

Reducción de Escala Espacial de Datos Diarios de Precipitación y Temperatura para Estudios de
Salud Pública en el Departamento del Cauca, Colombia

Kevin David Blanco Silva

Trabajo de Grado para Optar el título de Magíster en Recursos Hídricos y Saneamiento
Ambiental

Director

Sandra Rocío Villamizar Amaya

Ph.D. in Environmental Systems

Codirector

Álvaro Javier Ávila Díaz

Ph.D. in Applied Meteorology

Universidad Industrial de Santander

Facultad de Ingenierías Fisicomecánicas

Escuela de Ingeniería Civil

Maestría en Recursos Hídricos y Saneamiento Ambiental

Bucaramanga

2023

Dedicatoria

A mi familia, quienes son mi mayor motivación día a día; a mi madre Ángela a quien le debo todo; a mi abuela Nelly y mi tía Martha quienes me han apoyado toda mi vida; a mi novia Alejandra, que me ha acompañado y apoyado en cada paso en este proceso; y a mis gatos Joey, Penny, Angus, Chandler, Jess y Barney, que me hicieron compañía a lo largo del proyecto.

Agradecimientos

Agradezco a mis directores de tesis Dra. Sandra R. Villamizar y Dr. Álvaro Ávila, por la orientación y cada una de las enseñanzas recibidas en todo este proceso, las cuales me permitieron culminar con éxito todas las metas propuestas en el proyecto y tener un crecimiento profesional y académico.

Adicionalmente agradezco a la Secretaría de Salud Departamental del Cauca, la Universidad de Ciencias Aplicadas y Ambientales, el Instituto Nacional de Salud por el apoyo en la ejecución del proyecto y al Ministerio de Ciencia, Tecnología e Innovación de Colombia por la financiación de la tesis y el proyecto 210484467217.

Contenido

	Pág.
Introducción	13
1. Descripción del Problema	15
2. Justificación	18
3. Marco de Referencia	20
3.1 Fuentes de Datos Climáticos.....	21
3.2 Cambio de Escala Espacial de Variables Climáticas	24
3.2.1 Métodos Estadísticos	24
3.2.2 Métodos de Inteligencia Artificial	26
4. Objetivos	27
4.1 Objetivo General	27
4.2 Objetivos Específicos.....	27
5. Zona de Estudio	27
6. Metodología	30
6.1 Fase I - Selección de una Metodología Estadística Apropriada para la Reducción de Escala Espacial de Precipitación y Temperatura para el Caso de Estudio	31
6.1.1 Caracterización de las Metodologías Estadísticas de Reducción de Escala de Variables Climatológicas	31
6.1.2 Caracterización del Sitio de Estudio e Información Relevante para el Estudio.	31
6.1.3 Selección de la Metodología de Reducción de Escala a Implementar.....	32

6.2 Fase II - Aplicación de la Metodología Estadística Seleccionada para la Reducción de Escala Espacial de las Variables de Interés Generando un Banco de Datos Climatológicos del Departamento del Cauca, Colombia	33
6.2.1 Procesamiento de la Información Obtenida de las Variables Meteorológicas de Interés....	33
6.2.2 Implementación de la metodología seleccionada para la reducción de escala espacial	36
6.3 Fase III Determinación del Grado de Incertidumbre de los Datos Escalados por Medio de Métricas Estadísticas.....	40
7. Resultados.....	42
7.1 Metodología Estadística Seleccionada para la Reducción de Escala Espacial	42
7.2 Banco de Datos Climatológicos Escalados del Departamento del Cauca, Colombia.....	50
7.2.1 Base de Datos Cruda de las Variables Meteorológicas de Interés.....	51
7.2.2 Base de Datos Escalada de las Variables Meteorológicas de Interés	58
7.3 Grado de Incertidumbre	64
7.3.1 Temperatura Máxima.....	64
7.3.2 Temperatura Mínima	68
7.3.3 Temperatura Promedio.....	72
7.3.4 Precipitación	74
8. Discusión.....	80
8.1 Datos Sin Escalar	80
8.2 Datos Escalados	82
8.3 Beneficios Proyecto INS.....	85
9. Conclusiones	86
10. Recomendaciones	87

Referencias Bibliográficas	89
----------------------------------	----

Lista de Tablas

	Pág.
Tabla 1. <i>Bases de Datos Climatológicas Globales Relevantes</i>	32
Tabla 2. <i>Metodologías de Reducción de Escala Espacial</i>	42
Tabla 3. <i>Datos Estaciones IDEAM</i>	52

Lista de Figuras

	Pág.
Figura 1. <i>Ubicación de la Zona de Estudio</i>	28
Figura 2. <i>Distribución Espacial Estaciones IDEAM en el Departamento del Cauca</i>	30
Figura 3. <i>Portal web de consulta y descarga de datos Hidrometeorológicos del IDEAM (DHIME)</i>	34
Figura 4. <i>Portal Web de Descarga de Datos CHIRPS</i>	35
Figura 5. <i>Diagrama de Flujo Reducción de Escala</i>	40
Figura 6. <i>Estaciones IDEAM con Datos Disponibles</i>	51
Figura 7. <i>Porcentaje de Datos Faltantes Estaciones IDEAM Seleccionadas</i>	55
Figura 8. <i>Precipitación Total Diaria CHIRPS</i>	56
Figura 9. <i>Variables Disponibles ERA5-Land</i>	57
Figura 10. <i>Variables Disponibles MSWX</i>	58
Figura 11. <i>Temperatura Máxima Diaria Escalada ERA5-Land</i>	59
Figura 12. <i>Temperatura Máxima Diaria Escalada MSWX</i>	59
Figura 13. <i>Temperatura Mínima Diaria Escalada ERA5-Land</i>	60
Figura 14. <i>Temperatura Mínima Diaria Escalada MSWX</i>	60
Figura 15. <i>Temperatura Promedio Diaria Escalada ERA5-Land</i>	61
Figura 16. <i>Temperatura Promedio Diaria Escalada MSWX</i>	61
Figura 17. <i>Precipitación Total Diaria Escalada ERA5-Land</i>	62
Figura 18. <i>Precipitación Total Diaria Escalada MSWX</i>	63
Figura 19. <i>Precipitación Total Diaria Escalada CHIRPS</i>	64
Figura 20. <i>Métricas Temperatura Máxima Diaria</i>	66

