se verá a detalle, cómo es el comportamiento estacional de la zona de estudio, y cómo influyen los principales sistemas atmosféricos.

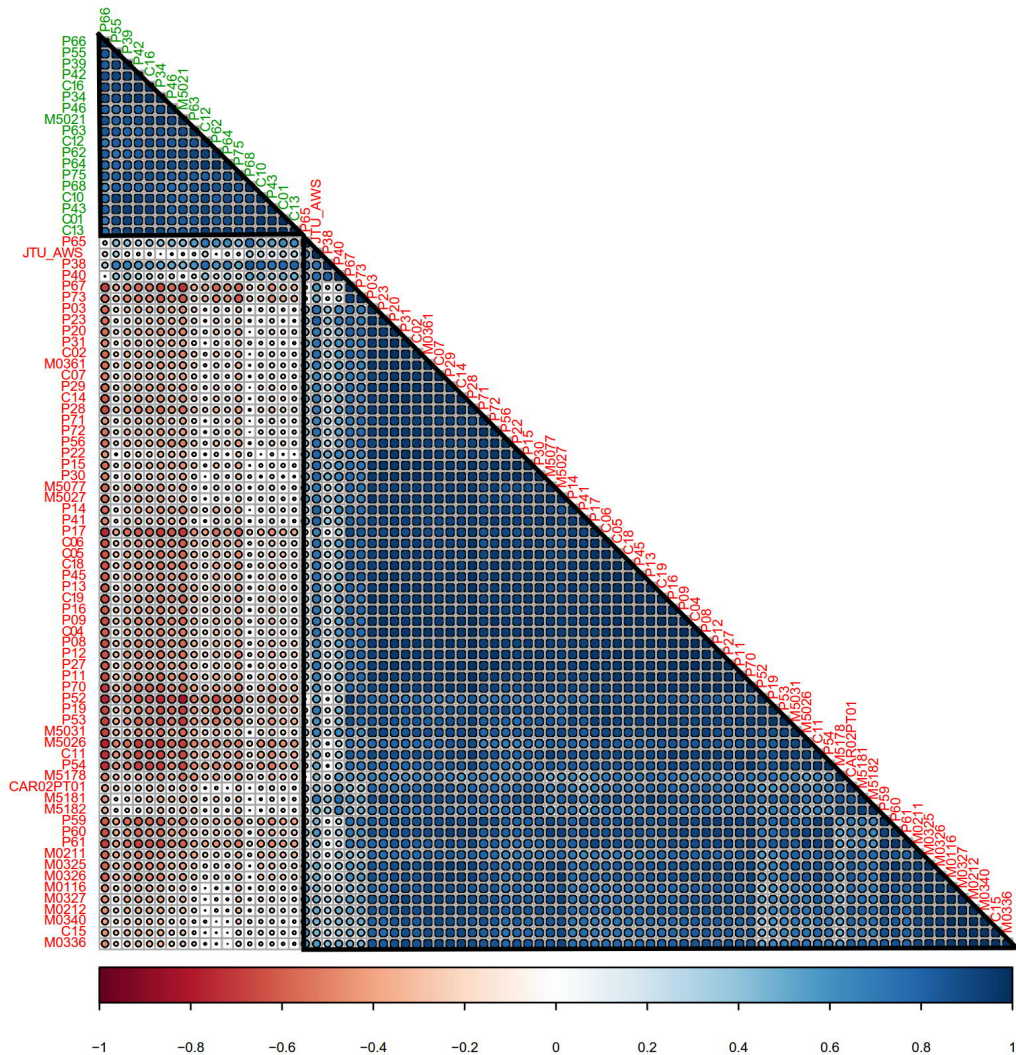
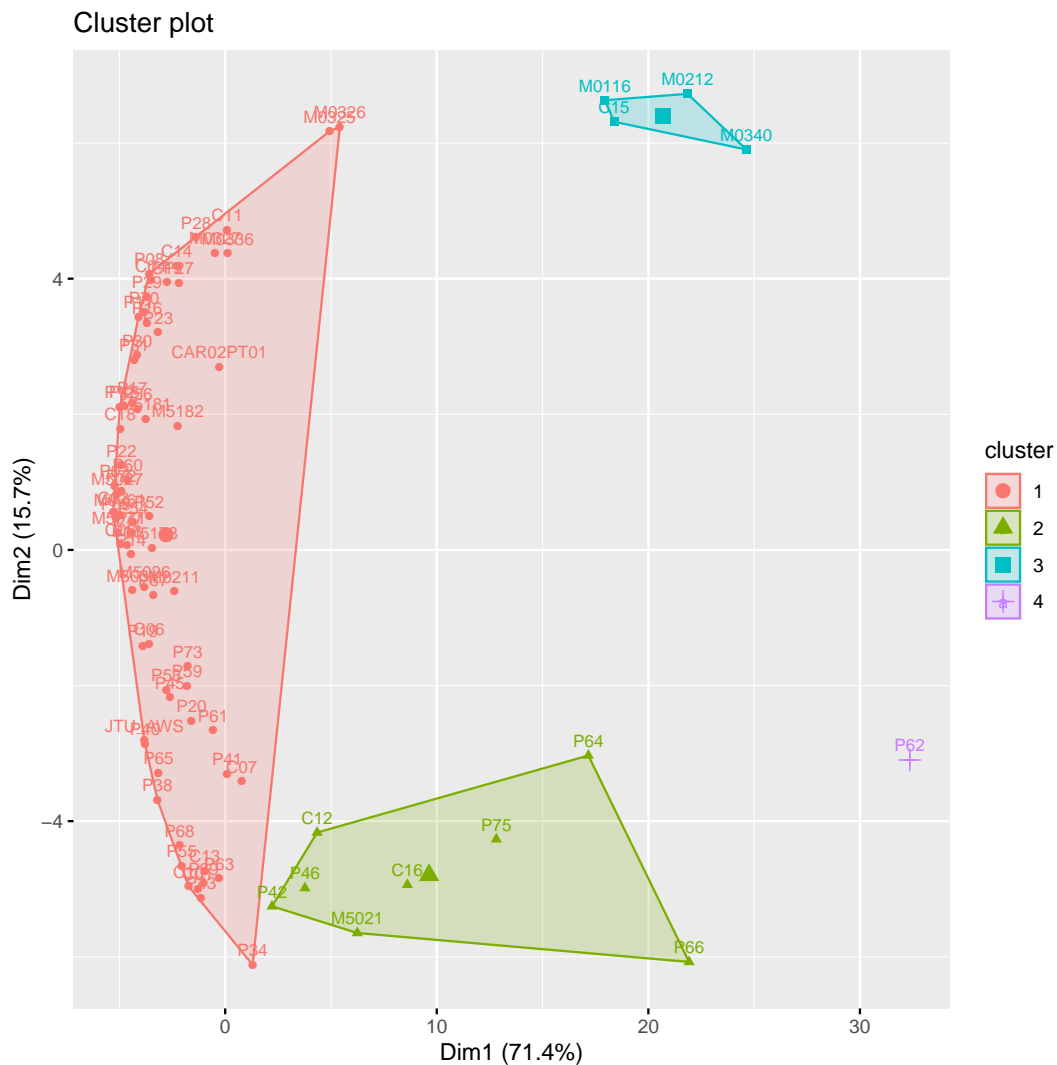


Figura N° 4.2: Visualización de la correlación entre de las serie de precipitación mensual de las estaciones seleccionadas en las Tablas 3.3 y 3.4.

En la Figura 4.3, se pueden apreciar las componentes principales (PC) donde el PC1 y PC2, en conjunto describen el 87.1% de la varianza de los Grupos 1 y 2 (G1 y G2) respectivamente, es decir, estos componentes explican un 87.1% del comportamiento de la información de la red de estaciones analizadas, se seleccionaron manualmente la visualización

de cuatro grupos, en el grupo (G3), con estaciones como: C15 Nanegalito o M0212, es decir, estaciones que se encuentran en en el Noroccidente del DMQ y que al parecer tienen un comportamiento que caracteriza esta zona, en el Grupo 4 (G4) se nota solamente a una estación: P62, con un comportamiento modulado principalmente por la Amazonía, pero que no se asocia fácilmente a los demás clusters por la gran cantidad de lluvia registrada.



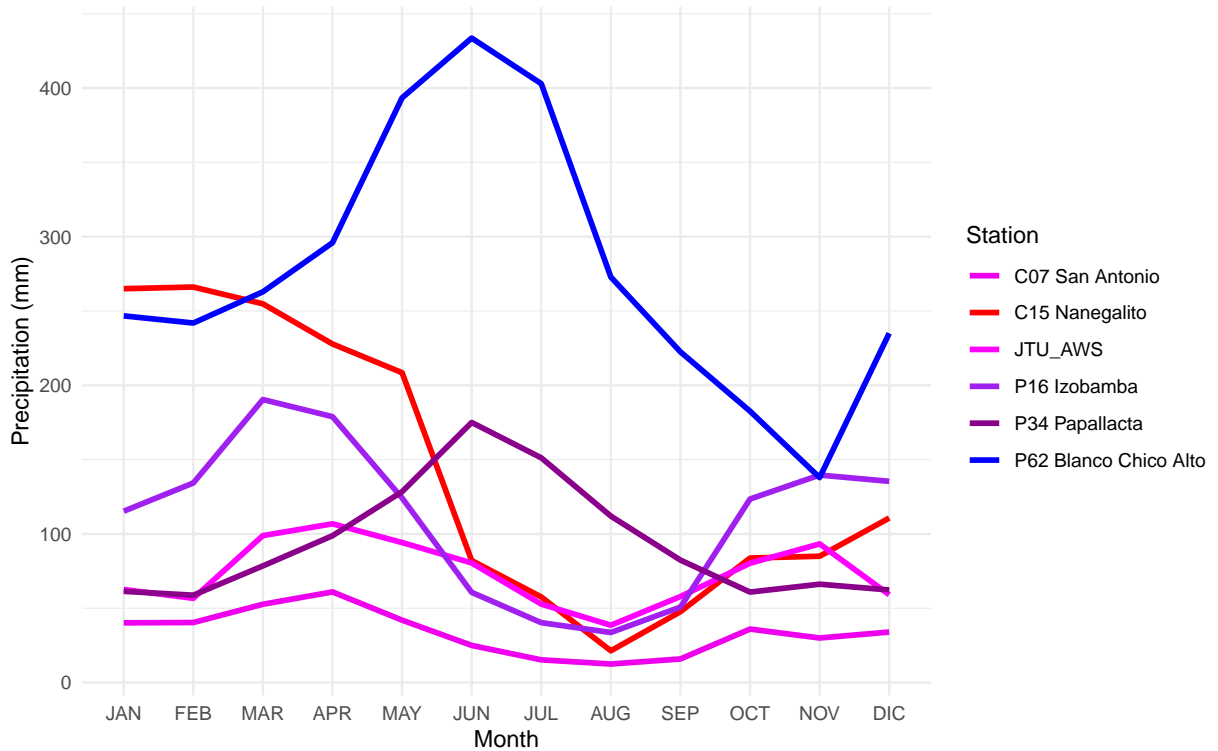
**Figura N° 4.3:** Visualización de las componentes principales de precipitación mensual de las estaciones seleccionadas en las Tablas 3.3 y 3.4..

En la Figura 4.4, se puede apreciar el resultado de aplicar el agrupamiento no supervisado jerárquico aglomerativo, para el conjunto de estaciones analizado, utilizando el método de Ward, descrito en la metodología y utilizando la distancia euclidiana, como métrica de comparación, la presentación de resultados es fácil comprender aunque son 83 estaciones,



- La estación C07 San Antonio ubicada en la Mitad del Mundo posee un comportamiento bimodal, con sus picos máximos en abril y octubre con una precipitación media anual de 389 mm.
- La estación C15 Nanegalito tiene un comportamiento usual del régimen de la costa del pacífico, con bajas precipitaciones con respecto a las medias mensuales desde junio a noviembre y sus valores máximos entre los meses de diciembre a mayo, donde su precipitación media anual aproximada es 1958 mm.
- La microcuenca de estudio JTU\_01 Jatunhuayco presenta un comportamiento bimodal, con sus máximos en abril y noviembre, con sus valores más bajos entre los meses de julio a septiembre y de diciembre a febrero, con una media anual aproximada de 881 mm.
- La estación P16 Izobamba presenta un comportamiento casi bimodal, donde en el año hidrológico la época de lluvias comienza en septiembre, con su primer máximo en noviembre y una ligera atenuación de diciembre a enero, para continuar con el incremento de la precipitación hasta alcanzar su segundo máximo en marzo, luego la época de precipitación más baja, está entre los meses de julio a agosto, con la precipitación media anual de 1327 mm.
- La estación P34 Papallacta tiene un comportamiento pluviométrico unimodal, con su pico máximo en junio y época de baja precipitación entre los meses de octubre a febrero, donde la precipitación media anual es de 1136 mm.
- La estación P62 Blanco Chico Alto, tiene un comportamiento unimodal influenciado directamente por la Amazonía, donde la época de mayor lluvia se registra de marzo a agosto con su pico máximo en junio, con la época de menor precipitación a mediados de noviembre y la precipitación media anual es 3328 mm.

En la Figura 4.6, se presenta la distribución geográfica de la precipitación multianual mensual en el Distrito Metropolitano de Quito y sus principales sistemas de abastecimiento



**Figura N° 4.5:** Comparación de estacionalidades en el Distrito Metropolitano de Quito y sus principales sistemas de abastecimiento hídrico.

hídrico, las imágenes de precipitación mensual de las diferentes zonas de interés se encuentran en la misma escala, para notar las diferencias entre la cantidad de precipitación registrada, los grupos de estaciones homogéneas se presentan con una misma tonalidad.

Es claro que la zona de estudio y sus zonas de influencia, se encuentran ubicados en la cordillera de los Andes, misma que “actúa como una barrera topográfica entre el frío - seco pacífico oriental y la cálida - húmeda región amazónica” (Bendix et al., 1992), influenciando un alto gradiente zonal de precipitaciones dentro de la compleja topografía analizada.

Es interesante notar que la precipitación en las zonas más orientales tiene un comportamiento influenciado desde la Amazonía del norte de Ecuador, por los vientos alisios que provienen del sur-este, llevando consigo la humedad que provoca un régimen pluviométrico estacional predominantemente unimodal, con su máximo de precipitación en junio (Bendix et al., 1992; Ruiz-Hernández et al., 2021).

Mientras que en las zonas más occidentales de la Figura 4.6, se pudo notar que el

régimen estacional corresponde al usual del oriente del pacífico ecuatorial, es decir, que “la época de lluvias comienza en diciembre y se extiende hasta mayo, lo cual está influenciado por la circulación de Walker debilitada” (Segura et al., 2019), misma que “está inducida por la presencia de bajas (parte occidental) y altas (parte oriental) presiones del pacífico sur, con inversiones de carácter regular conocidas como Oscilación del Sur- El Niño (ENSO)” (Segura et al., 2019).

Por otro lado, se encuentra una zona de transición simulando un sándwich, mixtura o mezcla de los regímenes pluviométricos oriental y occidental, descritos en los dos párrafos anteriores respectivamente, y que se encuentra principalmente en los valles interandinos, con un comportamiento estacional pluviométrico cuasi-bimodal, “con sus máximos de precipitación normalmente en abril y octubre, y que está influenciado por la convergencia del sistema Walker debilitado” (Segura et al., 2019) y la “presencia del transporte de humedad hacia el oeste desde la Amazonía ecuatorial” (Bendix et al., 1992; Ruiz-Hernández et al., 2021; Segura et al., 2019).

Resulta llamativo notar que la microcuenca de estudio JTU\_01, así como algunas estaciones ubicadas en el oriente de la Figura 4.6, como: P38, P43, P55; también tienen un comportamiento pluviométrico bimodal, esto podría justificarse principalmente por la barrera orográfica del volcán Antisana, minimizando la influencia pluviométrica proveniente de la Amazonía ecuatorial.

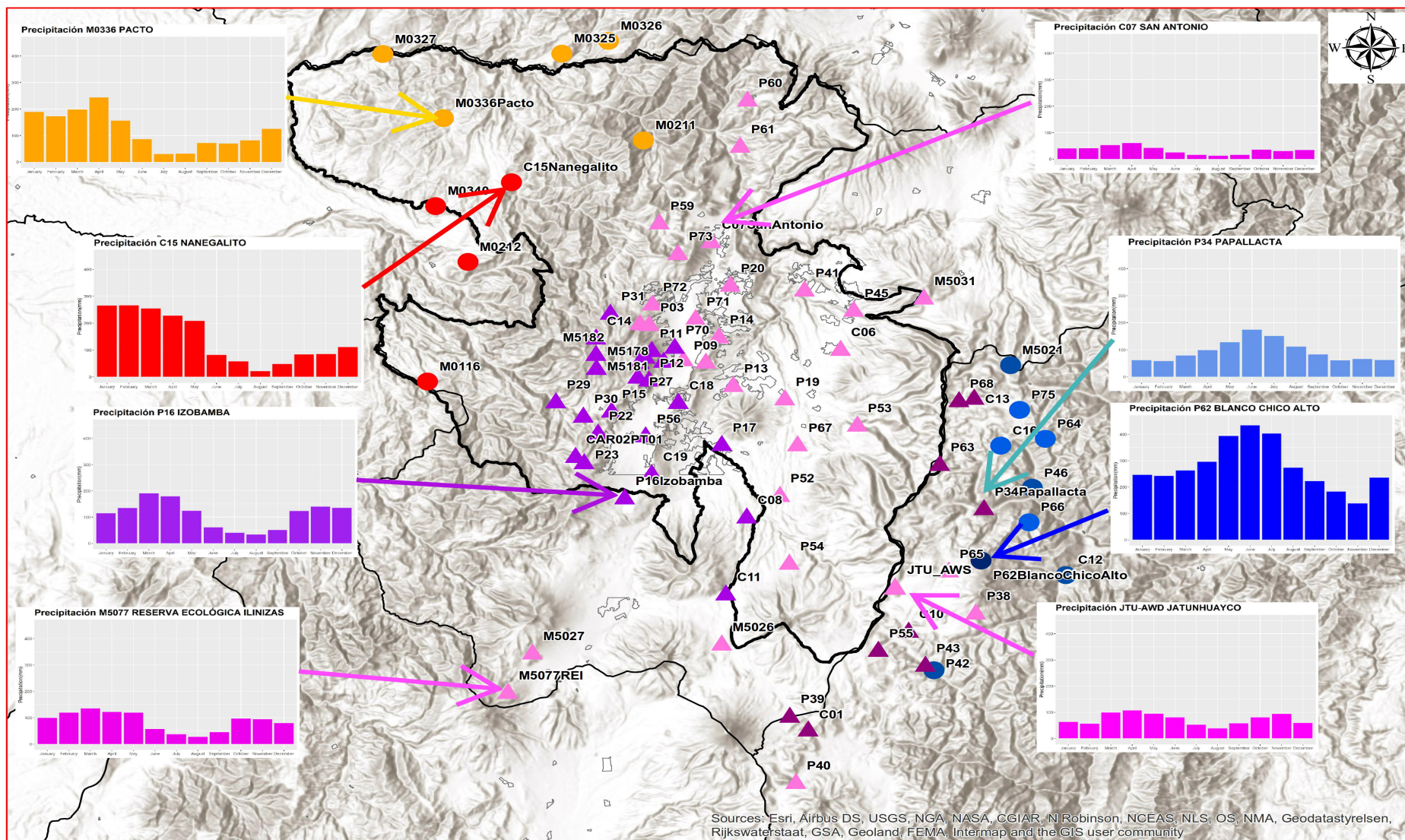


Figura N° 4.6: Tipos de estacionalidad en el DMQ y sus principales sistemas de abastecimiento hídrico, considerando los clusters del dendrograma circular de la Figura 4.4.

#### **4.1.4 Distribución de la precipitación horaria en la estación JTU\_AWS**

En la presente sección, se busca encontrar el momento donde es más frecuente la máxima cantidad de lluvia acumulada horaria, dentro de la microcuenca JTU\_AWS, lo cual podría ayudar a comprender la relación que existe con respecto a la humedad del suelo resultante, para este propósito se construyó la Figura 4.7, donde se presentan varios heatmaps, uno por año, con la información de precipitación acumulada horaria de 2019 a 2022, para finalmente graficar el promedio horario de toda la serie temporal.

La Figura 4.7, relaciona los valores más altos de precipitación con respecto a la paleta de colores asignada, los ejes verticales a la izquierda indican la hora local durante el día, a la derecha el año al cual corresponden y en la parte baja los meses del año subdividido en días.

La distribución horaria de los heat maps de cada uno de los años, no indica un patrón específico. En 2019 se aprecia que la precipitación no es uniforme, pesar de ello, por simple inspección, los valores máximos de lluvia se dan entre las 14h00 y 16h00, excepto entre los meses de junio a agosto, por otro lado, también se aprecian valores considerables de precipitación en la madrugada de los meses de mayo y octubre entre las 02h00 y 06h00.

Un comportamiento un tanto similar se aprecia en el año 2020, donde los valores máximos de precipitación, también se dan en horas de la tarde, aproximadamente entre las 14h00 y 17h00, excepto de marzo a octubre, no obstante se presentan valores altos de lluvia acumulada en horas de la madrugada principalmente entre los meses de abril a agosto.

En 2021 se puede notar que también la precipitación más intensa se da principalmente en horas de la tarde, entre las 12h00 y 17h00, excepto de julio a agosto, donde se presentan valores de mediana y baja intensidad de precipitación durante el día. Se debe notar también, la presencia de algunos registros horarios de mediana intensidad durante la madrugada, principalmente en los meses de junio a julio.

En 2022 la información fue procesada hasta el 7 de octubre de 2022, sin embargo, se puede apreciar que en el mes de enero se registró poca lluvia, posteriormente hay un patrón de los máximos horarios de febrero a julio en horas de la tarde entre las 12h00 y 18h00, además se puede apreciar que entre agosto y septiembre se presentaron lluvias de baja intensidad.



Consolidando la información en el último heat map (Hourly mean) de la Figura 4.7, considerando el promedio de cada hora, durante todos los días de los años analizados, se obtiene un patrón más estable y se puede notar que en los meses de enero y de julio a agosto, no se presentan lluvias muy intensas y las pocas que se presentan se encuentran distribuidas durante todo el día, lo cual coincide en parte con los meses de baja precipitación de la Figura 4.6, donde la precipitación multianual mensual en JTU\_AWS Jatunhuayco es bimodal. Por otro lado, entre los meses de febrero a junio y de septiembre a diciembre se presentan los máximos registros en horas de la tarde aproximadamente entre las 13h00 y 17h00, llama la atención la presencia de lluvias de mediana intensidad considerando la paleta de colores, en horas de la madrugada. Este análisis podría tener un patrón más estable si se contaría con más años de información de lluvia.

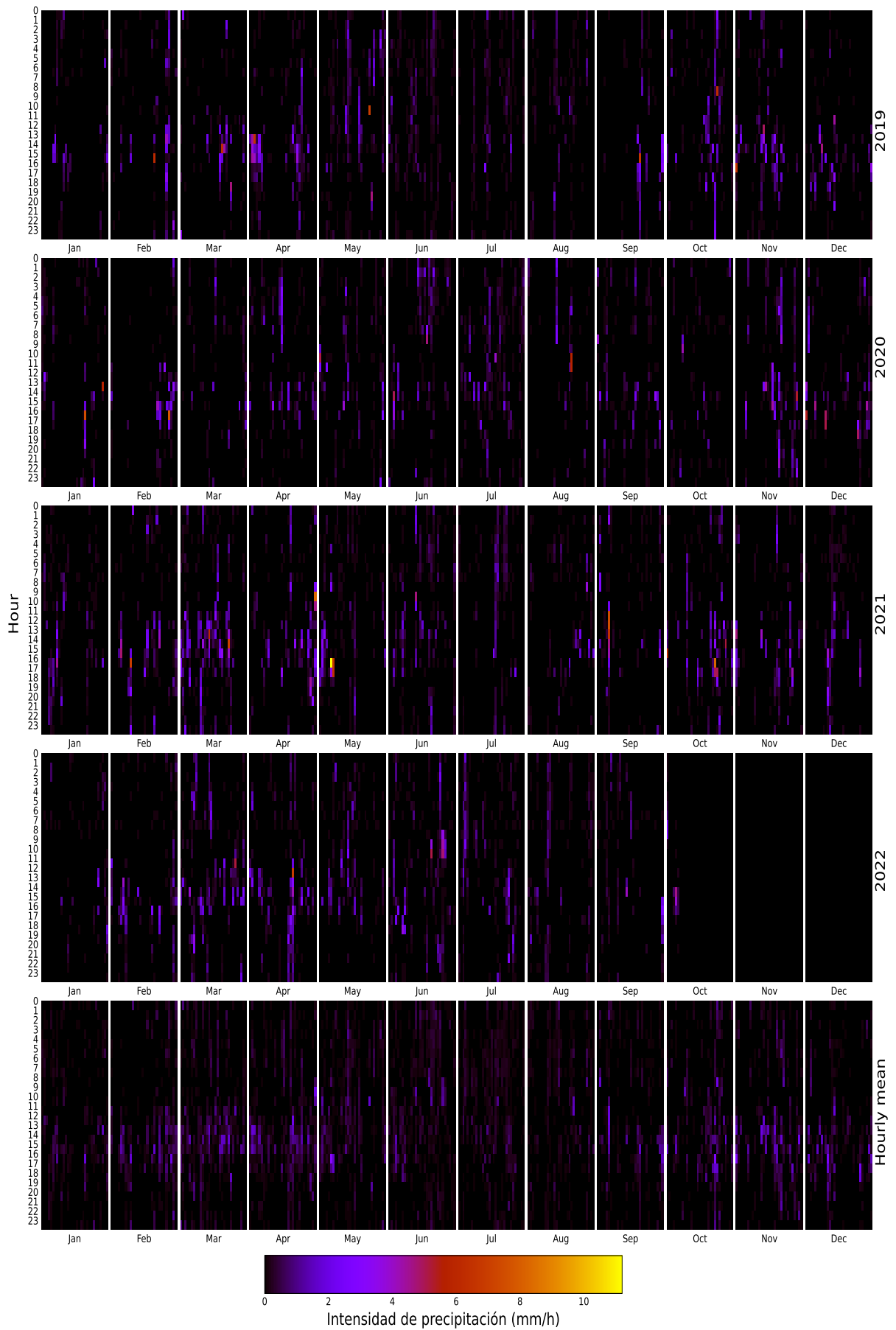


Figura N° 4.7: Precipitación horaria en Jatunhuayco JTU\_AWS, enero 2019 - septiembre 2022

#### 4.1.5 Curvas intensidad duración frecuencia en JTU\_AWS

Con el propósito de establecer un patrón más detallado de la lluvia, durante los casi cuatro años de registros, se construyeron las curvas de intensidad duración y frecuencia (IDFs), utilizando la metodología planteada en el capítulo anterior, para los eventos de lluvia de 5 a 1800 minutos. Es claro que al contar con pocos años de registros de lluvia, la incertidumbre aumenta, además considerando Klemeš (1986), se tiene que la utilización de estas extrapolaciones son indefendibles desde el punto de vista hidrológico, sin embargo, se presentan los resultados con estas consideraciones.

La ecuación (4.1) representa el comportamiento de eventos de lluvia de la estación JTU\_AWS, la cual está dentro de la zona de estudio y se describe en la Tabla 4.6, donde se muestra la cantidad de precipitación máxima, para un evento con una duración y periodo de retorno dados.

$$I(t, T) = \frac{0.6124 \ln T + 2.7390}{(28.0095 + t)^{0.6722}} \quad (4.1)$$

Se puede apreciar en la Tabla 4.6 y la Figura 4.8, que la precipitación de eventos con duraciones menores a dos horas no son muy intensas, para tener una idea, la precipitación acumulada máxima esperada en 5 años, para una duración de 2 horas, es de 15.5 mm; por otro lado, se puede apreciar que las precipitaciones más intensas esperadas, considerando duraciones entre 12 a 30 horas, fluctúan entre 31.4 y 43.0 mm, para un período de retorno de 5 años, lo cual sugiere que, las precipitaciones extremas más usuales se dan en eventos con duraciones largas, lo cual sugiere un comportamiento usual de páramo donde los eventos de lluvia son generalmente de larga duración y baja intensidad, como se indica en Sklenar et al. (2015), lo cual contrasta con la precipitación en la ciudad de Quito, donde se han registrado eventos de lluvia que superan los 43 mm en una hora, como se indica en Escobar-González et al. (2022).

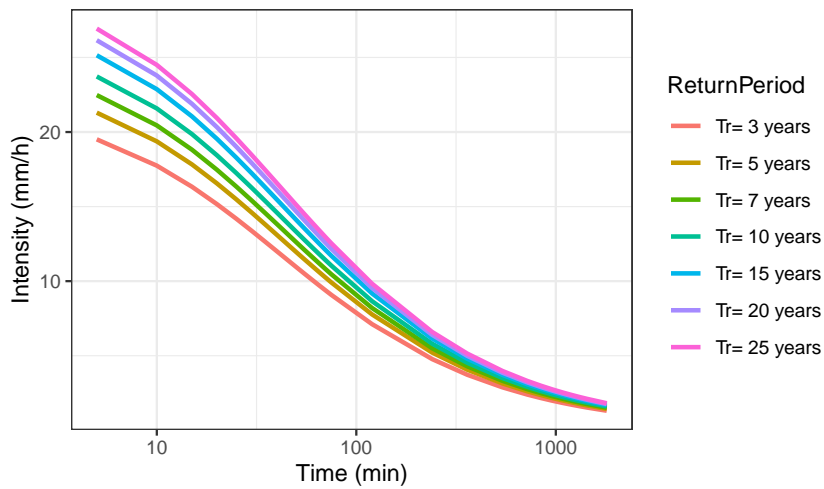


Figura N° 4.8: Intensidad máxima de precipitación (mm/min), escala logarítmica.

Tabla N° 4.6:

Cantidad de precipitación máxima estimada (mm) por duración y período de retorno.

t(min)	Período de retorno $T$ (años)						
	1	3	5	10	15	20	25
5	1.3	1.6	1.8	2.0	2.1	2.2	2.2
10	2.4	3.0	3.2	3.6	3.8	4.0	4.1
15	3.3	4.1	4.5	5.0	5.3	5.5	5.6
20	4.1	5.1	5.5	6.1	6.5	6.8	7.0
25	4.7	5.9	6.5	7.2	7.6	7.9	8.2
30	5.4	6.7	7.3	8.1	8.6	9.0	9.2
35	5.9	7.4	8.0	9.0	9.5	9.9	10.2
40	6.4	8.0	8.7	9.7	10.3	10.7	11.0
45	6.9	8.6	9.4	10.4	11.1	11.5	11.8
50	7.3	9.1	10.0	11.1	11.8	12.2	12.6
55	7.7	9.6	10.5	11.7	12.4	12.9	13.3
60	8.1	10.1	11.0	12.3	13.0	13.5	13.9
120	11.4	14.2	15.5	17.3	18.3	19.1	19.6
240	15.3	19.1	20.8	23.2	24.6	25.6	26.4
360	17.9	22.3	24.4	27.2	28.8	29.9	30.8
540	20.8	25.9	28.3	31.5	33.4	34.8	35.8
720	23.1	28.7	31.4	34.9	37.0	38.5	39.7
960	25.5	31.8	34.7	38.6	41.0	42.6	43.9
1200	27.6	34.3	37.5	41.7	44.2	46.0	47.4
1440	29.3	36.5	39.9	44.4	47.1	49.0	50.4
1800	31.6	39.4	43.0	47.9	50.8	52.8	54.4

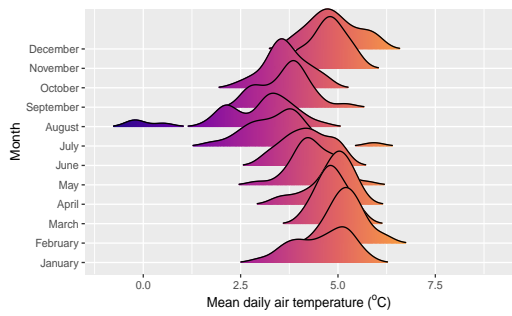
#### **4.1.6 Distribución temporal de la información climatológica**

##### 4.1.6.1 Distribución multianual mensual de la temperatura

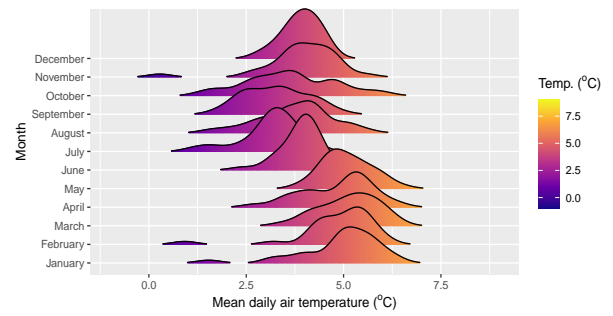
Para describir la distribución temporal de la temperatura, se calculó la temperatura media diaria y se graficó la distribución de frecuencias correspondientes de cada año en las Figuras 4.9(a) a 4.9(d), mientras que en la Figura 4.9(e), se consolidó la información registrada en todo el período de estudio, graficando la distribución de frecuencias con respecto a los 12 meses.

Evidentemente los picos de cada una de las distribuciones de frecuencias en la Figura 4.9, indican donde se concentra la mayor cantidad de información similar, sin embargo, existen meses cuya distribución tiene dos picos, esto puede indicar un cambio del régimen estacional, es decir, un cambio de época fría a más cálida o viceversa, además, existen distribuciones con más de dos picos, como por ejemplo: agosto 2019 y enero-febrero 2021; esto tiene sentido si consideramos la Figura 4.10, donde el número de días secos más altos podrían influir directamente en la temperatura ambiente, además de coincidir con el cambio de estacionalidad con respecto a la lluvia.

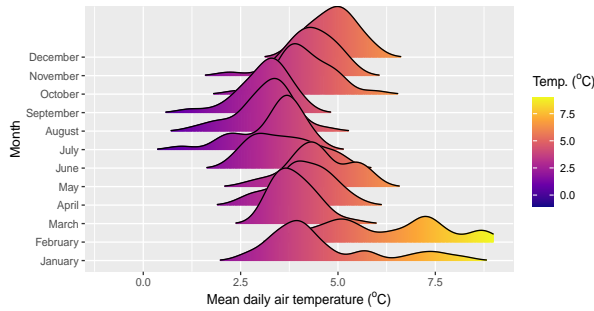
Ahora, si se consideran los picos de la distribución de frecuencias en cada mes, de toda la serie temporal de temperatura media de la Figura 4.9(e), se puede tener una idea del comportamiento interanual de la temperatura, es así que se presenta un período más frío entre los meses de junio a septiembre, mientras que en los meses de octubre a mayo considerando el año hidrológico, el comportamiento es más cálido.



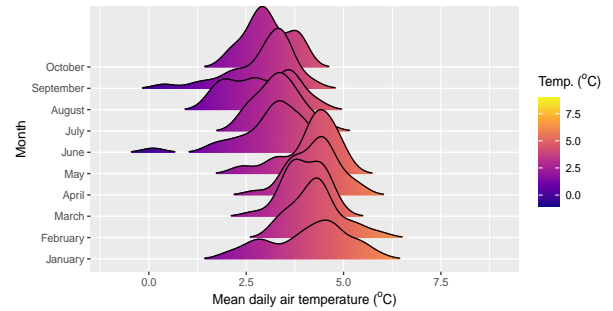
(a) Distribución de temperatura del aire ( $^{\circ}C$ ) 2019.



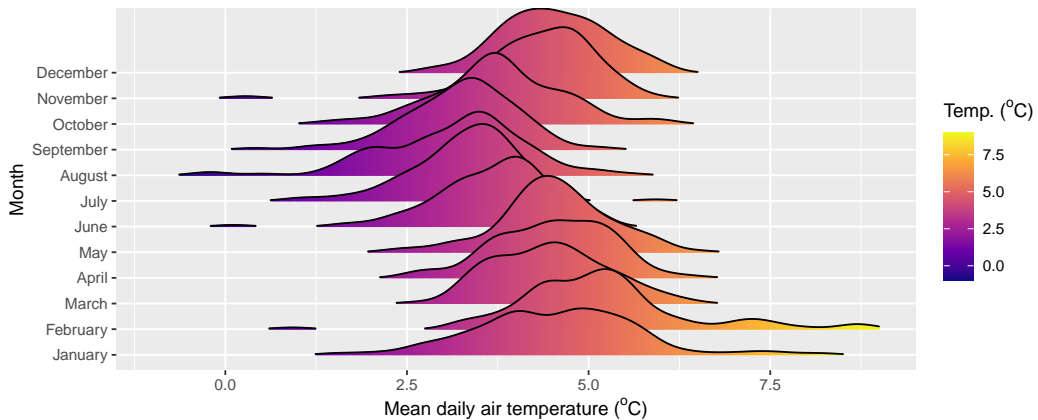
(b) Distribución de temperatura del aire ( $^{\circ}C$ ) 2020.