Figura 21. <i>Distribución MAE Noviembre - Temperatura Máxima Diaria</i>	67
Figura 22. <i>Métricas Temperatura Mínima Diaria</i>	70
Figura 23. <i>Distribución MAE marzo - Temperatura Mínima Diaria</i>	71
Figura 24. <i>Métricas Temperatura Promedio Diaria</i>	73
Figura 25. <i>Métricas Precipitación Total Diaria</i>	76
Figura 26. <i>Distribución MAE enero - Precipitación Total Diaria</i>	77
Figura 27. <i>Métricas Precipitación Total Diaria Escalada (DEM y NDVI)</i>	79

Lista de Apéndices

Los Apéndices están adjuntos en un archivo comprimido y puede visualizarlos en base de datos de la biblioteca UIS.

Apéndice A. Scripts Escalamiento de Datos

Apéndice B. Consolidación Mitologías para la Reducción de Escala

Apéndice C. Bases de Datos Estaciones IDEAM e INS

Apéndice D. Base de Datos CHIRPS

Apéndice E. Base de Datos ERA5-Land

Apéndice F. Base de Datos MSWX

Apéndice G. Datos Escalados Temperatura Máxima Diaria

Apéndice H. Datos Escalados Temperatura Mínima Diaria

Apéndice I. Datos Escalados Temperatura Promedio Diaria

Apéndice J. Datos Escalados Precipitación Total Diaria

Apéndice K. Métricas Estadísticas Estimadas

Resumen

Título: Reducción de Escala Espacial de Datos Diarios de Precipitación y Temperatura para Estudios de Salud Pública en el Departamento del Cauca, Colombia*

Autor: Kevin David Blanco Silva**

Palabras Claves: Hidrología, meteorología, clima, Kriging.

Descripción

La variación de los parámetros climáticos en el planeta afecta los servicios ecosistémicos que la humanidad recibe del medio ambiente. Este es el caso de la salud humana pues esta variabilidad puede generar un ambiente propicio para la propagación de enfermedades sensibles al clima como es el caso del dengue. Se han desarrollado sistemas de alerta temprana (SAT) para el control, prevención y monitoreo de la propagación de este tipo de enfermedades permitiendo a las entidades responsables de la toma de decisiones trabajar hacia la minimización de los picos de contagio. Para la construcción de estos SAT se requiere la generación de mapas de favorabilidad ambiental que se alimentan de información meteorológica de alta resolución espacial con el objeto de ser aplicados a escalas locales; esto representa una gran limitante en Colombia pues, si bien se cuenta con información meteorológica, mucha de ella presenta fallas o datos ausentes, y no tiene la resolución espacial requerida para la implementación de dichos sistemas. En este proyecto se identificó e implementó la metodología geoestadística Kriging para la reducción de escala espacial de series diarias de las variables meteorológicas precipitación total y temperatura máxima, mínima y promedio en el departamento del departamento del Cauca, para su uso en el proyecto que adelanta el Instituto Nacional de Salud (INS) denominado “Estratificación espacial del dengue basado en la identificación de factores de riesgo: un ensayo piloto en el departamento del Cauca”. El proceso de escalamiento se aplicó a las bases de datos disponibles en las fuentes de información satelital y de reanálisis CHIRPS (Climate Hazards Group InfraRed Precipitation With Station Data), ERA5-Land (European Centre for Medium-Range Weather Forecasts) y MSWX (Multi-Source Weather), logrando una reducción de escala de la resolución de 0.1° ofrecida por las bases de datos MSWX y ERA5-Land y 0.05° de CHIRPS a una escala de 0.01° en todas las variables. Los datos escalados se evaluaron respecto a las mediciones in situ de las estaciones IDEAM (Instituto de Hidrología, Meteorología y Estudios Ambientales) usando métricas de desempeño relevantes (*RMSE*, *MAE*, *r*, *MBE*). Las métricas obtenidas de las bases de datos escaladas, presentaron un mejor rendimiento respecto a los bancos de datos originales. Adicionalmente, no se verificó una única fuente de datos que tuviera las mejores métricas de desempeño para todas las variables analizadas. Usando la métrica MAE, (diferencia absoluta entre el valor observado y el estimado) se encontró que ERA5-Land ofrece los mejores datos escalados para las variables precipitación total y temperatura mínima diaria (MAE oscilando entre 3 mm a 10 mm y 1 °C a 2.5 °C, respectivamente) mientras que MSWX es una mejor fuente para el caso de las variables temperatura máxima y promedio diaria (MAE oscilando entre 1.5 °C a 3 °C y 0.5 °C a 1.5 °C, respectivamente). Sumado a lo anterior, se identificó que las métricas obtenidas para temperatura se encuentran cercanos a los valores presentados por otras metodologías, sin embargo, para el caso de la variable precipitación se obtuvieron datos con un rendimiento inferior al presentado por otros estudios. No obstante, las bases de datos meteorológicas generadas en el proceso de escalamiento permitieron al Instituto Nacional de Salud (INS) mejorar los resultados en la estimación de modelos de reproducción del vector *Aedes Aegypti*, requeridos para la construcción de los sistemas de alerta temprana de dengue en los municipios de Patía, Piamonte y Miranda del departamento del Cauca.