(c) Distribución de temperatura del aire ( $^{\circ}C$ ) 2021.



(d) Distribución de temperatura del aire ( $^{\circ}C$ ) 2022.



(e) Distribución de temperatura del aire ( $^{\circ}C$ ) enero 2019-octubre 2022.

**Figura N° 4.9:** Distribución de la temperatura del aire media diaria ( $^{\circ}C$ ) en Jatunhuayco de enero 2019 a octubre 2022.

#### 4.1.6.2 Distribución multianual mensual de la humedad relativa

Para describir la distribución temporal de la humedad relativa, se calculó la humedad relativa media diaria y se graficó la distribución de frecuencias correspondientes a cada año de las Figuras 4.11(a) a 4.11(c) y 4.11(d), mientras que en la Figura 4.11(e), se consolidó la información registrada en todo el período de estudio enero 2019 - octubre 2022, graficando

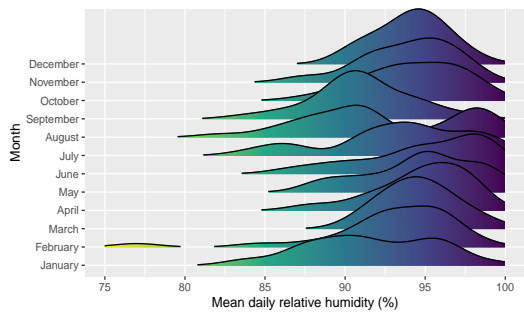


**Figura N° 4.10:** Media mensual del número de días secos en Jatunhuayco.

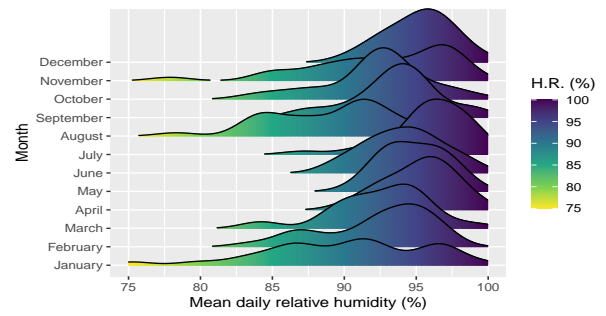
la distribución de frecuencias con respecto a los 12 meses.

Es claro que los picos de cada una de las distribuciones de frecuencias en la Figura 4.11, indican donde se distribuye la mayor cantidad de información similar, sin embargo, existen meses cuya distribución tiene dos picos, esto puede indicar un cambio del régimen estacional, de la misma forma que ocurre con la temperatura del aire descrita en la subsección anterior, es decir, un cambio de época húmeda a más seca o viceversa, además, se puede observar la relación directa que existe con los meses correspondientes al número de días secos, mostrados en la Figura 4.10, donde el número de días secos más altos implica un ambiente más seco y viceversa.

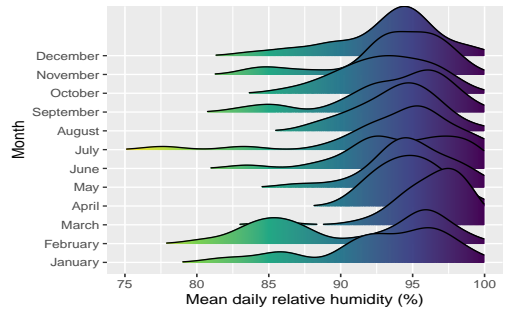
Ahora, si se consideran los picos de la distribución de frecuencias en cada mes, de toda la serie temporal en la Figura 4.11(e), se puede tener una idea del comportamiento interanual de la humedad relativa, es así, que se presenta un período más seco entre los meses de enero a febrero y de julio a septiembre, mientras que en los meses de marzo a junio y de octubre a diciembre se tiene mayor humedad, considerando el año hidrológico, lo cual coincide con el número de días secos Figura 4.10 y la precipitación multianual mensual de JTU\_AWS Jatunhuayco descrita en la Figura 4.6.



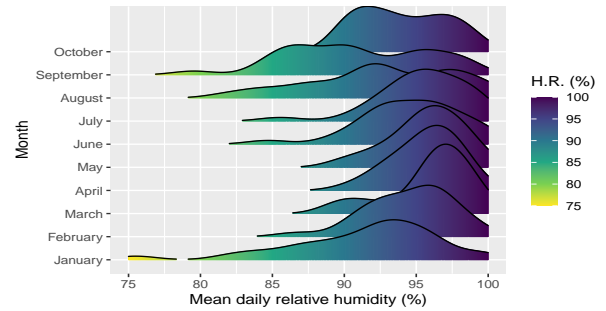
(a) Distribución de humedad relativa (%) 2019.



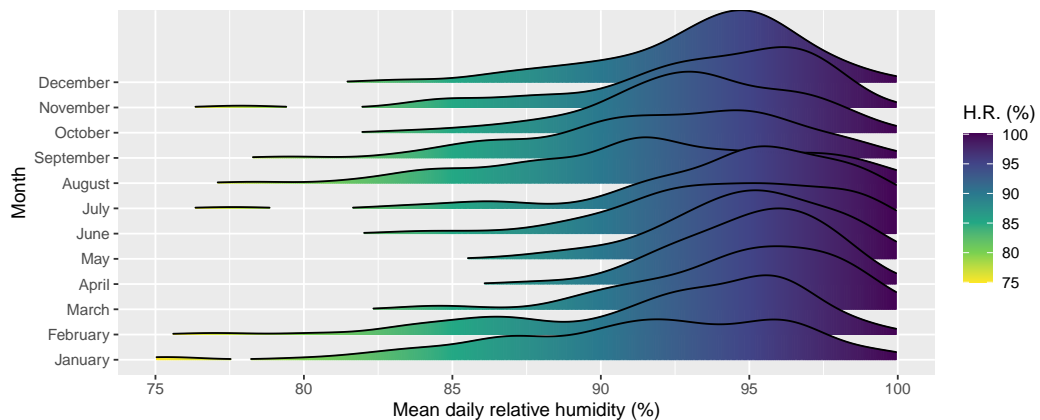
(b) Distribución de humedad relativa (%) 2020.



(c) Distribución de humedad relativa (%) 2021.



(d) Distribución de humedad relativa (%) 2022.



(e) Distribución de humedad relativa (%) enero 2019-octubre 2022.

**Figura N° 4.11:** Distribución de la humedad relativa media diaria (%) en Jatunhuayco 2019 - octubre 2022.

## 4.2 Descripción de la relación entre precipitación y humedad del suelo

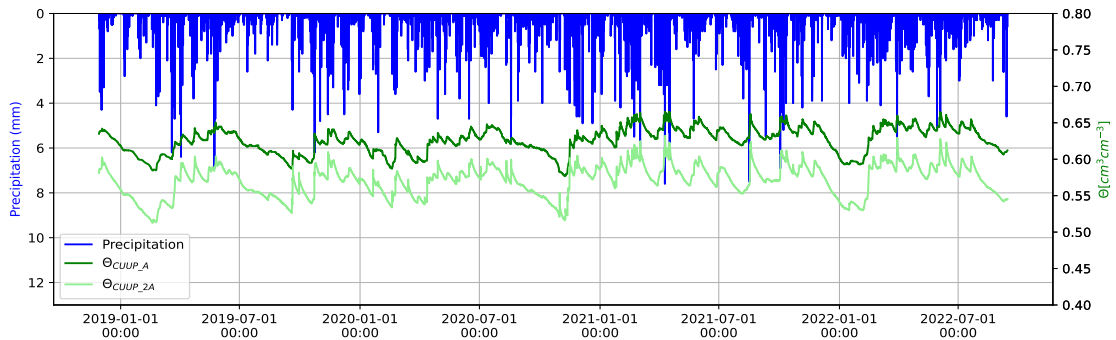
En la presente sección se busca relacionar la precipitación y humedad del suelo para almohadilla en los horizontes *A* y *2A*, para ello, se presenta en las Figuras 4.12(a), 4.12(b), 4.12(c) y 4.12(d), la precipitación acumulada y humedad del suelo media: horaria, diaria, quincenal y mensual respectivamente.

Es claro que, la respuesta de la señal de la humedad del suelo, depende de la lluvia y su relación es directa, es decir que a mayor cantidad de lluvia acumulada, la humedad del

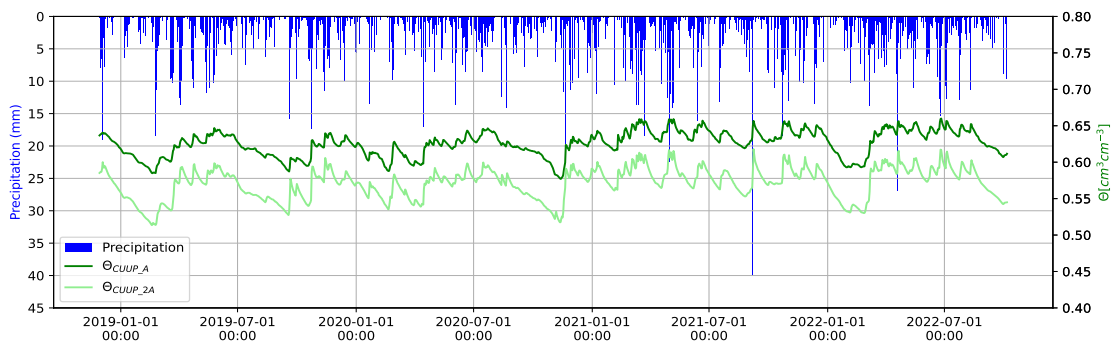


suelo aumentará y si disminuye la lluvia también lo hará la humedad, esto se puede apreciar en todas las subfiguras de la Figura 4.12. Una observación importante se puede notar en la Figuras 4.12(a), donde las señales de la humedad del suelo horaria en almohadilla, de los perfiles *A* y *2A*, son prácticamente las mismas, que en la humedad del suelo diaria de e los perfiles *A* y *2A* de la Figura 4.12(b), esto indica que los registros de humedad del suelo horarios y diarios permanecen casi invariantes.

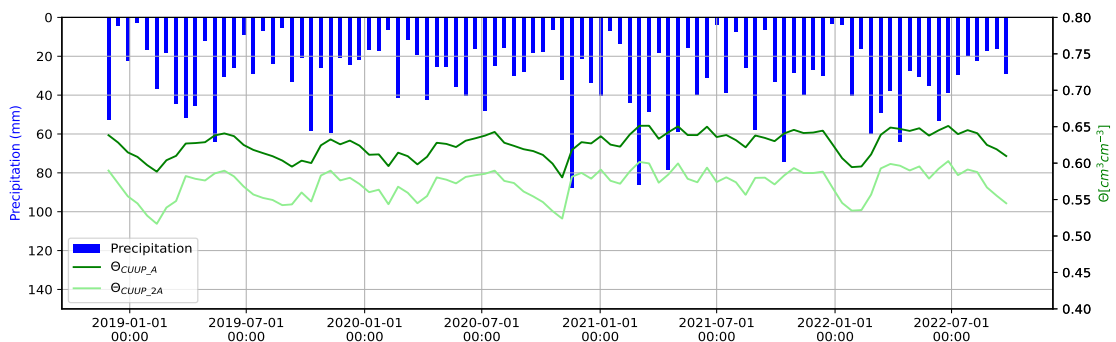
Otra observación importante se presenta en la frecuencia quincenal observada en la Figura 4.12(c), donde las señales de la precipitación y humedad del suelo comienzan a parecerse, mientras que al aumentar el período a mensual en la Figura 4.12(d), el parecido es mayor todavía, pero considerando un desfase temporal de la humedad del suelo de aproximadamente un mes.



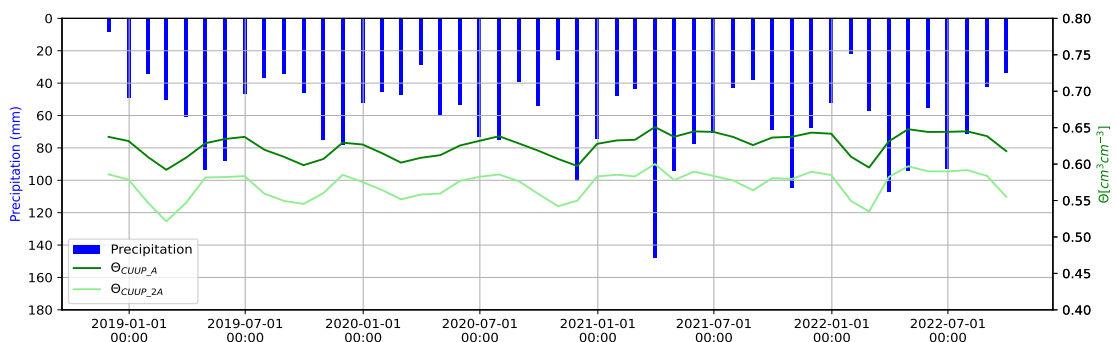
(a) Precipitación y humedad del suelo horaria, horizontes *A* y *2A* en almohadilla 2019 - 2022.



(b) Precipitación y humedad del suelo diaria, horizontes *A* y *2A* en almohadilla 2019 - 2022.

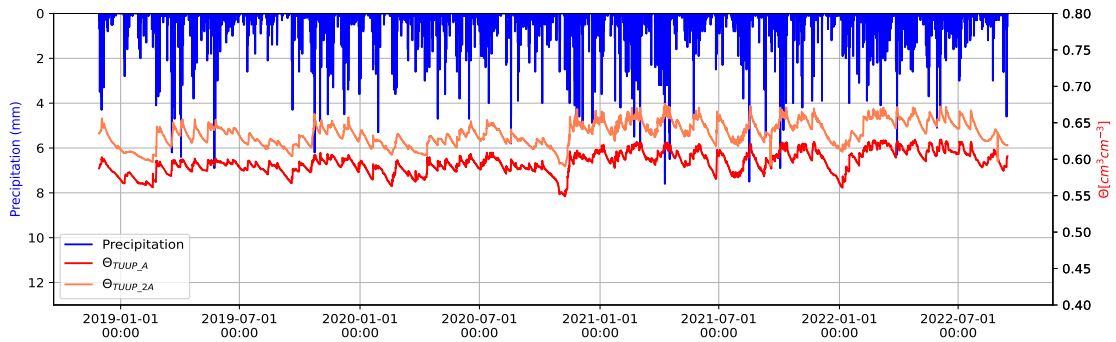


(c) Precipitación y humedad del suelo quincenal, horizontes *A* y *2A* en almohadilla 2019 - 2022.

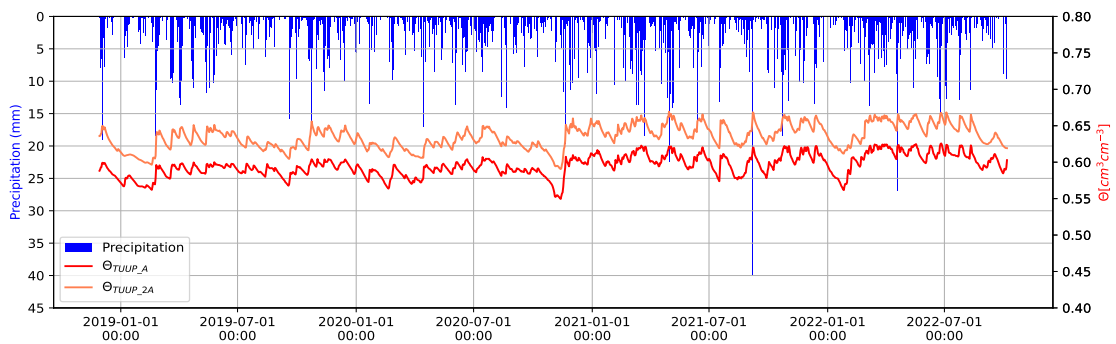


(d) Precipitación y humedad del suelo mensual, horizontes *A* y *2A* en almohadilla 2019 - 2022.

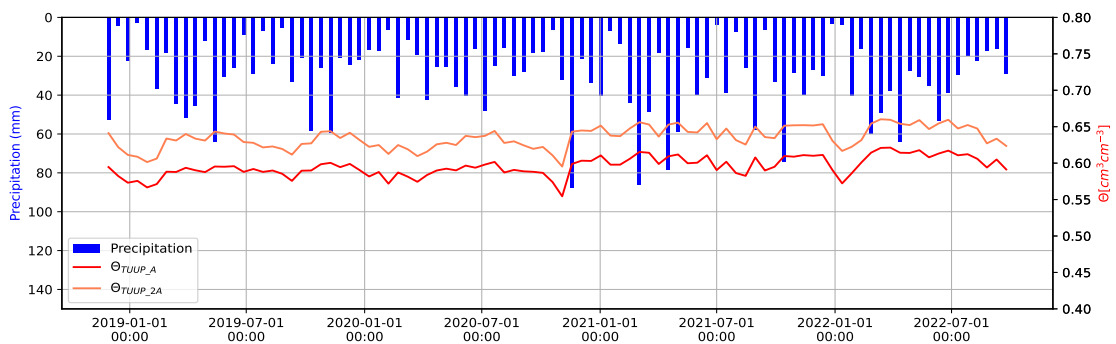
**Figura N° 4.12:** Precipitación y humedad del suelo en almohadilla en los horizontes *A* y *2A* a diferentes escalas temporales,  $\theta_{CU\_UPA}$  y  $\theta_{CU\_UP2A}$ , Jatunhuayco enero 2019 - septiembre 2022.



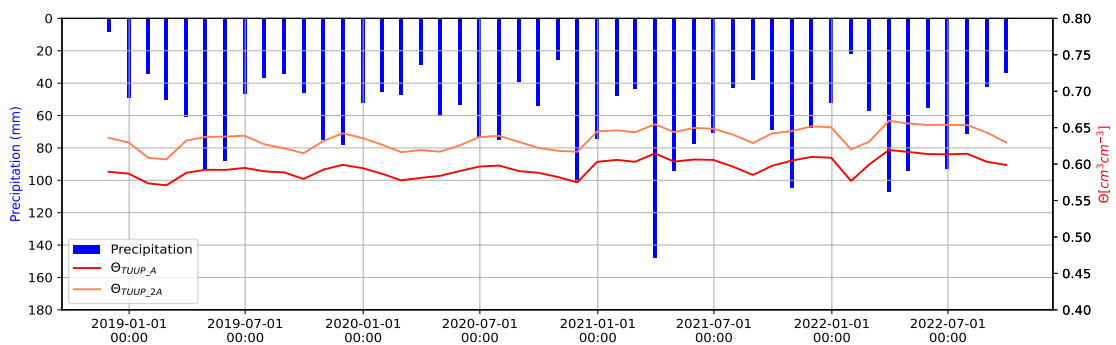
(a) Precipitación y humedad del suelo horaria, horizontes A y 2A en pajonal 2019 - 2022.



(b) Precipitación y humedad del suelo diaria, horizontes A y 2A en pajonal 2019 - 2022.



(c) Precipitación y humedad del suelo quincenal, horizontes A y 2A en pajonal 2019 - 2022.



(d) Precipitación y humedad del suelo mensual, horizontes A y 2A en pajonal 2019 - 2022.

**Figura N° 4.13:** Precipitación y humedad del suelo en pajonal en los horizontes A y 2A a diferentes escalas temporales,  $\theta_{TU\_UPA}$  y  $\theta_{TU\_UP2A}$ , Jatunhuayco enero 2019 - septiembre 2022.

### 4.2.1 Ciclo diario de la humedad del suelo

Para estudiar el comportamiento de la humedad del suelo durante el día, se procedió a calcular las diferencias consecutivas de la información horaria y graficar en un heatmap por año, con sus respectivos meses, días y horas, además, se graficó un último heatmap de la media de las diferencias consecutivas horarias de todo el período de estudio (Hourly mean). Se debe notar que la paleta de colores no es uniforme y va de tonalidades azules para valores negativos, pasando por blanco y a tonalidades rojas en los valores positivos, nótese que el color blanco no corresponde a los valores cero ya que la visualización se dificulta.

Se graficó el mismo tipo de heatmap en el horizonte *A* para los dos tipos de suelos: almohadilla y pajonal en las Figuras 4.14 y 4.15 respectivamente, y se consideraron los extremos en las paletas de color, tomando en cuenta el mínimo y máximo de las dos series temporales de las variaciones consecutivas de humedad del suelo como una sola serie, para de esta manera identificar en qué tipo de cobertura vegetal son más favorables este tipo de fluctuaciones.

#### 4.2.1.1 Ciclo diario de la humedad del suelo en almohadilla

En la Figura 4.14, se presentan las diferencias consecutivas horarias, de los registros de humedad del suelo en la cumbre de almohadilla en el horizonte *A* ( $\Delta\theta_{CU\_UPA}$ ), se puede notar según la paleta de colores que los valores de las diferencias consecutivas son muy bajos, además oscilan entre -0.005801742 y 0.014 ( $cm^3cm^{-3}$ ), por lo tanto, son imperceptibles y en escalas temporales más grandes, como en las Figuras 4.12(a) y 4.12(b) no se pueden visualizar. Sin embargo, con este nuevo enfoque se puede describir con más detalle, la respuesta de la humedad del suelo, después de un evento intenso de lluvia.

Es claro que la presencia de lluvia en un momento dado, implica aumento en la humedad en el suelo, sin embargo, este comportamiento depende de las condiciones antecedentes

de lluvia y humedad en el suelo, un primer vistazo de la Figura 4.14, indica que existen señales de similitud con el heatmap de precipitación acumulada horaria de la Figura 4.7, es decir, tras la identificación de un evento de lluvia, existe un tiempo de respuesta para que la humedad del suelo comience a aumentar, hasta llegar a un valor máximo y luego disminuye hasta estabilizarse, hasta la presencia de un nuevo evento de precipitación. Se puede apreciar en la Figura 4.14, que las variaciones de las señales que indican el inicio de un aumento en la humedad del suelo, se dan principalmente en horas de la tarde y considerando un tiempo de desfase, coinciden con las señales de precipitación de la Figura 4.7.

#### 4.2.1.2 Ciclo diurno de la humedad del suelo en pajonal

En la Figura 4.15, se presentan las diferencias consecutivas horarias, de los registros de humedad del suelo de pajonal en el horizonte A ( $\Delta\theta_{TU\_UPA}$ ), se puede notar según la paleta de colores que los valores de las diferencias consecutivas son muy bajos, además oscilan entre  $-0.005801742$  y  $0.014$  ( $cm^3 cm^{-3}$ ).

En principio la Figura 4.15, indica que existen señales de similitud con el heatmap de precipitación acumulada horaria de la Figura 4.7 al igual que en el caso de almohadilla, es decir, tras la identificación de un evento de lluvia, existe un tiempo de respuesta para que la humedad del suelo comience a aumentar, hasta llegar a un valor máximo y luego disminuye hasta estabilizarse, hasta la presencia de un nuevo evento de precipitación y al igual que en caso anterior, las variaciones de las señales que indican el inicio de un aumento en la humedad del suelo, se dan principalmente en horas de la tarde y considerando un tiempo de desfase coinciden con las señales de precipitación de la Figura 4.7.

Una observación muy importante es que las señales de humedad de suelo en el horizonte A para pajonal, tienen una reacción más intensa que en almohadilla y esto se puede visualizar de mejor manera por las tonalidades más intensas, correspondientes a la paleta de color.

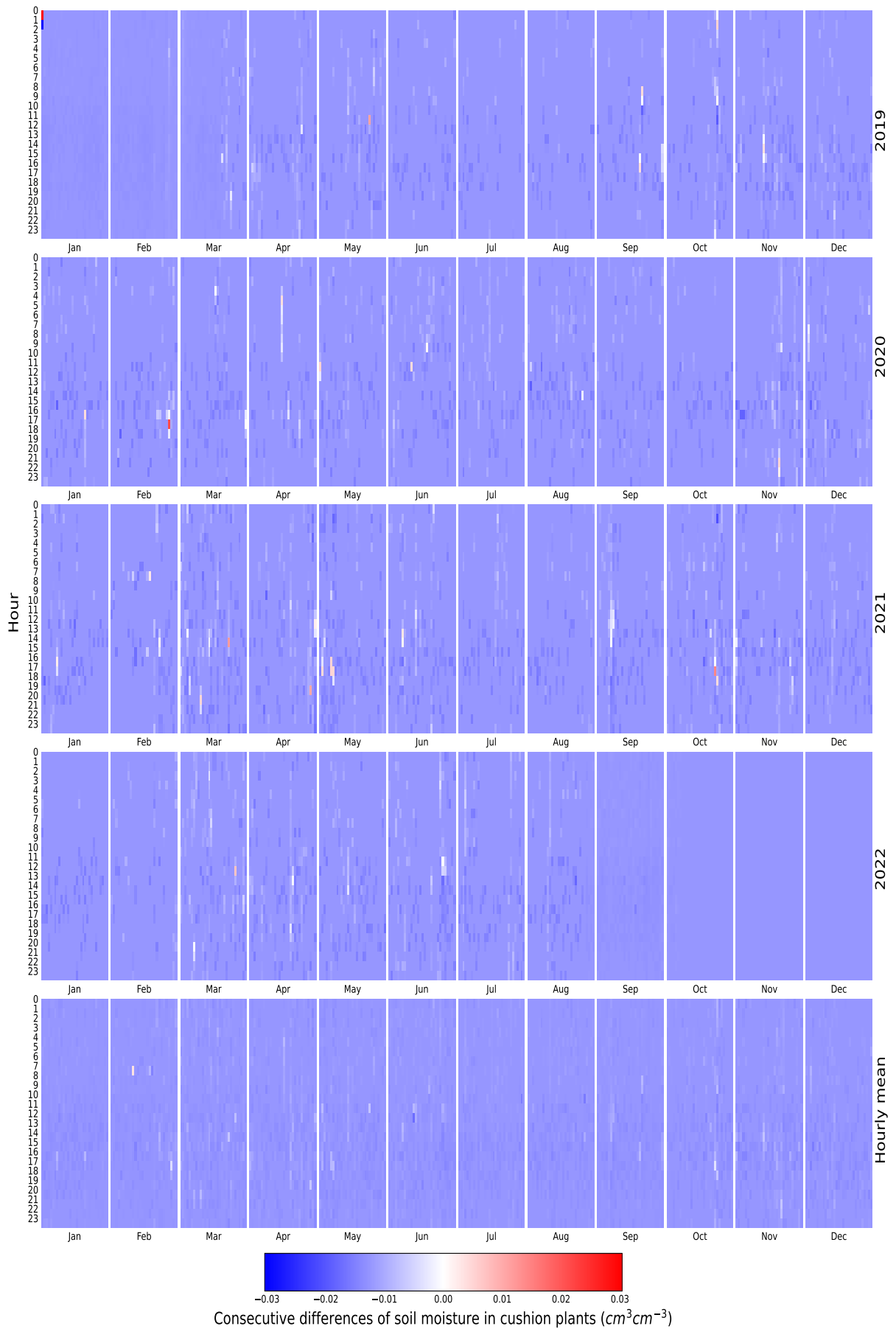


Figura N° 4.14: Serie de  $\Delta\theta_{CU\_UPA}$  en Jatunhuayco, enero 2019 - septiembre 2022

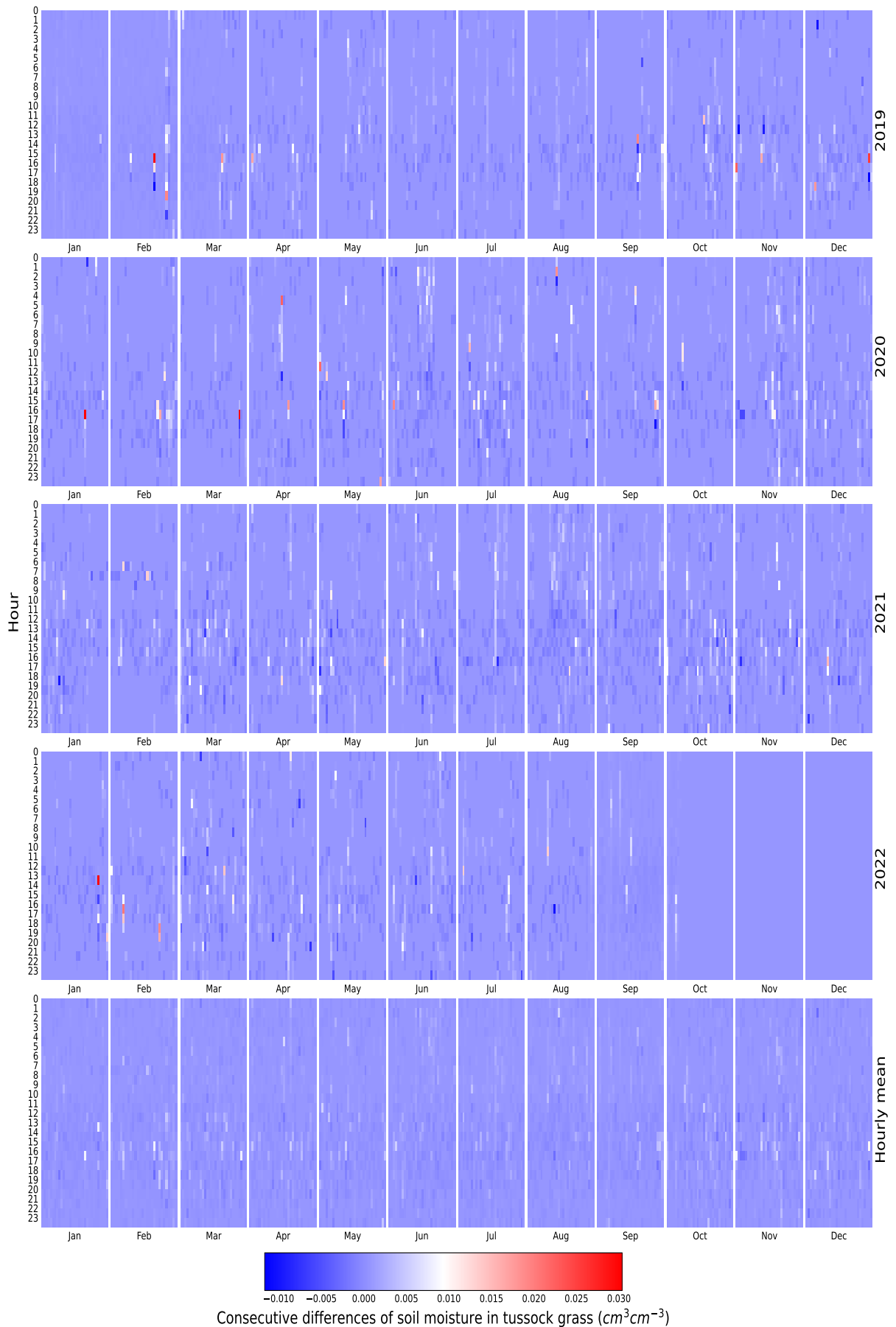


Figura N° 4.15: Serie de  $\Delta\theta_{TU\_UPA}$  en Jatunhuayco, enero 2019 - septiembre 2022

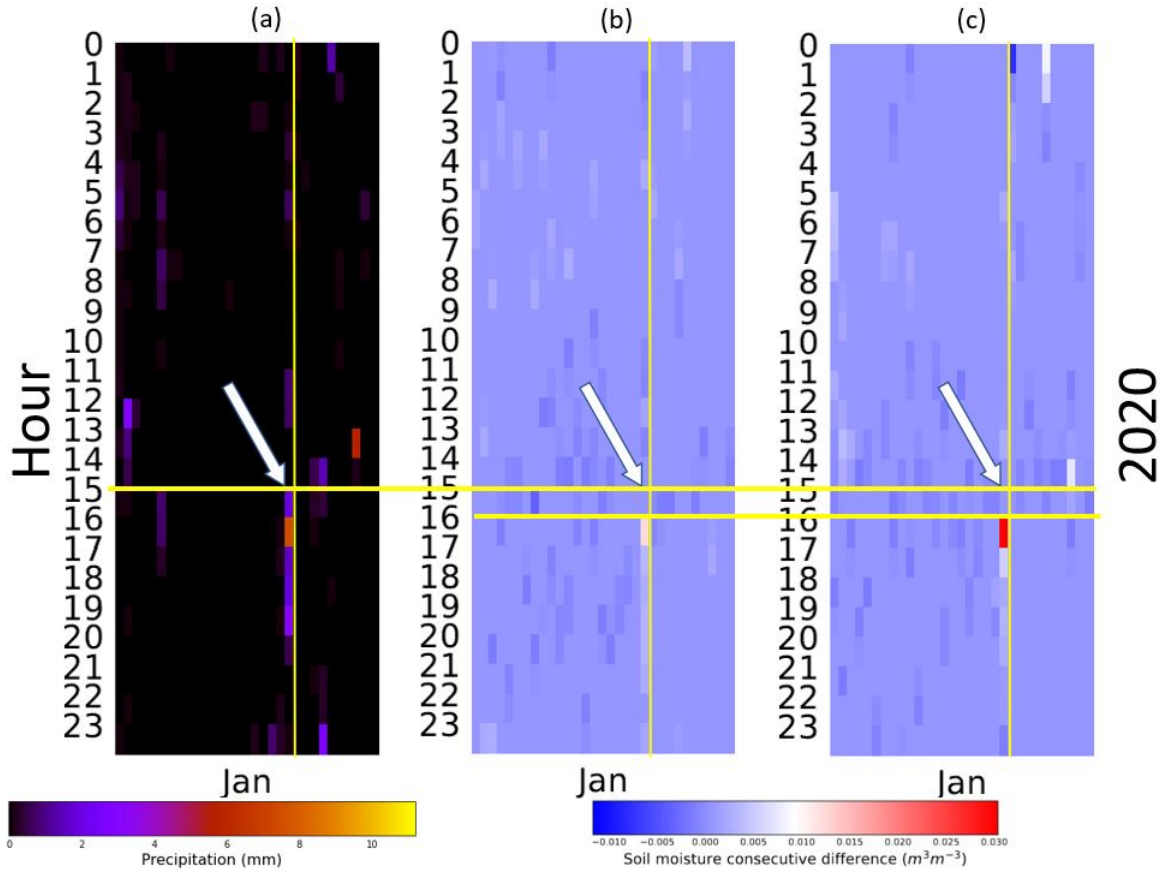
#### ***4.2.2 Tiempo de respuesta de la humedad del suelo frente a un evento de lluvia intensa***

En esta sección se presenta, una breve descripción del comportamiento de la humedad del suelo y su respuesta frente a un evento de precipitación intensa, para esto se presenta un zoom de las Figuras 4.16(a), 4.16(b) y 4.16(c), correspondientes al heatmap de precipitación horaria acumulada de las diferencias consecutivas del humedad del suelo en almohadilla y pajonal respectivamente, para el horizonte *A* del mes de enero de 2020.

En la Figura 4.16(a), se puede apreciar el evento de precipitación del 21 de enero de 2020 en JTU\_AWS, mismo que inicio a las 15h00 hasta las 20h00 (ubicado en la intersección de las líneas amarillas horizontal y vertical, en la primera flecha blanca), con una lluvia acumulada de 16.7 mm en ese período, mientras que en la Figura 4.16(b), se muestra que el inicio de la respuesta de la humedad del suelo en almohadilla, se da una hora después del inicio de la lluvia y coincide con el máximo valor de humedad del suelo de almohadilla en ese período, luego se aprecia claramente como se va atenuando la humedad del suelo hasta que se estabiliza; por otro lado, en la Figura 4.16(c), se muestra que el inicio de la respuesta de la humedad del suelo en pajonal, se da también una hora después del inicio del lluvia y coincide con el máximo valor de humedad del suelo durante ese período, nótese que este valor es más grande que en la almohadilla, después se aprecia claramente como va disminuyendo la humedad del suelo, hasta que se estabiliza durante las primeras horas del 22 de enero de 2020.

Es claro que, las señales de precipitación intensa coinciden con los registros de mayor variación de las diferencias consecutivas horarias de humedad del suelo de almohadilla y pajonal, por otro lado, el ejemplo descrito en esta sección, podría generalizarse haciendo una captura automática de eventos de precipitación, considerando la duración e intensidad, como se describe en Escobar-González et al. (2022), de esta manera se podrían encontrar patrones de variación temporal de humedad del suelo, para diferentes tipos de eventos intensos de lluvia.