* Trabajo de Grado

** Facultad de Ingenierías Fisicomecánicas. Escuela de Ingeniería Civil. Maestría en Recursos Hídricos y Saneamiento Ambiental. Director: Sandra Rocío Villamizar Amaya, Ph.D. in Environmental Systems. Codirector: Álvaro Javier Ávila Díaz, Ph.D. in Applied Meteorology

Abstract

Title: Spatial Downscaling of Daily Precipitation and Temperature Data for Public Health Studies in the Department of Cauca, Colombia*

Author: Kevin David Blanco Silva**

Keywords: Hydrology, meteorology, climate, Kriging.

Description

The variation of climatic parameters on the planet affects the ecosystem services that humanity receives from the environment. This is the case of human health, as this can create an environment conducive to the spread of climate-sensitive diseases such as dengue. Early warning systems (SAT) have been developed for the control, prevention, and monitoring of the spread of this type of disease that allow the entities responsible for decision-making to work towards minimizing contagion peaks. For the construction of these SATs, the generation of environmental favorability maps that are fed with high spatial resolution meteorological information is required in order to be applied at local scales; this represents a great limitation in Colombia because, although meteorological information is available, much of it presents faults or missing data, and does not have the spatial resolution required for the implementation of said systems. In this project, the Kriging geostatistical methodology was identified and implemented for the reduction of the spatial scale of daily series of the meteorological variables total precipitation and maximum, minimum and average temperature in the department of Cauca, for use in the project carried out by the National Institute of Health (INS) called "Spatial stratification of dengue based on the identification of risk factors: a pilot trial in the department of Cauca". The scaling process was applied to the databases available in the satellite information and reanalysis sources CHIRPS (Climate Hazards Group InfraRed Precipitation with Station Data), ERA5-Land (European Center for Medium-Range Weather Forecasts) and MSWX (Multi -Source Weather), achieving a 0.1° resolution downscaling offered by the MSWX and ERA5-Land databases and 0.05° from CHIRPS to a 0.01° scale in all variables. Scaled data was evaluated against in situ measurements from IDEAM (Institute of Hydrology, Meteorology and Environmental Studies) stations using relevant performance metrics (RMSE, MAE, r, MBE). The metrics obtained from the escalation databases presented a better performance with respect to the original databases. In addition, there was no single data source verified that had the best performance metrics for all the variables analyzed. Using the MAE metric (absolute difference between the observed and the estimated value) it was found that ERA5-Land offers the best scaled data for the variables total precipitation and minimum daily temperature (MAE ranging between 3 mm to 10 mm and 1°C to 2.5°C , respectively) while MSWX is a better source for the case of the maximum and daily average temperature variables (MAE oscillating between 1.5°C to 3°C and 0.5°C to 1.5°C , respectively). Added to the above, it was identified that the metrics obtained for temperature are close to the values presented by other methodologies, however, in the case of the precipitation variable, data was obtained with a lower performance than that presented by other studies. However, the meteorological databases generated in the scaling process allowed the National Institute of Health (INS) to improve the results in the estimation of reproduction models of the *Aedes Aegypti* vector, required for the construction of dengue early warning systems. in the municipalities of Patía, Piamonte and Miranda in the department of Cauca.

* Degree work

** Faculty of Physical-Mechanical Engineering. School of Civil Engineering. Master's Degree in Water Resources and Environmental Sanitation. Director: Sandra Rocío Villamizar Amaya, Ph.D. in Environmental Systems. Co-director: Álvaro Javier Ávila Díaz, Ph.D. in Applied Meteorology

Introducción

Las alteraciones de las variables meteorológicas promueven ambientes propicios para la transmisión de enfermedades sensibles al clima como aquellas propagadas mediante vectores (OPS, 2021). Debido a esto, ha crecido la importancia del desarrollo de sistemas de alerta temprana (SAT) dentro del sistema de salud, con el objeto de brindar a las entidades las herramientas necesarias para la prevención y control de estas patologías (PAHO, 2019). En Colombia, el dengue ha sido una de las enfermedades propagadas mediante vectores que más ha afectado a la población, registrando un total de 1.4 millones de casos entre los años 1990 – 2016 (Padilla et al., 2017); por esto, surgió la necesidad de crear un SAT para la construcción de mapas de riesgo de dengue que permitiera la toma de decisiones institucionales en busca de disminuir las afectaciones de esta enfermedad sobre el sistema de salud.

Con el objeto de crear un SAT, el Instituto Nacional de Salud de Colombia (INS), se encuentra desarrollando un proyecto denominado “Estratificación espacial del dengue basado en la identificación de factores de riesgo: un ensayo piloto en el departamento del Cauca”. Este proyecto busca “Determinar la relación entre los factores ambientales, biológicos y sociodemográficos que influyen en el aumento de la carga de dengue y su variación espacial en los municipios de alto riesgo del departamento del Cauca”. Sin embargo, la creación del SAT requiere de información climatológica continua y con una escala espacial que permita representar las condiciones ambientales en los procesos de reproducción y supervivencia de los vectores dentro de los centros poblados de los municipios de interés (Louis et al., 2014). Lo anterior representa un reto en el país debido a que, aunque existe una red de estaciones climatológicas administrada por el Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM), esta red no tiene una

densidad suficiente (número de estaciones por km^2), las estaciones no están distribuidas apropiadamente en todo el territorio nacional (la mayor concentración se ubica en la región andina y cerca a centros urbanos o agroindustriales), los datos en muchos casos no son continuos en el tiempo, y las ventanas de tiempo para las que hay datos disponibles son altamente irregulares.

Debido a lo anterior, surge la necesidad de buscar fuentes alternas de información, evaluar su grado de incertidumbre, y seleccionar un banco de datos que permitan la construcción de los SAT. Dentro de estas fuentes alternas se encuentran las bases de datos de productos satelitales y reanálisis como CHIRPS, ERA5-Land y MSWX, que permiten obtener información meteorológica continua y a una resolución espacial entre 0.05° a 0.1° (equivalente a 5.6 km y 11.1 km en el ecuador terrestre). Sin embargo, esta resolución ofrecida limita la creación de los mapas de favorabilidad ambiental requeridos en los SAT. Por lo tanto, este estudio implementa una metodología estadística para la reducción de escala espacial de fuentes alternas de información para las variables climáticas precipitación y temperatura, generando insumos que permitan al INS obtener mejores resultados en la creación del SAT en los centros poblados de los municipios de Patía, Piamonte y Miranda del departamento del Cauca.

1. Descripción del Problema

Los patrones climáticos del planeta han venido presentando variaciones debido a los efectos directos e indirectos de las actividades humanas sobre el medio ambiente; todo esto se ha visto reflejado en alteraciones globales como cambios de las temperaturas superficiales (atmosféricas y oceánicas), derretimiento de los glaciares, reducción de la capa de nieve, disminución del hielo marino, incremento del nivel del mar, acidificación oceánica y aumento del vapor de agua atmosférico (USGCRP, 2017). Estas alteraciones del clima han generado afectaciones en el ciclo hidrológico terrestre; una de ellas es la reducción de la disponibilidad (en calidad y cantidad) de agua apta para el abastecimiento de las necesidades humanas, ahondando en la actual problemática de acceso al agua potable que afecta cerca de 2200 millones de personas alrededor del mundo, e intensificada por la problemática de crecimiento poblacional desmedido (UNESCO, 2020). Estas variaciones en el ciclo del agua influyen negativamente sobre los servicios ecosistémicos de los cuales se benefician las comunidades, generando así riesgos en la producción de energía, la salud humana y el desarrollo económico (UNESCO, 2020).

Las alteraciones de los patrones hidro-climatológicos a nivel local tienen un impacto sobre la salud pública debido a que estas pueden promover un ambiente propicio para la propagación de enfermedades sensibles al clima como la malaria, el dengue, el cólera, la filariasis, la leptospirosis y la esquistosomiasis; todas ellas con afectaciones importantes en los índices de mortalidad y morbilidad, especialmente de las poblaciones más vulnerables (OPS, 2021). Como respuesta a esta problemática, en el marco de la tercera conferencia mundial sobre salud y cambio climático, la Organización Mundial de la Salud (OMS) ha planteado dentro de sus objetivos la construcción de

sistemas de salud resilientes al clima, mediante la integración de sistemas de alerta temprana climática (SATc) en los sistemas de información de salud (PAHO, 2019).

El dengue es una enfermedad arboviral transmitida por el *Aedes Aegypti* y *Ea. Albopictus*. En Colombia se presentan, en promedio, 312.85 casos de dengue por cada cien mil habitantes ubicados en las zonas de riesgo (Acosta Cardona, 2015). Su transmisión se debe a diversos factores entomológicos, socioeconómicos y climáticos, siendo estos últimos los que impactan significativamente en la distribución espacial y temporal de la enfermedad debido a la sensibilidad que presentan las poblaciones de *A. Aegypti* a los cambios de las variables climatológicas (temperatura, humedad relativa, precipitación, evaporación). Los patrones de clima y su variación inciden en la cantidad y calidad de agua estancada, la cual es empleada por el vector para el proceso de implantación o difusión de huevos (Acosta Cardona, 2015). El dengue tiene patrones espaciales y temporales dinámicos, lo que dificulta atribuir los cambios epidemiológicos a un solo factor, por lo que se han desarrollado herramientas de vigilancia como los mapas de incidencia (Louis et al., 2014). Para la construcción de estos mapas se requieren datos climatológicos que permitan describir la idoneidad de las condiciones ambientales para la reproducción y supervivencia de los vectores; estos datos pueden ser registrados en campo o tomados a través de la teledetección. Sin embargo, la baja resolución espacial de esta información, independientemente de su fuente, dificulta la implementación de estas herramientas a una escala local (Louis et al., 2014).

Para el caso de Colombia, existe una red de estaciones climatológicas *in-situ* administrada por el Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM), que abarca un rango de elevaciones desde el nivel del mar de las costas Caribe y Pacífico hasta 4150 metros sobre el nivel del mar (m s. n. m.) en la cordillera de los Andes (Urrea et al., 2019). Sin embargo, esta

red de estaciones es dispersa; por ejemplo, para el caso de la variable precipitación, cuenta con 2046 estaciones meteorológicas distribuidas en todo el territorio nacional, lo que representa, en promedio, una estación por cada 558.651 km² (IDEAM, 2018). Adicionalmente, las estaciones no se encuentran distribuidas uniformemente en todo el país; el 76% de ellas están ubicadas entre las elevaciones de 0 a 2000 m s. n. m., concentrándose principalmente en las regiones Andina, Caribe y Pacífico (Condom et al., 2020).