**Figura N° 4.16:** (a) Heatmap precipitación horaria acumulada Jatunhuayco, enero-2020. (b) Heatmap de la serie de  $\Delta\theta_{CU\_UPA}$  en Jatunhuayco enero-2020, (c) Heatmap de la serie de  $\Delta\theta_{TU\_UPA}$  en Jatunhuayco enero-2020.

#### 4.2.3 Efecto de la precisión del reflectómetro CS650 en las series de humedad del suelo

En este apartado se verificó que los valores de las diferencias consecutivas sean coherentes con la precisión de los reflectómetros correspondiente a equipos calibrados *in-situ*.

Sea la serie  $\{\theta\}_{k=\{1,\dots,n\}} = \{\theta_1, \theta_2, \dots, \theta_n\}$  correspondiente a la serie temporal de humedad del suelo, en almohadilla (Cu) de la zona más alta (Up) en el horizonte A, definimos la serie de diferencias consecutivas

$$\{\Delta\theta\}_{k=\{1,\dots,n\}} = \{\Delta\theta_1, \Delta\theta_2, \dots, \Delta\theta_n\},$$

tal que:  $\Delta\theta_1 = 0$  y  $\Delta\theta_k = \theta_k - \theta_{k-1}$ .

Sean ahora  $\theta_M = \max\{\theta\}_{k=\{1,\dots,n\}}$  y  $\theta_m = \min\{\theta\}_{k=\{1,\dots,n\}}$ , es claro que, para todo

$k$ :

$$\theta_m - \theta_M \leq \Delta\theta_k \leq \theta_M - \theta_m,$$

luego, si consideramos que la precisión del reflectómetro *CS650* es  $\pm 1\%$ , dado que sí se realizó una calibración del equipo, se tiene:

$$(\theta_m \pm 1\%) - (\theta_M \pm 1\%) \leq \Delta\theta_k \leq (\theta_M \pm 1\%) - (\theta_m \pm 1\%),$$

se puede afirmar que:

$$\theta_m - \theta_M - 0.02 \leq \Delta\theta_k \leq \theta_M - \theta_m + 0.02,$$

de las series temporales analizadas, se tiene que en almohadilla para el horizonte y perfil dados:  $\theta_M = 0.665$  y  $\theta_m = 0.577$ , de donde:

$$[-0.108 \leq \Delta\theta_k \leq 0.108] \text{cm}^3 \text{cm}^{-3}.$$

Finalmente, si consideramos la Figura 4.17, es claro que el intervalo resultante contiene a todos los valores de las diferencias consecutivas de humedad en almohadilla. Se puede realizar el mismo procedimiento para otros perfiles y horizontes e incluir al pajonal, encontrándose similares resultados.

Otro resultado que se puede apreciar de la Figura 4.17, es la distribución de las diferencias consecutivas para almohadilla, donde se puede notar que la mayoría de los datos, se encuentran más concentrados en un intervalo más pequeño, lo cual contrasta con el box-plot del pajonal, cuya distribución es más amplia, esto sugiere que, en el perfil más alto (Up), para el horizonte A, la almohadilla tenga mayores fluctuaciones de humedad que en pajonal.

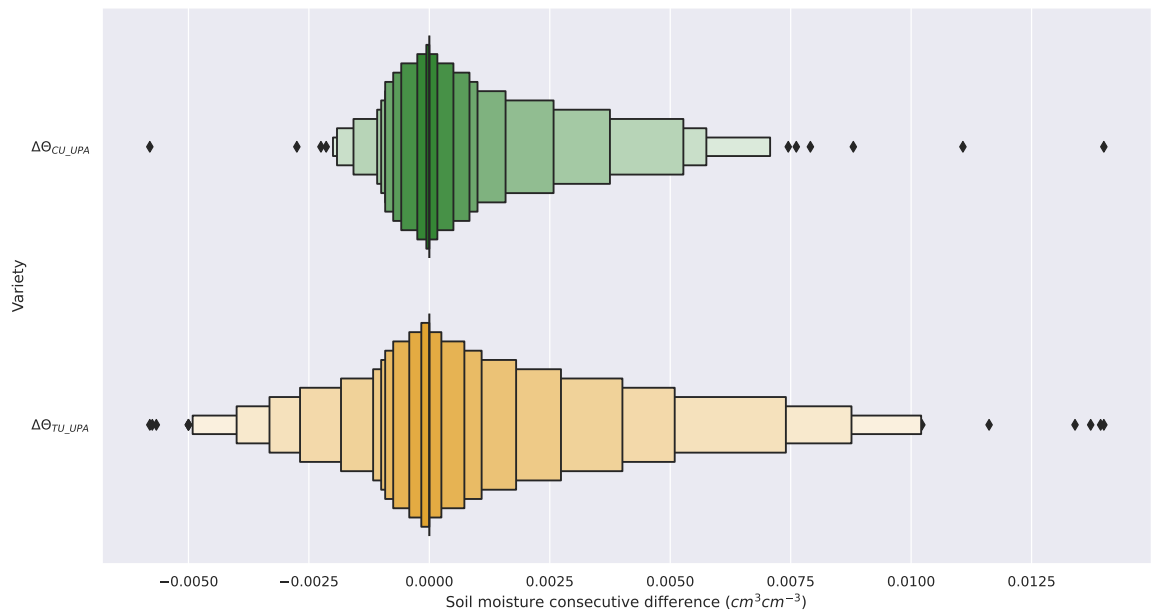


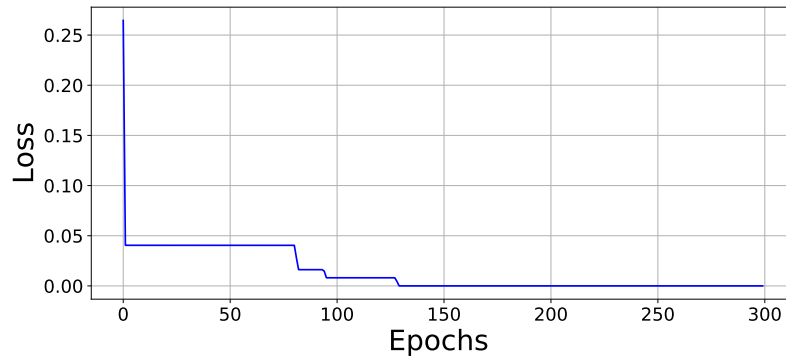
Figura N° 4.17: Box-plot de las series de diferencias consecutivas:  $\Delta\theta_{CU\_UPA}$  y  $\Delta\theta_{TU\_UPA}$ .

### 4.3 Red Neuronal Profunda

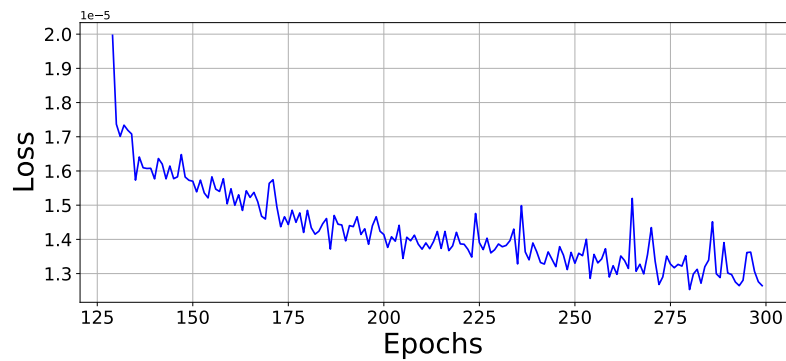
La técnica de transfer-learning utilizada para que el conocimiento de un modelo ya entrenado sea útil en un modelo con dominio similar, brinda resultados con errores de orden entre  $1 \times 10^{-6} < \epsilon < 1 \times 10^{-3}$ . Para datos de entrenamiento, el pronóstico de la red base presentó buenos resultados, siendo la métrica de error con menor magnitud RMSE igual a  $4.77 \times 10^{-6}$ , NSE y KGE igual a 0.97.

#### 4.3.1 Optimización de la función de pérdida

Durante el entrenamiento de la red neuronal base, usada para pronosticar los valores de las siguientes 48 horas, dado siete días de información, la función de pérdida encontró un mínimo local a partir de la época 100 y 150, Figura 4.18(a). Sin embargo, al ser una red profunda, se sigue optimizando lentamente hasta la época 300, Figura 4.18(b).



(a) Valor de la función de pérdida durante el entrenamiento.



(b) Zoom del valor de la función de pérdida después de 127 épocas.

**Figura N° 4.18:** Función de pérdida de la red que pronostica los valores horarios de  $\theta_{CU\_LOA}$ , para las siguientes 48 horas, dada información horaria de:  $P$ ,  $T$ ,  $HR$ ,  $\theta_{CU\_LOA}$  y  $T_{CU\_LOA}$ , de siete días.

#### 4.3.2 Red Base para pronóstico de humedad del suelo de almohadilla en la zona baja (LO) del horizonte A

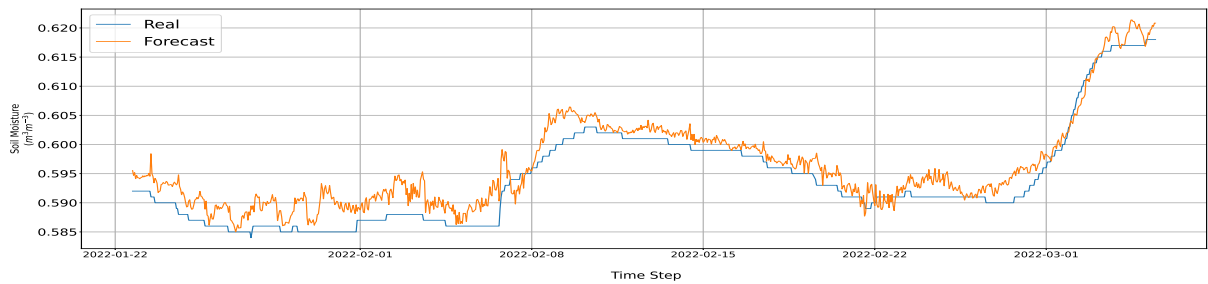
En esta sección se presentan los resultados de la red base que pronostica valores horarios de 2 días, humedad del suelo de almohadilla en la zona baja (LO) del horizonte A ( $\theta_{CU\_LOA}$ ), utilizando como variables de entrada: precipitación, temperatura ambiente, humedad relativa, radiación solar, con humedad y temperatura del suelo de la almohadilla en la zona baja, para el horizonte A. Además, se presenta la evaluación de la red base, considerando: entrenamiento, validación y evaluación.

#### 4.3.2.1 Pronóstico de humedad del suelo de almohadilla en la zona baja del horizonte A, para el conjunto de datos de evaluación

A continuación, se presenta en la Figura 4.19a, el pronóstico de 48 horas de  $\theta_{CU\_LOA}$ , para la red que toma como entrada información de siete días de información horaria de las variables ambientales para el conjunto de datos de evaluación:  $P$ ,  $T$ ,  $HR$ ,  $RS$ ,  $\theta_{CU\_LOA}$  y  $T_{CU\_LOA}$ , mientras que en la Figura 4.19(b), se puede ver un zoom de la Figura 4.19(a) desde el 2022-01-22 al 2022-03-04.



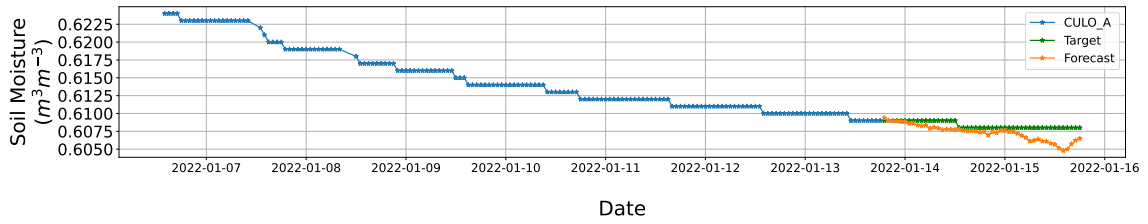
(a) Pronóstico para el conjunto de datos de evaluación



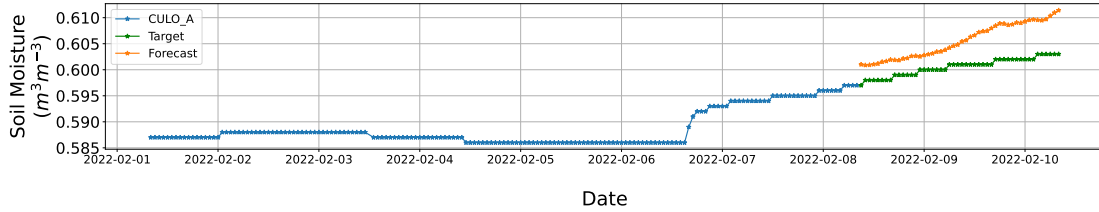
(b) Zoom de 4.19a del 2022-01-22 al 2022-03-04.

**Figura N° 4.19:** Pronóstico de 48 horas de  $\theta_{CU\_LOA}$ , para el conjunto de datos de evaluación horarios de 7 días de:  $P$ ,  $T$ ,  $HR$ ,  $RS$ ,  $\theta_{CU\_LOA}$  y  $T_{CU\_LOA}$ .

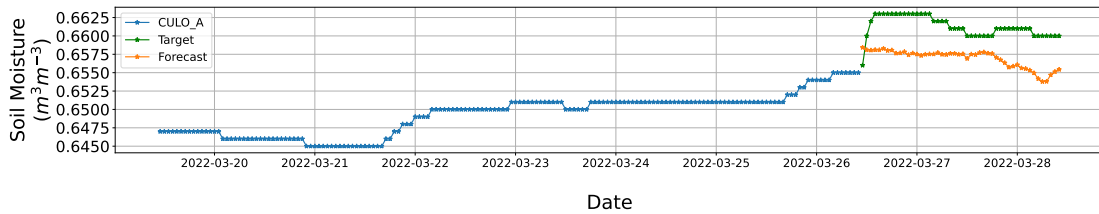
Adicionalmente, en las Figuras 4.20(a) a (d), se presentan los pronósticos de cuatro diferentes momentos, para las series de evaluación en  $\theta_{CU\_LOA}$ .



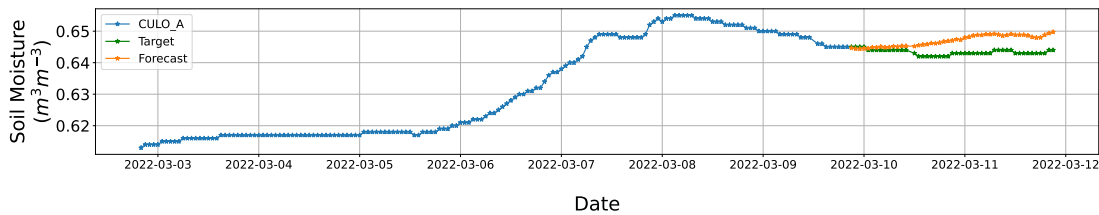
(a) Pronóstico de la serie de evaluación al 14 de enero de 2022.



(b) Pronóstico de la serie de evaluación al 8 de febrero de 2022.



(c) Pronóstico de la serie de evaluación al 26 de marzo de 2022.



(d) Pronóstico de la serie de evaluación al 9 de marzo de 2022.

**Figura N° 4.20:** Pronóstico para diferentes momentos de la serie de evaluación  $\theta_{CU\_LOA}$ .

#### 4.3.2.2 Evaluación de la red base que pronostica $\theta_{CU\_LOA}$

La Red Base que se evaluó fue capaz de predecir  $\theta_{CU\_LOA}$ , para las próximas 48 horas, utilizando las variables ambientales  $P$ ,  $T$ ,  $HR$ ,  $RS$ ,  $\theta_{CU\_LOA}$  y  $T_{CU\_LOA}$  de los siete días anteriores. En cuanto a las métricas de evaluación de desempeño y según la Tabla 4.7, se encontró que la función de pérdida utilizada durante el entrenamiento de la red tuvo un valor mínimo de  $1.26E-06$  en el conjunto de entrenamiento. En términos de precisión en la predicción de  $\theta_{CU\_LOA}$ , la red base tuvo un error absoluto medio (MAE) de 0.0025, un error cuadrático medio (RMSE) de 0.0036 y un error cuadrático medio en escala logarítmica (MSLE) de  $4.77E-06$  en el conjunto de entrenamiento. Estos valores aumentaron en los conjuntos

de validación y evaluación, lo que sugiere una cierta cantidad de sobreajuste de la red a los datos de entrenamiento, sin embargo los resultados son alentadores.

En cuanto a la capacidad de la red para ajustarse a los valores reales, se utilizaron los coeficientes de eficiencia de Nash-Sutcliffe (NSE) y el de Kling-Gupta (KGE). El valor más alto de NSE (0.97) se obtuvo en el conjunto de entrenamiento, lo que indica un buen ajuste entre las predicciones y los valores reales. Sin embargo, en el conjunto de evaluación, se encontró que la red tenía un valor negativo de NSE (-21.4), lo que indica que la red no se ajustó bien a los datos de prueba. El valor más alto de KGE (0.97) se obtuvo en el conjunto de entrenamiento, mientras que el valor más bajo (0.87) se obtuvo en el conjunto de evaluación.

En general, estos resultados indican que la red base evaluada tuvo un buen desempeño en el conjunto de entrenamiento, pero no tanto en los conjuntos de validación y evaluación, lo que sugiere un pequeño grado de sobreajuste a los datos de entrenamiento. Además, se encontró que la capacidad de la red para ajustarse a los valores reales varió según la métrica utilizada y el conjunto de datos evaluado.