Por otra parte, existen productos de teledetección a escala global que permiten comprender los patrones generales de las variables climáticas y sus extremos. Por ejemplo, los datos satelitales de precipitación global basados en la Técnica de Morphing (CMORPH - Climate Prediction Center Morphing Technique) del Centro de Predicción Climática (CPC) de la Administración Nacional Oceánica y Atmosférica (NOAA) (Xavier et al., 2016) y la base de datos CHIRPS (Climate Hazards Group Infrared Precipitation with Station Data), que usa una combinación de datos obtenidos por sensores remotos y calibra sus estimaciones haciendo uso las observaciones “in-situ” (Funk et al., 2015). La base de datos CHIRPS ha sido empleada para la evaluación del ciclo anual de precipitaciones en Colombia (Urrea et al., 2019) y la construcción de series temporales anuales de precipitación para la evaluación de la variabilidad climática producto de las fases cálida y fría del fenómeno El Niño-Oscilación del Sur (ENOS) (Giraldo-Osorio et al., 2022). Sin embargo, este tipo de recursos presenta una limitante en cuanto a la resolución de pixel disponible pues usualmente varía entre 0.05° y 0.1° (equivalente a 5.6 km y 11.1 km en el ecuador terrestre).

Por lo tanto, surge la necesidad de establecer una estrategia de trabajo que permita integrar las diferentes fuentes de información por medio de una reducción de escala para generar una base de datos de variables meteorológicas de alta resolución que posibilite el desarrollo de análisis a

escala local de fenómenos de interés como es el caso de la ocurrencia y propagación de enfermedades endémicas tales como el dengue en Colombia.

2. Justificación

El virus del dengue es considerado como la enfermedad viral transmitida por vectores más importante dado que anualmente cerca de 400 millones de personas son infectadas y, además, tiene los mayores índices de mortalidad y morbilidad en el mundo (Muñoz et al., 2021). En Colombia, por ejemplo, se registraron 1.4 millones de casos de dengue entre los años 1990 – 2016 (52 mil casos/año), 774 mil casos de Chikunguña entre los años 2014 al 2016 (258 mil casos/año), y 117 mil casos de fiebre de Zika entre los años 2015 – 2016 (58.5 mil casos/año) (Padilla et al., 2017). Datos más recientes para el caso específico del dengue indican que se reportaron 44171 casos en el 2018 (Reyes & Instituto Nacional de Salud, 2019), un total de 256605 casos entre los años 2019 al 2021 (85.5 mil casos/año) y, para el año 2022, entre los meses de enero a noviembre se presentaron 62987 casos (INS, 2022).

Adicionalmente, el aumento en la incidencia del dengue que se ha evidenciado en las últimas décadas se ha relacionado con factores como el crecimiento de la población, la urbanización y el cambio climático (Muñoz et al., 2021), y se ha identificado que los picos epidemiológicos están influenciados por la variabilidad de las condiciones climáticas y ambientales (Ceccato et al., 2018). Por ejemplo, a finales del 2010 se registró el brote epidémico de dengue más importante presentado en el departamento del Cauca (Padilla et al., 2012) que coincidió con la presencia de la fase fría del ENOS (“La Niña”) en el territorio nacional (IDEAM, 2023). El estudio de la ocurrencia y propagación del dengue en este departamento resulta de alto

interés pues presenta las condiciones geográficas y ecológicas ideales para favorecer la propagación hipo- y meso-endémica del dengue en todo su territorio, lo cual, genera la presencia de comportamientos temporales, estacionales y cíclicos de la enfermedad en poblaciones de riesgo (INS, 2020).

En este orden de ideas, para entender los mecanismos relacionados con la ocurrencia de la enfermedad en Colombia, el Instituto Nacional de Salud (INS) desarrolló un proyecto denominado “Estratificación espacial del dengue basado en la identificación de factores de riesgo: un ensayo piloto en el departamento del Cauca”. Este proyecto buscaba “Determinar la relación entre los factores ambientales, biológicos y sociodemográficos que influyen en el aumento de la carga de dengue y su variación espacial en los municipios de alto riesgo del departamento del Cauca”. El INS creó un Sistema de Alerta Temprana (SAT) soportado en sistemas de información geográfica que permitió la construcción de mapas de riesgo. Dado que el proyecto se desarrolló a una escala local (sectores de cabecera municipal), el INS recolectó información de campo relacionada con los factores biológicos y sociodemográficos en los diferentes sectores de los municipios de interés para el proyecto (Patía, Piamonte y Miranda) desde el año 2015 hasta el 2021. Para el caso de los factores ambientales, específicamente precipitación y temperatura, era necesario establecer una estrategia metodológica que permitiera generar información de alta resolución horizontal, con datos continuos tanto en el espacio como en el tiempo que se integrara al análisis de riesgo propuesto por el proyecto de investigación.

En este orden de ideas este estudio identificó, aplicó y evaluó una metodología estadística que permitió obtener datos de alta resolución, produciendo datos continuos tanto en el espacio como en el tiempo de las variables precipitación y temperatura con un grado de confiabilidad que

permitió al INS avanzar en la construcción del SAT de propagación de dengue en el departamento del Cauca, Colombia.