**Tabla N° 4.7:**

**Métricas obtenidas para la red base que pronostica 48 horas de  $\theta_{CU\_LOA}$ , dada la información de las variables:  $P, T, HR, RS, \theta_{CU\_LOA}$  y  $T_{CU\_LOA}$  de siete días.**

	<b>Entrenamiento</b>	<b>Validación</b>	<b>Evaluación</b>
Loss	<b>1.26E-06</b>	2.03E-05	2.54E-05
MAE	<b>0.0025</b>	0.0033	0.0039
RMSE	<b>0.0036</b>	0.0045	0.005
MSLE	<b>4.77E-06</b>	7.47E-06	9.59E-06
NSE	<b>0.97</b>	0.89	-21.4
KGE	<b>0.97</b>	0.94	0.87

### ***4.3.3 Aplicación de la red base para pronóstico de humedad del suelo en pajonal usando transfer-learning***

#### **4.3.3.1 Pronóstico de humedad del suelo en pajonal ( $\theta_{TU}$ )**

En este apartado podemos ver el conjunto de Figuras 4.21, donde gracias al conocimiento adquirido desde la red base y con las técnicas de *transfer-learning*, se pudo pro-

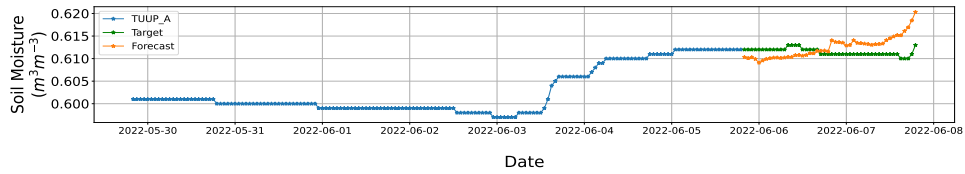


nosticar 48 horas de humedad del suelo en:  $\theta_{TU\_UPA}$ ,  $\theta_{TU\_UP2A}$ ,  $\theta_{TU\_URA}$ ,  $\theta_{TU\_UR2A}$ ,  $\theta_{TU\_MIA}$ ,  $\theta_{TU\_MI2A}$ ,  $\theta_{TU\_MIA}$ ,  $\theta_{TU\_LOA}$  y  $\theta_{TU\_LO2A}$ , donde se utilizó información horaria de 7 días como entrada, con el mismo dominio que la red base:  $P$ ,  $T$ ,  $HR$ ,  $RS$ ,  $\theta_{TU}$  y  $T_{TU}$ .

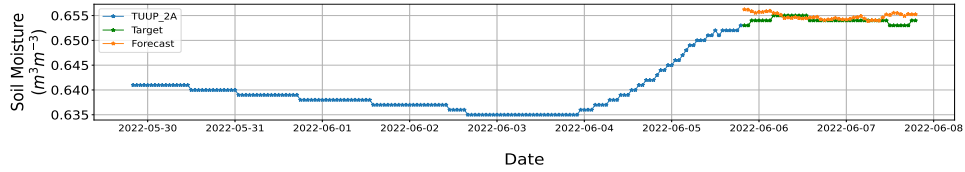
#### 4.3.3.2 Evaluación de los modelos de pronóstico de humedad del suelo en pajonal ( $\theta_{TU}$ )

La Tabla 4.8, presenta métricas de rendimiento para modelos de redes profundas que pronostican 48 horas de  $\theta_{TU}$  en función de:  $P$ ,  $T$ ,  $HR$ ,  $RS$ ,  $\theta_{TU}$  y  $T_{TU}$ . Las métricas incluyen los estadísticos: Loss, MAE, RMSE y MSLE; para los conjuntos de datos de: entrenamiento, validación y evaluación. Se observan resultados altamente prometedores en esta tabla. Las cifras resaltadas en negrita indican un rendimiento particularmente destacado. Estos valores resaltados, por ejemplo, Loss de 3.6e-6 para  $\theta_{TU\_UPA}$  en el conjunto de entrenamiento, señalan una excelente capacidad del modelo para ajustarse a los datos de entrenamiento.

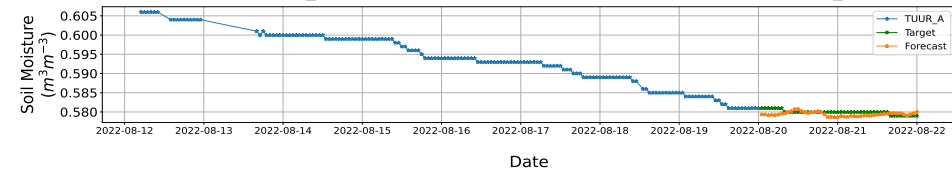
El MAE y RMSE también son notables, con valores tan bajos como 1.4e-3 y 1.9e-3 respectivamente. Estos números indican que los modelos tienen una precisión sobresaliente al predecir  $\theta_{TU}$  en condiciones diversas. El MSLE, con valores en la escala de 1e-6, muestra que los modelos son especialmente efectivos al predecir en un entorno logarítmico, lo que es crucial en muchas aplicaciones.



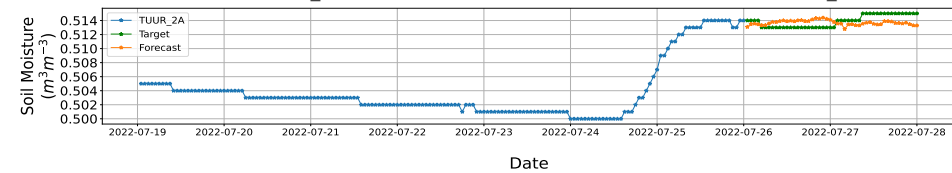
(a) Pronóstico de  $\theta_{TU\_UPA}$ , usando transfer-learnig de la red base  $\theta_{CU\_LOA}$ .



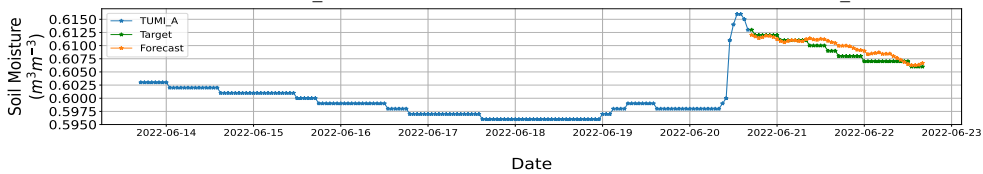
(b) Pronóstico de  $\theta_{TU\_UP2A}$ , usando transfer-learnig de la red base  $\theta_{CU\_LOA}$ .



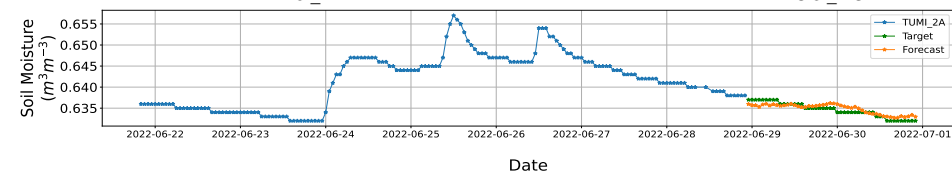
(c) Pronóstico de  $\theta_{TU\_URA}$ , usando transfer-learnig de la red base  $\theta_{CU\_LOA}$ .



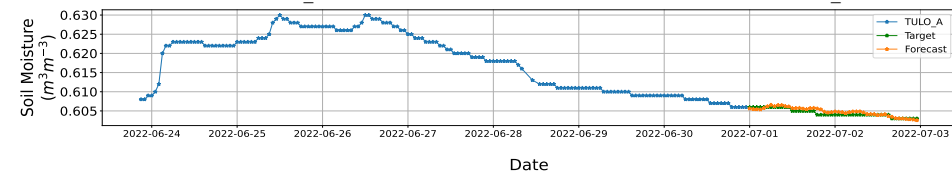
(d) Pronóstico de  $\theta_{TU\_UR2A}$ , usando transfer-learnig de la red base  $\theta_{CU\_LOA}$ .



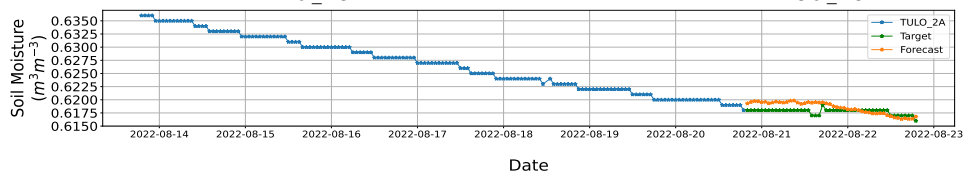
(e) Pronóstico de  $\theta_{TU\_MIA}$ , usando transfer-learnig de la red base  $\theta_{CU\_LOA}$ .



(f) Pronóstico de  $\theta_{TU\_MI2A}$ , usando transfer-learnig de la red base  $\theta_{CU\_LOA}$ .



(g) Pronóstico de  $\theta_{TU\_LOA}$ , usando transfer-learnig de la red base  $\theta_{CU\_LOA}$ .



(h) Pronóstico de  $\theta_{TU\_LO2A}$ , usando transfer-learnig de la red base  $\theta_{CU\_LOA}$ .

**Figura N° 4.21:** Pronósticos en diferentes momentos de humedad del suelo en pajonal  $\theta_{TU}$ , en todos los perfiles y horizontes analizados.

**Tabla N° 4.8:** Métricas obtenidas para las redes profundas que pronostican 48 horas de  $\theta_{TU}$ , dada la información de las variables:  $P$ ,  $T$ ,  $HR$ ,  $RS$ ,  $\theta_{TU}$  y  $T_{TU}$ .

Variable objetivo	Conjunto de datos	Métricas			
		Loss	MAE	RMSE	MSLE
$\theta_{TU\_UPA}$	Entrenamiento	<b>3.6e-6</b>	<b>1.4e-3</b>	<b>1.9e-3</b>	<b>1.42e-6</b>
	Validación	2.49e-5	3.7e-3	5e-3	9.69e-5
	Evaluación	2.11e-5	3.39e-3	4.6e-3	8.16e-6
$\theta_{TU\_UP2A}$	Entrenamiento	<b>3.9e-6</b>	<b>1.4e-3</b>	<b>2e-3</b>	<b>1.46e-06</b>
	Validación	2.58e-5	3.7e-3	5.1e-3	9.5e-6
	Evaluación	2.84e-5	3.6e-3	5.3e-3	1.04e-5
$\theta_{TU\_URA}$	Entrenamiento	<b>4.95e-6</b>	<b>1.5e-3</b>	<b>2.2e-3</b>	<b>1.94e-6</b>
	Validación	3.82e-5	3.9e-3	6.2e-3	1.52e-5
	Evaluación	2.38e-5	3.2e-3	4.88e-3	9.63e-6
$\theta_{TU\_UR2A}$	Entrenamiento	<b>7.64e-6</b>	<b>1.7e-3</b>	<b>2.8e-3</b>	<b>3.34e-6</b>
	Validación	3.22e-5	2.7e-3	5.7e-3	1.42e-5
	Evaluación	9.13e-6	2.02e-3	3.03e-3	4.07e-6
$\theta_{TU\_MIA}$	Entrenamiento	<b>3.65e-6</b>	<b>1.4e-3</b>	<b>1.9e-3</b>	<b>1.44e-6</b>
	Validación	2.38e-5	3.5e-3	4.9e-3	9.32e-6
	Evaluación	2.70e-5	3.37e-3	5.20e-3	1.05e-5
$\theta_{TU\_MI2A}$	Entrenamiento	<b>5.27e-6</b>	<b>1.5e-3</b>	<b>2.3e-3</b>	<b>2.13e-6</b>
	Validación	3.03e-5	3.0e-3	5.7e-3	1.39e-5
	Evaluación	1.77e-5	2.21e-3	4.21e-3	6.95e-6
$\theta_{TU\_LOA}$	Entrenamiento	<b>3.43e-6</b>	<b>1.3e-3</b>	<b>1.9e-3</b>	<b>1.40e-6</b>
	Validación	2.14e-5	3.1e-3	4.6e-3	8.4e-6
	Evaluación	2.51e-5	2.87e-3	5.01e-3	9.70e-6
$\theta_{TU\_LO2A}$	Entrenamiento	<b>3.27e-6</b>	<b>1.2e-3</b>	<b>1.8e-3</b>	<b>1.26e-6</b>
	Validación	1.49e-5	2.2e-3	3.9e-3	5.64e-6
	Evaluación	2.70e-5	2.24e-3	4.54e-3	7.77e-6

#### 4.3.4 Aplicación de la red base para pronóstico de humedad del suelo en almohadilla usando transfer-learning

En esta sección se presentan los resultados de la aplicación de las técnicas de *transfer-learning*, en dominios similares, presentando pronósticos y la evaluación de los resultados

para almohadilla, con las métricas establecidas en la metodología.

#### 4.3.4.1 Pronóstico de humedad del suelo en almohadilla ( $\theta_{CU}$ )

En este apartado podemos ver el conjunto de Figuras 4.22, donde gracias al conocimiento adquirido desde la red base y con las técnicas de *transfer-learning*, se pudo pronosticar 48 horas de humedad del suelo de todos los perfiles y horizontes de almohadilla, donde se utilizó 7 días de información horaria de entrada, con el mismo dominio que la red base:  $P, T, HR, RS, \theta_{CU}$  y  $T_{CU}$

Las Figuras 4.22(a) a 4.22(g), muestran pronósticos para diferentes momentos de las series de evaluación:  $\theta_{CU\_UPA}$ ,  $\theta_{CU\_UP2A}$ ,  $\theta_{CU\_URA}$ ,  $\theta_{CU\_UR2A}$ ,  $\theta_{CU\_MIA}$ ,  $\theta_{CU\_MI2A}$ ,  $\theta_{CU\_MIA}$  y  $\theta_{CU\_LO2A}$ .

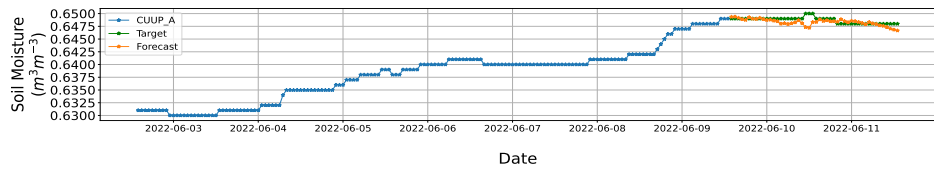
#### 4.3.4.2 Evaluación de los modelos de pronóstico de humedad del suelo en almohadilla ( $\theta_{CU}$ )

La Tabla 4.9, presenta métricas de rendimiento para modelos de redes profundas que pronostican 48 horas de  $\theta_{CU}$  en función de las variables:  $P, T, HR, RS, \theta_{CU}$  y  $T_{CU}$ . Las métricas incluyen los estadísticos: Loss, MAE, RMSE y MSLE; para los conjuntos de datos de: entrenamiento, validación y evaluación.

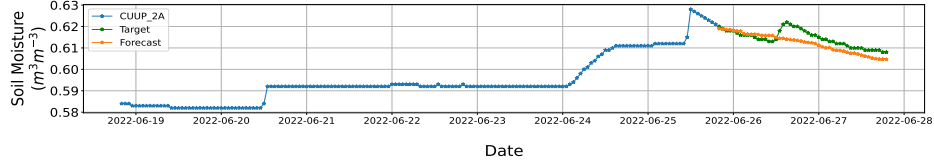
Una característica destacable de esta tabla son los valores resaltados en negrita. Por ejemplo, en el caso de  $\theta_{CU\_UPA}$  en el conjunto de entrenamiento, se observa una pérdida (Loss) extraordinariamente baja de  $1.4e-6$ , lo que refleja la capacidad excepcional del modelo para ajustarse a los datos de entrenamiento. Este nivel de precisión se refuerza con un MAE y un RMSE de  $2.3e-3$  y  $3.8e-3$  respectivamente, indicando un rendimiento sobresaliente en la predicción de  $\theta_{CU}$ .

Los valores MSLE también son notables, con cifras en el orden de  $1e-6$ , lo que sugiere que los modelos son altamente efectivos al predecir en un entorno logarítmico.

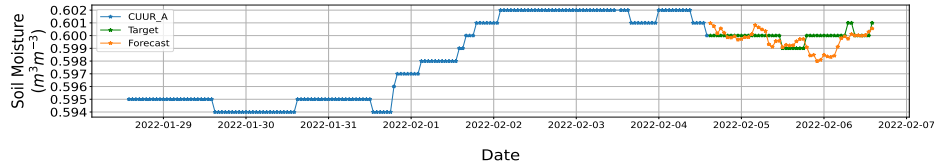
En general, esta tabla subraya el éxito de los modelos de redes profundas en la predicción de  $\theta_{CU}$ , demostrando una precisión excepcional en una variedad de conjuntos de datos. Estos resultados respaldan la utilidad potencial de estos modelos en aplicaciones prácticas



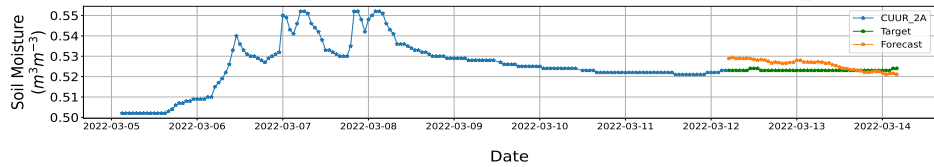
(a) Pronóstico de  $\theta_{CU\_UPA}$ , usando transfer-learnig de la red base  $\theta_{CU\_LOA}$ .



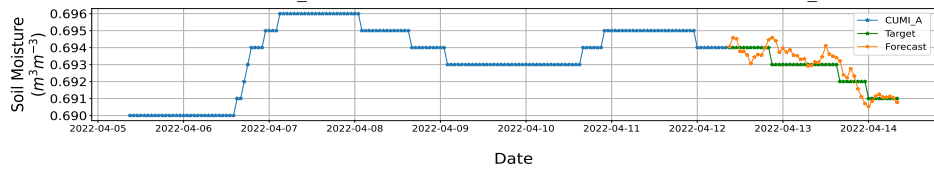
(b) Pronóstico de  $\theta_{CU\_UP2A}$ , usando transfer-learnig de la red base  $\theta_{CU\_LOA}$ .



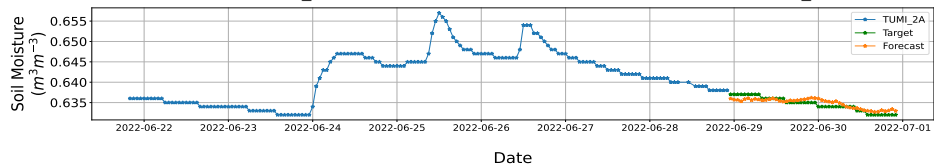
(c) Pronóstico de  $\theta_{CU\_URA}$ , usando transfer-learnig de la red base  $\theta_{CU\_LOA}$ .



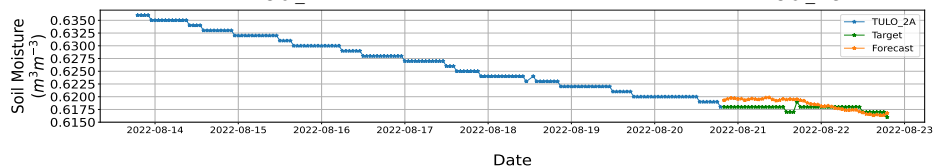
(d) Pronóstico de  $\theta_{CU\_UR2A}$ , usando transfer-learnig de la red base  $\theta_{CU\_LOA}$ .