3. Marco de Referencia

Los SAT permiten monitorear una amenaza y pronosticar la ocurrencia de un evento mediante un conjunto de procedimientos e instrumentos estructurados y planificados para tal fin (UNESCO, 2011); han sido aplicados en diferentes campos como la gestión del riesgo de desastres naturales (Makinoshima et al., 2021), procesos de adaptación al cambio climático (Gil-Guirado et al., 2022) y procesos de prevención y control de enfermedades (Linze et al., 2022b). El desarrollo de SAT aplicados a la enfermedad del dengue requiere la elaboración de mapas de favorabilidad ambiental que permitan detectar las zonas donde existan las condiciones ambientales necesarias para el desarrollo del vector; esta herramienta permite comprender el proceso de distribución del virus, posibilitando la toma de decisiones en los programas de control y prevención de enfermedades por parte de los organismos de Salud (Rojas, 2016). Para la construcción de mapas de favorabilidad ambiental se requiere de información meteorológica (i.e. temperatura, humedad relativa, precipitación) de alta resolución espacial (1 km x 1 km o menor), que permita su aplicación a escala local con el objeto de monitorear el aumento de la población del vector a partir de la influencia del clima sobre los procesos de desarrollo y supervivencia, los ciclos gonadotróficos, así como de desarrollo y eclosión de los huevos (Rotela, 2019).

3.1 Fuentes de Datos Climáticos

Para la obtención de la información requerida en la construcción de los mapas de favorabilidad ambiental existen datos meteorológicos generados por sensores remotos que permiten complementar la información recolectada *in-situ* y, en muchos casos, se constituyen en la única fuente de información disponible. Este tipo de información se ha ido fortaleciendo con el mejoramiento de la tecnología para la captura y procesamiento de datos espaciales. En el caso de la precipitación para las zonas tropicales, la base de datos CHIRPS ofrece información de alta resolución espacial (0.05°) y temporal (desde 1981 hasta la actualidad) derivada de observaciones infrarrojas de duración de nubes frías (Funk et al., 2015), y ya ha sido utilizada para estudios de enfermedades propagadas mediante vectores, como es el caso de la evaluación del vínculo existente entre la transmisión de la malaria y los factores climáticos en Senegal (Fall et al., 2022).

Las estimaciones de precipitación mensual con base en datos CHIRPS han arrojado resultados de alta precisión en las cuencas de los ríos Beijiang, Huai y Liao en China (Zhang et al., 2022); sin embargo, se ha evidenciado que los productos satelitales varían su precisión debido a factores como la pendiente, el viento, la elevación y la evapotranspiración del sitio de interés (Wiwoho et al., 2021). No obstante, esta fuente de información ha sido validada para su uso en la zona tropical de América del Sur, obteniendo resultados satisfactorios para la estimación de precipitación mensual en el noreste de Brasil (Paredes-Trejo et al., 2017) y en los departamentos de Antioquia (López-Bermeo et al., 2022) y Nariño (Ocampo-Marulanda et al., 2022) en Colombia. Para el caso de datos diarios, la base de datos CHIRPS ha sido evaluada en el departamento de Antioquia, Colombia, permitiendo validar un buen funcionamiento en la mayoría de las condiciones climáticas de este departamento, incluyendo los periodos más intensos ENOS (López-Bermeo et al., 2022), en la provincia de Sumatra Septentrional en Indonesia, los datos

CHIRPS presentaron una baja correlación y un error promedio grande respecto a los datos de observación in-situ (Saragih et al., 2022), y en el continente africano presentando, para este último caso, el sesgo más alto en comparación con las fuentes de datos TAMSAT y GsMap (Macharia et al., 2022).

Adicional a la herramienta CHIRPS, se dispone de diversas fuentes para la obtención de datos de alta resolución de precipitación como MSWEP V2, que proporciona datos desde 1979 hasta 2017, con resolución horizontal de 0.1° y temporal de 3 horas, y fusiona de forma óptima una extensa gama de datos de precipitación (reanálisis y satelital), aunque su uso se encuentra limitado al público debido a la necesidad del licenciamiento que debe ser solicitado a través de www.gloh2o.org (Beck et al., 2019) (MSWEP - GloH2O, 2020). Otro recurso disponible para la obtención de información meteorológica es el Multi-Source Weather (MSWX), que ofrece datos con una resolución de 0.1° cada 3 horas, disponiendo de diez variables meteorológicas (precipitación, temperatura del aire, temperatura del aire mínima y máxima diaria, presión superficial, humedad relativa y específica, velocidad del viento y radiación descendente de onda corta y onda larga), desde el 1 de enero hasta la actualidad; estos datos se construyen a partir de la corrección y reducción del sesgo de datos ERA5 a partir de referencias climatológicas de alta resolución, y su acceso se encuentra limitado a la aprobación de la solicitud realizada mediante el portal web www.gloh2o.org/mswx (Beck et al., 2021) (MSWX - GloH2O, 2021).

Para la obtención de información satelital de temperatura existen fuentes como NOAA GHCN_CAMS, que ofrece series de datos de alta resolución ($0.5^\circ \times 0.5^\circ$) desde 1948 hasta el presente (Fan & van den Dool, 2008); NASA GISS Surface Temperature Analysis, que dispone de cuadrículas de $2^\circ \times 2^\circ$ de resolución con una escala temporal mensual entre 1880 hasta el presente (GISTEMP Team, 2022); la base de datos desarrollada por University of Delaware que brinda

cuadrículas mensuales de alta resolución ($0.5^{\circ} \times 0.5^{\circ}$) de temperatura y precipitación desde 1900 hasta 2014 (NOAA, 2021); y CPC Global Temperature, que entrega valores diarios de temperatura del alta resolución ($0.5^{\circ} \times 0.5^{\circ}$) desde 1979 hasta el presente (Xie et al., 2007). Además de las herramientas descritas anteriormente, existen diversos recursos para la obtención de datos meteorológicos como es el caso de WorldClim (WorldClim, 2022), que ha sido empleada para la consecución de datos de temperatura promedio anual, temperatura promedio máxima y mínima anual y precipitación anual, con una resolución aproximada de 1 km, dentro del estudio de vulnerabilidad de los ecosistemas tropicales al cambio climático en el sur de Ecuador (Eguiguren-Velepucha et al., 2016).