(e) Pronóstico de  $\theta_{CU\_MIA}$ , usando transfer-learnig de la red base  $\theta_{CU\_LOA}$ .



(f) Pronóstico de  $\theta_{CU\_MI2A}$ , usando transfer-learnig de la red base  $\theta_{CU\_LOA}$ .



(g) Pronóstico de  $\theta_{CU\_LO2A}$ , usando transfer-learnig de la red base  $\theta_{CU\_LOA}$ .

**Figura N° 4.22:** Pronósticos en diferentes momentos de humedad del suelo en almohadilla  $\theta_{CU}$ , en todos los perfiles y horizontes analizados.

que requieren pronósticos precisos basados en múltiples entradas.

**Tabla N° 4.9:** Métricas obtenidas para las redes profundas que pronostican 48 horas de  $\theta_{CU}$ , dada la información de las variables:  $P$ ,  $T$ ,  $HR$ ,  $RS$ ,  $\theta_{CU}$  y  $T_{CU}$ .

Variable objetivo	Conjunto de datos	Métricas			
		Loss	MAE	RMSE	MSLE
$\theta_{CU\_UPA}$	Entrenamiento	<b>1.4e-6</b>	<b>2.3e-3</b>	<b>3.8e-3</b>	<b>5.26e-6</b>
	Validación	2.64e-5	3.7e-3	5.1e-3	9.7e-5
	Evaluación	2.87e-5	3.55e-3	5.36e-3	1.08e-5
$\theta_{CU\_UP2A}$	Entrenamiento	<b>5.27e-6</b>	<b>1.5e-3</b>	<b>2.3e-3</b>	<b>2.13e-06</b>
	Validación	3.3e-5	3e-3	5.7e-3	1.3e-5
	Evaluación	1.77e-5	2.21e-3	4.21e-3	6.95e-6
$\theta_{CU\_URA}$	Entrenamiento	<b>4.95e-6</b>	<b>1.5e-3</b>	<b>2.2e-3</b>	<b>1.94e-6</b>
	Validación	3.82e-5	3.9e-3	6.2e-3	1.52e-5
	Evaluación	2.38e-5	3.2e-3	4.88e-3	9.63e-6
$\theta_{CU\_UR2A}$	Entrenamiento	<b>1.59e-5</b>	<b>2.5e-3</b>	<b>4e-3</b>	<b>6.95e-6</b>
	Validación	2.96e-5	3.4e-3	5.4e-3	1.28e-5
	Evaluación	4.03e-5	3.84e-3	6.35e-3	1.76e-5
$\theta_{CU\_MIA}$	Entrenamiento	<b>9.32e-6</b>	<b>2.1e-3</b>	<b>3.1e-3</b>	<b>3.36e-6</b>
	Validación	3.08e-5	3.8e-3	5.6e-3	1.09e-5
	Evaluación	7.03e-5	5.55e-3	8.39e-3	2.61e-5
$\theta_{CU\_MI2A}$	Entrenamiento	<b>7.84e-6</b>	<b>1.9e-3</b>	<b>2.8e-3</b>	<b>3.27e-6</b>
	Validación	2.96e-5	3.5e-3	5.4e-3	1.2e-5
	Evaluación	7.03e-5	5.55e-3	8.39e-3	2.61e-5
$\theta_{CU\_LO2A}$	Entrenamiento	<b>3.27e-6</b>	<b>1.2e-3</b>	<b>1.8e-3</b>	<b>1.26e-6</b>
	Validación	1.49e-5	2.2e-3	3.9e-3	5.64e-6
	Evaluación	2.70e-5	2.24e-3	4.54e-3	7.77e-6

## CONCLUSIONES

### • Descripción de la climatología y humedad del suelo

Las técnicas de clusterización aplicadas a la información de precipitación, permitió establecer zonas homogéneas de lluvia con respecto a la estacionalidad, además se verificó que el Distrito Metropolitano de Quito y sus principales sistemas de abastecimiento hídrico poseen tres principales zonas de estacionalidad asociadas a: comportamiento amazónico (principalmente unimodal, con su pico máximo entre mayo-junio y su mínimo entre diciembre-enero), comportamiento interandino (normalmente bimodal con sus picos en abril y octubre y sus mínimos aproximadamente en julio-agosto y diciembre-enero) y un comportamiento asociado a la lluvia del pacífico del norte de sudamérica (principalmente unimodal con una zona de alta pluviosidad entre marzo-abril y baja precipitación de junio-octubre); se debe señalar también que la distribución de la precipitación planteada en la Figura 4.6, muestra que existen zonas intermedias donde se presenta un cambio de fase.

Se debe destacar que la estacionalidad de JTU\_AWD en Jatunhuayco es claramente bimodal, con sus picos en abril y noviembre, sus zonas de baja pluviosidad se encuentran en enero y agosto, además de demostró que la lluvia está asociada al comportamiento usual de páramo, donde predominan los eventos de baja intensidad pero alta duración, como se muestra en la Figura 4.8, además el ciclo diurno de la lluvia muestra según la Figura 4.7, los eventos más intensos de lluvia se producen entre los meses de febrero a junio y de septiembre a diciembre, donde se presentan los máximos registros en horas de la tarde aproximadamente entre las 13h00 y 17h00.

El análisis exploratorio de los datos de suelo registrados, muestra que la microcuenca JTU\_AWD en Jatunhuayco, posee una temperatura media del suelo que en almohadilla oscila entre  $[6.44 < T_{CU} < 7.21]^{\circ}C$ , mientras que en pajonal varía entre  $[6.1 < T_{TU} < 7.54]^{\circ}C$ , por lo tanto, en los dos tipos de suelo y considerando todos los horizontes y perfiles analizados, es claro que la temperatura media del suelo no presenta grandes

variaciones. Por otra parte, la información de humedad media del suelo da cuenta de que en almohadilla está dentro del intervalo  $[0.5 < \Theta_{CU} < 0.67] \text{cm}^3 \text{cm}^{-3}$ , mientras que en pajonal varía entre  $[0.5 < \Theta_{TU} < 0.64] \text{cm}^3 \text{cm}^{-3}$ , por lo tanto, es claro que la humedad media del suelo tampoco presenta grandes variaciones y predomina una alta concentración volumétrica del agua.

- **Descripción de la relación entre precipitación y humedad del suelo**

Se verificó que la humedad del suelo analizada a escalas temporales: horaria, diaria y hasta quincenal es casi invariante, esto se puede comprobar observando la Figura 4.12, es claro también que los efectos en la humedad del suelo son directos como consecuencia del comportamiento pluviométrico de eventos de lluvia y días secos, es decir, se requieren de varios días de lluvia o días secos para comenzar a observar cambios notorios en la humedad del suelo. Es importante notar que desde la frecuencia quincenal observada en la Figura 4.12c, las señales de la precipitación y humedad del suelo comienzan a parecerse, mientras que al aumentar el período a mensual en la Figura 4.12d, el parecido es mayor todavía, pero considerando un desfase temporal de la humedad del suelo de aproximadamente un mes.

Del análisis del ciclo diario de las diferencias consecutivas de la humedad del suelo en almohadilla y pajonal, es claro que las señales en pajonal son más intensas, esto se traduce diciendo que las fluctuaciones son también mayores, es decir que los flujos de agua en almohadilla son menos notorios que en pajonal, por otro lado, si se comparan estas señales con las del ciclo diario pluviométrico, se verifica que coinciden, lo cual tiene sentido ya que la humedad del suelo es consecuencia de la lluvia, además en la Figura 4.16, se muestra con un ejemplo, cómo reacciona la humedad del suelo en los dos diferentes tipos de cobertura vegetal analizados, es así que en la Figura 4.16b, se aprecia que el inicio de la respuesta de la humedad del suelo en almohadilla, se da una hora después del inicio de la lluvia y coincide con el máximo valor de humedad del suelo de almohadilla en ese período, luego se aprecia claramente como se va atenuando la humedad del suelo hasta que se estabiliza. Se debe indicar que el análisis realizado con información a baja frecuencia es válido y se justifica gracias al análisis realizado



del efecto de la precisión del reflectómetro y la propagación de posibles errores.

- **Redes neuronales de humedad del suelo**

Se emplearon modelos del estado del arte aplicando redes neuronales, CNN-LSTM-ANN, para pronosticar 48 horas de humedad del suelo a varias profundidades y en diferentes ubicaciones usando información horaria de siete días. La red neuronal usada tiene una arquitectura profunda con hiperparámetros optimizados, permitiendo que las neuronas interpreten correctamente la relación entre las variables hidrometeorológicas y humedad del suelo. La técnica de transfer-learning empleada, network-based transfer learning, ha servido de doble propósito: se demostró la capacidad de asociar el conocimiento de un dominio de particular, *i.e.* un perfil de suelo, hacia otro relacionado, *i.e.* un perfil de suelo diferente ubicado en otro sitio; por otro lado, de manera indirecta se optimizó el tiempo de entrenamiento de las redes neuronales que usaron la red base entrenada. Las métricas obtenidas para el pronóstico indican que los resultados son alentadores, con errores de orden entre  $1 \times 10^{-6} < \epsilon < 1 \times 10^{-3}$ .

## RECOMENDACIONES Y TRABAJO FUTURO

El presente estudio ha demostrado la eficacia del uso de redes neuronales y técnicas de transfer-learning en el pronóstico de la humedad del suelo en los Andes tropicales del norte de Ecuador. Sin embargo, existen numerosas oportunidades para ampliar y aplicar estos resultados en diferentes áreas de interés.

### RECOMENDACIONES

1. **Ampliar el alcance espacial y temporal:** Considerar la inclusión de más estaciones de monitoreo y aumentar el período de estudio para obtener una visión más completa de las variaciones climáticas y de humedad del suelo en la región. Esto podría mejorar la capacidad de generalización de los modelos.
2. **Optimización o inclusión de variables de entrada:** Analizar cuáles son las variables que más influyen en el pronóstico de humedad del suelo y explorar la inclusión de otras variables ambientales, para mejorar la capacidad de los modelos para capturar las interacciones complejas entre el clima y la humedad del suelo.
3. **Optimización continua de hiperparámetros:** Continuar ajustando los hiperparámetros de los modelos de redes neuronales para obtener un rendimiento aún mejor. La optimización de hiperparámetros puede mejorar significativamente la precisión de las predicciones.
4. **Integración con telemetría:** Incorporar sistemas de telemetría para obtener información a tiempo real y poder hacer predicciones casi al instante.
5. **Aplicaciones prácticas:** Explorar aplicaciones prácticas de los pronósticos de humedad del suelo en áreas como la agricultura, la gestión del agua y la mitigación de desastres. Colaborar con expertos en estas áreas para implementar soluciones basadas en los resultados.

6. **Actualización de modelos:** Dado que las condiciones climáticas pueden cambiar con el tiempo, mantener los modelos actualizados y recalibrados regularmente para garantizar que sigan siendo precisos en condiciones cambiantes.

## **TRABAJO FUTURO**

Una de las posibles líneas de investigación futura es la aplicación de los modelos desarrollados en la predicción de derrumbes o movimientos en masa. La humedad del suelo desempeña un papel crucial en la estabilidad de los taludes y pendientes, y su pronóstico preciso podría proporcionar una herramienta valiosa para la evaluación y prevención de deslizamientos. La integración de datos de precipitación, temperatura y humedad del suelo en los modelos podría permitir identificar umbrales críticos y generar alertas tempranas para mitigar los riesgos asociados con estos eventos.

Otra área de aplicación potencial es el pronóstico de incendios. La disponibilidad de información precisa sobre la humedad del suelo puede ayudar a evaluar el nivel de riesgo de incendios forestales. Combinando los datos de humedad del suelo con variables climáticas relevantes, como precipitación, temperatura, humedad relativa y radiación solar, se podrían desarrollar modelos predictivos capaces de identificar condiciones propicias para la propagación del fuego. Estos pronósticos podrían contribuir a la planificación y la toma de decisiones en la gestión del riesgo de incendios, permitiendo una respuesta más efectiva y oportuna.

Además, se vislumbra la posibilidad de utilizar los resultados de esta investigación como insumo para estimar la cantidad de agua disponible en una cuenca. La humedad del suelo es un indicador clave de la disponibilidad de agua para la vegetación y los recursos hídricos en general. Al combinar los datos de humedad del suelo obtenidos a través de modelos de pronóstico con información adicional sobre las características hidrológicas de una cuenca, se podrían desarrollar estimaciones confiables de la cantidad de agua almacenada y su disponibilidad para diversos usos, como la agricultura, el abastecimiento de agua potable y la generación de energía hidroeléctrica.

Estas son solo algunas de las posibilidades que se abren a partir de los resultados obtenidos en este estudio, donde se podría explorar y ampliar estas aplicaciones, así como a

considerar otras áreas en las que el pronóstico preciso de la humedad del suelo pueda tener un impacto significativo. El desarrollo de modelos más complejos, la incorporación de nuevas variables y el refinamiento de las técnicas de transfer-learning son desafíos prometedores para futuras investigaciones en este campo.

## Bibliografía

- Ahmed, A., Deo, R. C., Ghahramani, A., Raj, N., Feng, Q., Yin, Z., Yang, L., et al. (2021). LSTM integrated with Boruta-random forest optimiser for soil moisture estimation under RCP4.5 and RCP8.5 global warming scenarios. *Stochastic Environmental Research and Risk Assessment*, 35(9), 1851-1881.
- Ahmed, A. M., Deo, R. C., Raj, N., Ghahramani, A., Feng, Q., Yin, Z., & Yang, L. (2021). Deep learning forecasts of soil moisture: convolutional neural network and gated recurrent unit models coupled with satellite-derived MODIS, observations and synoptic-scale climate index data. *Remote Sensing*, 13(4), 554.
- Alexander, L., & Herold, N. (2016). ClimPACT2. *Indices and software. A document prepared on behalf of the Commission for Climatology (CCI) Expert Team on Sector-Specific Climate Indices (ET-SCI)*, 46.
- Alpaydin, E. (2016). *Machine learning: the new AI*. MIT press.
- An, X., & Zhao, F. (2021). Prediction of soil moisture based on BP neural network optimized search algorithm. *IOP Conference Series: Earth and Environmental Science*, 714(2), 022046.
- Andrades, M., Moliner, A., & Masaguer, A. (2015). Prácticas de edafología: métodos didácticos para análisis de suelos. *Editorial Universidad La Rioja, Logroño*.
- Andrienko, N., Aandrienko, G., & Andrienko, G. (2006). *Exploratory analysis of spatial and temporal data: a systematic approach*. Springer Science & Business Media.
- Arnbjerg-Nielsen, K., Willems, P., Olsson, J., Beecham, S., Pathirana, A., Bülow Gregersen, I., Madsen, H., & Nguyen, V.-T.-V. (2013). Impacts of climate change on rainfall extremes and urban drainage systems: a review. *Water science and technology*, 68(1), 16-28.
- Bartels, G. K., dos Reis Castro, N. M., Pedrollo, O., & Collares, G. L. (2021). Soil moisture estimation in two layers for a small watershed with neural network models: Assessment of the main factors that affect the results. *CATENA*, 207, 105631.

- Bendix, J., Lauer, W., & Lauer, W. (1992). Die Niederschlagsjahreszeiten in Ecuador und ihre klimadynamische Interpretation (Rainy seasons in Ecuador and their climate-dynamic interpretation). *Erdkunde*, 118-134.
- Bendix, J., Rafiqpoor, M. D., & Rafiqpoor, M. D. (2001). Studies on the Thermal Conditions of Soils at the Upper Tree Line in the Páramo of Papallacta:(Eastern Cordillera of Ecuador)(Studien der thermischen Bedingungen der Böden an der oberen Waldgrenze des Páramo de Papallacta (Ostkordillere Ecuador)). *Erdkunde*, 257-276.
- Berg, A., & Sheffield, J. (2018). Climate change and drought: the soil moisture perspective. *Current Climate Change Reports*, 4(2), 180-191.
- Bonnesoeur, V., Locatelli, B., Guariguata, M. R., Ochoa-Tocachi, B. F., Vanacker, V., Mao, Z., Stokes, A., & Mathez-Stiefel, S.-L. (2019). Impacts of forests and forestation on hydrological services in the Andes: A systematic review. *Forest Ecology and Management*, 433, 569-584.
- Brutsaert, W. (2013). *Evaporation into the atmosphere: theory, history and applications* (Vol. 1). Springer Science & Business Media.
- Buytaert, W., Célleri, R., De Bièvre, B., Cisneros, F., Wyseure, G., Deckers, J., & Hofstede, R. (2006). Human impact on the hydrology of the Andean páramos. *Earth-Science Reviews*, 79(1-2), 53-72.
- Cadilhac, L., Torres, R., Calles, J., Vanacker, V., & Calderón, E. (2017). Desafíos para la investigación sobre el cambio climático en Ecuador. *Neotropical Biodiversity*, 3(1), 168-181.
- Casas-Castillo, C. (2005). *Análisis espacial y temporal de las lluvias extremas en Catalunya. Modelización y clasificación objetiva* (Ph.D. Tesis). Universidad de Barcelona. Departament d'Astronomia i Meteorologia.
- Chatterjee, S., Dey, N., & Sen, S. (2020). Soil moisture quantity prediction using optimized neural supported model for sustainable agricultural applications. *Sustainable Computing: Informatics and Systems*, 28, 100279.
- Chen, T. (1974). COMPARISON OF JARDI AND WORKMAN RATE-OF-RAINF ALL GAUGES. *Technical Note (Local) No, 18*.