Adicionalmente, existen fuentes que brindan conjuntos de datos producto de reanálisis, como ERA5-Land, que proporciona un panorama de la evolución de las variables meteorológicas durante varias décadas. El proceso de reanálisis combina datos de modelos climatológicos con observaciones de todo el mundo en un conjunto de datos globalmente completo, ofreciendo información de alta resolución ($0.1^{\circ} \times 0.1^{\circ}$) (Copernicus Climate Change Service, 2019).

La información disponible en los recursos descritos se encuentra con una gran variedad de resoluciones que están asociadas al tamaño de la entidad más pequeña que puede ser representada en la superficie (Pérez Navarro, 2011). Estos datos, generalmente, son suministrados en formato ráster (unidimensional o multidimensional), en los cuales se considera que la superficie es dividida en un conjunto uniforme de celdas o píxeles; cada uno de estos contiene un valor de una variable y asume que este es homogéneo en toda la celda (Ortega, 2016).

3.2 Cambio de Escala Espacial de Variables Climáticas

El escalamiento, entendido como la aplicación de un proceso que transforma la información de una escala a otra, surge por las diferencias de resolución espacial entre la información meteorológica disponible en los diversos recursos y los requerimientos de resolución espacial de los proyectos de investigación o aplicación. El sentido de la transferencia de datos clasifica el proceso en dos tipos totalmente opuestos: uno plantea la ampliación (*upscaling*) y el otro la reducción de la escala (*downscaling*). En el proceso de ampliación de escala se obtienen datos de baja resolución a partir de datos de alta resolución; este ha sido aplicado para la cuantificación de la evapotranspiración terrestre (Awada et al., 2021) o la estimación de caudales ambientales (Młyński et al., 2020). Por otra parte, el proceso de reducción de escala, consistente en obtener datos de alta resolución a partir de datos de menor resolución (Río San José, 2010) se emplea generalmente para la obtención de información de alta resolución requeridas en procesos de análisis hidrológico y meteorológico (F. Chen et al., 2020a).

Las metodologías de reducción de escala que se han desarrollado a la fecha se agrupan en dos categorías: estadísticas y computacionales o de “machine learning”.

3.2.1 Métodos Estadísticos

Estos métodos generalmente implican la implementación de procesos de interpolación y extrapolación que permiten la construcción de cuadrículas a partir de estaciones locales. Entre los métodos más conocidos están los polígonos de Thiessen, la interpolación condicional, la ponderación de distancia inversa, la ponderación de distancia angular, las regresiones lineales múltiples, y el Kriging, entre otros; sus resultados son altamente influenciados por la densidad de la red de observación y las escalas espaciales y temporales que se trabajan (Serrano-Notivoli & Tejedor, 2021). Un caso de éxito de generación de grillas de alta resolución (0.25×0.25) en una

zona con alta densidad de estaciones meteorológicas es el estudio de (Xavier et al., 2016) que implementó los métodos de promedio aritmético, thinplate spline, natural neighbour, ponderación de distancia inversa, ponderación de distancia angular y el punto ordinario Kriging para obtener datos meteorológicos (temperatura máxima y mínima, radiación solar, humedad relativa, velocidad del viento y precipitación diaria) requeridos para la estimación de la evapotranspiración, a partir de 3625 pluviómetros y 735 estaciones meteorológicas.

Sen Gupta y Tarboton (2016) desarrollaron una herramienta de reducción de escala espacial con el objetivo de suministrar información para la construcción de modelos hidrológicos que requieren cuadrículas de resolución fina (30 a 100 m). Su estrategia permite generar superficies de cuadrícula de 3 horas de datos meteorológicos (temperatura, precipitación, velocidad del viento, humedad relativa, radiación de onda corta y onda larga) mediante datos satelitales y de reanálisis (con resolución inicial de 100-200 km), utilizando la interpolación espacial, las relaciones físicas entre las variables meteorológicas y la elevación (Sen Gupta & Tarboton, 2016).

Existen estudios dedicados a la evaluación del desempeño de varios métodos de interpolación. El estudio realizado por Hofstra et al. (2008) en el Reino Unido y los Alpes comparó los métodos de interpolación Kriging global y local, dos versiones de ponderación de distancia angular, natural neighbour, regresión, thinplate spline 2D y 3D e interpolación condicional para la obtención de cuadrículas de 5×5 km de las variables precipitación diaria, temperatura promedio, máxima y mínima, a partir de información tomada de 10900 sensores locales (Hofstra et al., 2008). Este estudio concluyó que el método de Kriging global presentó el mejor rendimiento para su uso en la consecución de datos diarios de alta resolución en Europa y el Reino Unido. Por otra parte, en Italia se desarrolló un estudio similar en el cual se implementaron los métodos de regresión multilínea con mejoras locales, el Kriging de regresión y la regresión lineal ponderada local,

empleando 1484 estaciones para construir mapas de temperatura mensual de alta resolución. Para este caso, los autores encontraron un rendimiento levemente superior en el método de regresión lineal ponderada (Brunetti et al., 2014).