- Constante-Segura, J., & Palacios-Chacón, E. (2014). El recurso solar para generación de energía.
- Deza, M. M., Deza, E., & Deeza, E. (2009). Encyclopedia of distances. En *Encyclopedia of distances* (pp. 1-583). Springer.
- Dubois, A., Teytaud, F., & Verel, S. (2021). Short term soil moisture forecasts for potato crop farming: A machine learning approach. *Computers and Electronics in Agriculture*, *180*, 105902.
- Duffie, J. A., Beckman, W. A., & Blair, N. (2020). *Solar engineering of thermal processes, photovoltaics and wind*. John Wiley & Sons.
- EPMAPS. (2018). Actualización del plan de manejo del área de conservación hídrica Antisana. *Empresa Pública Metropolitana de Agua Potable y Saneamiento (EPMAPS), Fondo para la Protección del Agua (FONAG), Quito, Ecuador*.
- EPMAPS. (2022). Empresa Pública Metropolitana de Agua Potable Y Saneamiento.
- Escobar-González, D., Singaña-Chasi, M. S., González-Vergara, J., Erazo, B., Zambrano, M., Acosta, D., Villacís, M., Gualpa, M., Lahuatte, B., & Peluffo-Ordóñez, D. H. (2022). Intensity-Duration-Frequency Curve for Extreme Rainfall Event Characterization, in the High Tropical Andes. *Water*, *14*(19), 2998.
- Fang, K., & Shen, C. (2020). Near-real-time forecast of satellite-based soil moisture using long short-term memory with an adaptive data integration kernel. *Journal of Hydro-meteorology*, *21*(3), 399-413.
- FONAG. (2022). Fideicomiso Mercantil Fondo Ambiental para la Protección de las Cuen-cas y Agua. En *Fideicomiso Mercantil Fondo Ambiental para la Protección de las Cuencas y Agua*. Routledge.
- Gamboa, J. C. B. (2017). Deep learning for time-series analysis. *arXiv preprint arXiv:1701.01887*.
- Gao, P., Qiu, H., Lan, Y., Wang, W., Chen, W., Han, X., & Lu, J. (2021). Modeling for the Prediction of Soil Moisture in Litchi Orchard with Deep Long Short-Term Memory. *Agriculture*, *12*(1), 25.

- Gao, P., Xie, J., Yang, M., Zhou, P., Chen, W., Liang, G., Chen, Y., Han, X., & Wang, W. (2021). Improved soil moisture and electrical conductivity prediction of citrus orchards based on IOT using Deep Bidirectional LSTM. *Agriculture*, *11*(7), 635.
- García, A. (2006). *Hietogramas de diseño en zonas urbanas*. División de Ciencias Básicas e Ingeniería. <http://148.206.53.231/tesiuami/UAMI13294.pdf>
- Glickman, T. S. (2000). *Glossary of meteorology*. American Meteorological Society.
- González-Fernández, R. (2017). Nubes y otros meteoros. Vistos desde el Observatorio Meteorológico de Oviedo y su entorno.
- González-Vergara, J., Escobar-González, D., Chaglla-Aguagallo, D., & Peluffo-Ordóñez, D. H. (2020). A Data-Driven Approach for Automatic Classification of Extreme Precipitation Events: Preliminary Results, 197-209.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT press.
- Gronau, I., & Moran, S. (2007). Optimal implementations of UPGMA and other common clustering algorithms. *Information Processing Letters*, *104*(6), 205-210.
- Gultepe, I. (2008). Fog and boundary layer clouds: fog visibility and forecasting.
- Han, H., Choi, C., Kim, J., Morrison, R. R., Jung, J., & Kim, H. S. (2021). Multiple-depth soil moisture estimates using artificial neural network and long short-term memory models. *Water*, *13*(18), 2584.
- Hegazi, E. H., Yang, L., & Huang, J. (2021). A Convolutional Neural Network Algorithm for Soil Moisture Prediction from Sentinel-1 SAR Images. *Remote Sensing*, *13*(24), 4964.
- Hurwitz, J., Kirsch, D., & Kirsch, D. (2018). Machine learning for dummies, IBM Limited Edition. *IBM Limited Edition*, 75.
- Ibrahim, K. S. M. H., Huang, Y. F., Ahmed, A. N., Koo, C. H., & El-Shafie, A. (2022). A review of the hybrid artificial intelligence and optimization modelling of hydrological streamflow forecasting. *Alexandria Engineering Journal*, *61*(1), 279-303.
- INAMHI. (2022). Instituto Nacional de Meteorología e Hidrología.



- Iñiguez, V., Morales, O., Cisneros, F., Bauwens, W., & Wyseure, G. (2016). Analysis of the drought recovery of Andosols on southern Ecuadorian Andean páramos. *Hydrology and Earth System Sciences*, 20(6), 2421-2435.
- Irrgang, C., Boers, N., Sonnewald, M., Barnes, E. A., Kadow, C., Staneva, J., & Saynisch-Wagner, J. (2021). Towards neural Earth system modelling by integrating artificial intelligence in Earth system science. *Nature Machine Intelligence*, 3(8), 667-674.
- Johnson, S. C. (1967). Hierarchical clustering schemes. *Psychometrika*, 32(3), 241-254.
- Jones, L. V. (1987). *The Collected Works of John W. Tukey: Philosophy and Principles of Data Analysis 1965-1986* (Vol. 4). CRC Press.
- Joshi, A. V. (2020). Machine learning and artificial intelligence.
- Jove, F., Hernández, R., & Caballero, Á. (2020). IDF Curves and Maximum Rainfall in 24 hours in the Subregions of La Mojana and San Jorge in Northern Colombia. *International Journal of Engineering Research and Technology*, 13(10), 2884-2894. <https://doi.org/10.37624/IJERT/13.10.2020.2884-2894>
- Kelleher, J. D., Tierney, B., & Tierney, B. (2018). *Data science*. MIT Press.
- Kelly, E. F., Chadwick, O. A., & Hilinski, T. E. (1998). The effect of plants on mineral weathering. *Biogeochemistry*, 42(1), 21-53.
- Ketkar, N., & Santana, E. (2017). *Deep learning with Python* (Vol. 1). Springer.
- Kim, T., Yang, T., Gao, S., Zhang, L., Ding, Z., Wen, X., Gourley, J. J., & Hong, Y. (2021). Can artificial intelligence and data-driven machine learning models match or even replace process-driven hydrologic models for streamflow simulation?: A case study of four watersheds with different hydro-climatic regions across the CONUS. *Journal of Hydrology*, 598, 126423.
- Klemeš, V. (1986). Dilettantism in hydrology: Transition or destiny? *Water Resources Research*, 22(9S), 177S-188S.
- Kreyszig, E. (1991). *Introductory functional analysis with applications* (Vol. 17). John Wiley & Sons.
- Kuttippurath, J., Murasingh, S., Stott, P., Sarojini, B. B., Jha, M. K., Kumar, P., Nair, P., Varikoden, H., Raj, S., Francis, P., et al. (2021). Observed rainfall changes in the

- past century (1901–2019) over the wettest place on Earth. *Environmental Research Letters*, 16(2), 024018.
- Lahuatte, B., Mosquera, G. M., Páez-Bimos, S., Calispa, M., Vanacker, V., Zapata-Ríos, X., Muñoz, T., & Crespo, P. (2022). Delineation of water flow paths in a tropical Andean headwater catchment with deep soils and permeable bedrock. *Hydrological Processes*, 36(10), e14725.
- Lange, H. (2020). Machine learning applications in hydrology. En *Forest-water interactions* (pp. 233-257). Springer.
- Legates, D. R., & McCabe Jr, G. J. (1999). Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. *Water resources research*, 35(1), 233-241.
- Li, Q., Wang, Z., Shangguan, W., Li, L., Yao, Y., & Yu, F. (2021). Improved daily SMAP satellite soil moisture prediction over China using deep learning model with transfer learning. *Journal of Hydrology*, 600, 126698.
- Liu, D., Liu, C., Tang, Y., & Gong, C. (2022). A GA-BP Neural Network Regression Model for Predicting Soil Moisture in Slope Ecological Protection. *Sustainability*, 14(3), 1386.
- Mohyont, B., Demarée, G. R., & Faka, D. N. (2004). Establishment of IDF-curves for precipitation in the tropical area of Central Africa - comparison of techniques and results. *Natural Hazards and Earth System Sciences*, 4(3), 375-387. <https://doi.org/10.5194/nhess-4-375-2004>
- Molina, A., Govers, G., Vanacker, V., Poesen, J., Zeelmaekers, E., & Cisneros, F. (2007). Runoff generation in a degraded Andean ecosystem: Interaction of vegetation cover and land use. *Catena*, 71(2), 357-370.
- Molina, A., Vanacker, V., Corre, M. D., & Veldkamp, E. (2019). Patterns in soil chemical weathering related to topographic gradients and vegetation structure in a high Andean tropical ecosystem. *Journal of Geophysical Research: Earth Surface*, 124(2), 666-685.

- Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, 50(3), 885-900.
- Müllner, D. (2011). Modern hierarchical, agglomerative clustering algorithms. *arXiv preprint arXiv:1109.2378*.
- Murphy, K. P. (2012). *Machine learning: a probabilistic perspective*. MIT press.
- Murtagh, F., & Legendre, P. (2014). Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion? *Journal of classification*, 31(3), 274-295.
- Nicholson, S., Farrar, T., & Farrar, T. (1994). The influence of soil type on the relationships between NDVI, rainfall, and soil moisture in semiarid Botswana. I. NDVI response to rainfall. *Remote sensing of environment*, 50(2), 107-120.
- Noguchi, S., Nik, A. R., Yusop, Z., Tani, M., & Sammori, T. (1997). Rainfall-runoff responses and roles of soil moisture variations to the response in tropical rain forest, Bukit Tarek, Peninsular Malaysia. *Journal of Forest Research*, 2(3), 125-132.
- Ochoa-Tocachi, B. F., Buytaert, W., Antiporta, J., Acosta, L., Bardales, J. D., Célleri, R., Crespo, P., Fuentes, P., Gil-Ríos, J., Guallpa, M., et al. (2018). High-resolution hydro-meteorological data from a network of headwater catchments in the tropical Andes. *Scientific data*, 5(1), 1-16.
- Ochoa-Tocachi, B. F., Buytaert, W., De Bievre, B., Célleri, R., Crespo, P., Villacís, M., Llerena, C. A., Acosta, L., Villazón, M., Guallpa, M., et al. (2016). Impacts of land use on the hydrological response of tropical Andean catchments. *Hydrological Processes*, 30(22), 4074-4089.
- Oki, T., Kanae, S., & Kanae, S. (2006). Global hydrological cycles and world water resources. *science*, 313(5790), 1068-1072.
- Páez-Bimos, S. (2020). Control de calidad y análisis de datos hidro-meteorológicos subcuenca 1 - Jatunhuyaco.
- Páez-Bimos, S. (2022). *Soil-water process interactions in a landscape context in the Andean páramo* (Tesis doctoral). UCL-Université Catholique de Louvain.

- Páez-Bimos, S., Molina, A., Calispa, M., Delmelle, P., Lahuate, B., Villacís, M., Muñoz, T., & Vanacker, V. (2022). Soil-vegetation-water interactions controlling solute flow and transport in volcanic ash soils of the high Andes. *Hydrology and Earth System Sciences Discussions*, 1-37.
- Páez-Bimos, S., Villacís, M., Morales, O., Calispa, M., Molina, A., Salgado, S., de Bievre, B., Delmelle, P., Muñoz, T., & Vanacker, V. (2022). Vegetation effects on soil pore structure and hydraulic properties in volcanic ash soils of the high Andes. *Hydrological Processes*, e14678.
- Páliz Larrea, P., Zapata Ríos, X., & Campozano Parra, L. (2021). Application of Neural Network Models and ANFIS for Water Level Forecasting of the Salve Faccha Dam in the Andean Zone in Northern Ecuador. *Water*, 13(15), 2011.
- Patiño, D. T., Sánchez, P. C., & Rojas, G. M. (2018). Umbrales en la respuesta de humedad del suelo a condiciones meteorológicas en una ladera Altoandina. *Maskana*, 9(2), 53-65.
- Petneházi, G. (2019). Recurrent neural networks for time series forecasting. *arXiv preprint arXiv:1901.00069*.
- Prasad, R., Deo, R. C., Li, Y., & Maraseni, T. (2018). Soil moisture forecasting by a hybrid machine learning technique: ELM integrated with ensemble empirical mode decomposition. *Geoderma*, 330, 136-161.
- Prasad, R., Deo, R. C., Li, Y., & Maraseni, T. (2019). Weekly soil moisture forecasting with multivariate sequential, ensemble empirical mode decomposition and Boruta-random forest hybridizer algorithm approach. *Catena*, 177, 149-166.
- Quiring, S. M., Ford, T. W., & Yuan, S. (2015). Climate of the Critical Zone. En *Developments in Earth Surface Processes* (pp. 79-111). Elsevier.
- Ramos, M. (2001). Divisive and hierarchical clustering techniques to analyse variability of rainfall distribution patterns in a Mediterranean region. *Atmospheric Research*, 57(2), 123-138.
- Rincón, L. (2014). Introducción a la probabilidad.

- Rodríguez, R., Benito, Á., & Portela, A. (2004). Meteorología y climatología. *Fundación española para la ciencia y la tecnología. Villena Artes Gráficas. España*, 12-16.
- Rostami, S., He, J., & Hassan, Q. K. (2018). Riverine water quality response to precipitation and its change. *Environments*, 5(1), 8.
- Rueda Bayona, J., Elles Pérez, C., Sánchez, E., González, A., & Rivillas Ospina, G. (2016). Identificación de patrones de variabilidad climática a partir de análisis de componentes principales, Fourier y clúster k-medias. *Revista Tecnura*, 55-68. <https://doi.org/10.14483/udistrital.jour.tecnura.2016.4.a04>
- 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, 100924.
- Russell, S. J. (2010). *Artificial intelligence a modern approach*. Pearson Education, Inc.
- Santos, C. A. G., Brasil Neto, R. M., da Silva, R. M., & Costa, S. G. F. (2019). Cluster analysis applied to spatiotemporal variability of monthly precipitation over Paraíba state using tropical rainfall measuring mission (TRMM) data. *Remote Sensing*, 11(6), 637.
- Schmugge, T., Jackson, T., & McKim, H. (1980). Survey of methods for soil moisture determination. *Water Resources Research*, 16(6), 961-979.
- Scientific, C. (2017). Inc. CS650 and CS655 Water Content Reflectometers. *Campbell Scientific, Inc., Logan, UT*.
- Segura, H., Junquas, C., Espinoza, J. C., Vuille, M., Jauregui, Y. R., Rabatel, A., Condom, T., & Lebel, T. (2019). New insights into the rainfall variability in the tropical Andes on seasonal and interannual time scales. *Climate dynamics*, 53(1), 405-426.
- Seneviratne, S. I., Corti, T., Davin, E. L., Hirschi, M., Jaeger, E. B., Lehner, I., Orlowsky, B., & Teuling, A. J. (2010). Investigating soil moisture–climate interactions in a changing climate: A review. *Earth-Science Reviews*, 99(3-4), 125-161.

- Sheppard, B., & Joe, P. (2008). Performance of the precipitation occurrence sensor system as a precipitation gauge. *Journal of atmospheric and Oceanic technology*, 25(2), 196-212.
- Silva Robledo, H., Silva Candia, P., Garrido Salinas, M., & Acevedo Hinojosa, E. (2015). Manual de estudio y ejercicios relacionados con el contenido de agua en el suelo y su uso por los cultivos.
- Sit, M., Demiray, B. Z., Xiang, Z., Ewing, G. J., Sermet, Y., & Demir, I. (2020). A comprehensive review of deep learning applications in hydrology and water resources. *Water Science and Technology*, 82(12), 2635-2670.
- Sivasankaran, S., Mohan, K. J., & Nazer, G. M. (2021). Soil Moisture Quantity Prediction using Optimized Deep Learning Supported model for Sustainable cultivation of Groundnut plant. *NVEO-NATURAL VOLATILES & ESSENTIAL OILS Journal| NVEO*, 13184-13205.
- Sklenar, P., Kuècerová, A., Mackova, J., & Macek, P. (2015). Temporal variation of climate in the high-elevation páramo of Antisana, Ecuador. *Suppl. Geogr. Fis. Din. Quat*, 38, 67-78.
- Suebsombut, P., Sekhari, A., Sureephong, P., Belhi, A., & Bouras, A. (2021). Field data forecasting using LSTM and Bi-LSTM approaches. *Applied Sciences*, 11(24), 11820.
- Sun, Z., Sandoval, L., Crystal-Ornelas, R., Mousavi, S. M., Wang, J., Lin, C., Cristea, N., Tong, D., Carande, W. H., Ma, X., et al. (2022). A review of earth artificial intelligence. *Computers & Geosciences*, 105034.
- Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., & Liu, C. (2018). A survey on deep transfer learning. *International conference on artificial neural networks*, 270-279.
- Topp, G. C., Davis, J., & Annan, A. P. (1980). Electromagnetic determination of soil water content: Measurements in coaxial transmission lines. *Water resources research*, 16(3), 574-582.
- Trenberth, K. E., Fasullo, J. T., & Kiehl, J. (2009). Earth's global energy budget. *Bulletin of the American Meteorological Society*, 90(3), 311-324.

- Vanacker, V., Molina, A., Rosas-Barturen, M., Bonnesoeur, V., Román-Dañobeytia, F., Ochoa-Tocachi, B. F., & Buytaert, W. (2022). The effect of natural infrastructure on water erosion mitigation in the Andes. *Soil*, 8(1), 133-147.
- Vanacker, V., Molina, A., Torres, R., Calderon, E., & Cadilhac, L. (2018). Challenges for research on global change in mainland Ecuador.
- Vyas, A., & Bandyopadhyay, S. (2020). Dynamic Structure Learning through Graph Neural Network for Forecasting Soil Moisture in Precision Agriculture.
- Willems, P. (2000). Compound intensity/duration/frequency-relationships of extreme precipitation for two seasons and two storm types. *Journal of Hydrology*, 233(1-4), 189-205.
- Yapo, P. O., Gupta, H. V., & Sorooshian, S. (1996). Automatic calibration of conceptual rainfall-runoff models: sensitivity to calibration data. *Journal of hydrology*, 181(1-4), 23-48.
- Yuan, Y., Xiong, D., Wu, H., Zhang, S., Zhang, B., Dahal, N. M., Liu, L., Li, W., Zhang, W., & Shi, L. (2020). Spatial variation of soil physical properties and its relationship with plant biomass in degraded slopes in dry-hot valley region of Southwest China. *Journal of Soils and Sediments*, 20(5), 2354-2366.
- Zahumenský, I. (2004). Guidelines on quality control procedures for data from automatic weather stations. *World Meteorological Organization, Switzerland*, 1-10.
- Zhang, L., Xue, Z., Zhang, Y., Ma, J., & Li, H. (2022). Enhanced Generalized Regression Neural Network for Soil Moisture Estimation Over the Qinghai-Tibet Plateau. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15, 3815-3829.
- Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., & He, Q. (2020). A comprehensive survey on transfer learning. *Proceedings of the IEEE*, 109(1), 43-76.