3.2.2 Métodos de Inteligencia Artificial

Las redes neuronales artificiales son herramientas informáticas flexibles que pueden ser aplicadas en una amplia gama de problemas de predicción de series de tiempo con alto grado de precisión (Khashei & Bijari, 2010). Se han desarrollado estudios de aplicación de redes neuronales para el escalamiento de datos meteorológicos con resultados bastante satisfactorios. El estudio de Serifi et al., (2021) investigó la capacidad de aprendizaje de las redes neuronales para la reconstrucción de datos de alta resolución partiendo de simulaciones de baja resolución; el estudio de Retalis et al., (2017) desarrolló un método aplicando redes neuronales artificiales para la obtención de datos de precipitación de alta resolución ($1 \text{ km} \times 1 \text{ km}$) a partir de la correlación de datos CHIRPS, información de altitud y el índice de vegetación de diferencia normalizada de imágenes satelitales; finalmente, el estudio de Ma, (2022) empleó redes neuronales para la reducción de escala de datos de vapor de agua precipitable obtenidos del centro europeo de reanálisis 5 (ERA 5).

Recientemente se han desarrollado nuevos métodos computacionales fundamentados en el modelado de superficies mediante geometría diferencial. Por ejemplo, el método de reducción de escala basado en el modelado de superficie de alta precisión (HASMD por sus siglas en inglés), permite estimar series de datos de precipitaciones mensuales de alta calidad con resolución espacial de 0.01° (Zhao & Jiao, 2021); este ha sido empleado para mejorar la precisión de la estimación de la precipitación mensual en china (C.-L. Wang et al., 2015). Adicionalmente, con este tipo de métodos se logró calcular la distribución espacial de precipitaciones extremas en áreas

topográficamente complejas, obteniendo errores de precisión inferiores a otros métodos de interpolación (B. Q. Li et al., 2017). Si bien existe una tendencia creciente en el uso de los métodos de inteligencia artificial pues estos permiten mejorar los procesos de escalamiento, estos procesos requieren un alto consumo de recursos informáticos y alta experticia.

4. Objetivos

4.1 Objetivo General

Implementar una metodología de reducción de escala estadística para información diaria de precipitación y temperatura, con el fin de fortalecer los servicios climáticos para el sector de salud pública en el departamento del Cauca, Colombia.

4.2 Objetivos Específicos

Seleccionar una metodología estadística apropiada para la reducción de escala espacial de precipitación y temperatura para el caso de estudio.

Aplicar la metodología estadística seleccionada a las variables de interés, con el fin de crear un banco de datos climatológicos del departamento del Cauca, Colombia.

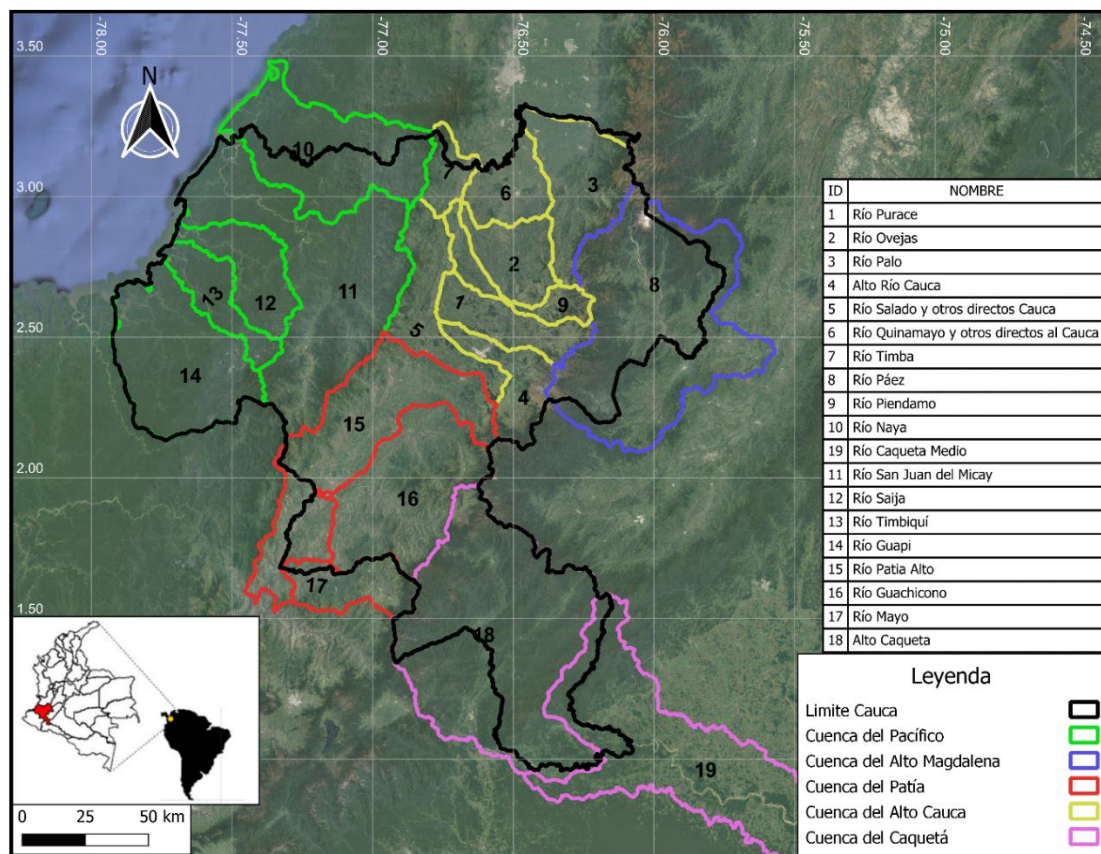
Determinar el grado de incertidumbre de los datos escalados por medio de métricas estadísticas.

5. Zona de Estudio

El departamento del Cauca cuenta con una superficie de 29308 Km² y se encuentra ubicado al suroeste de Colombia. En su superficie drenan cinco grandes cuencas: Alto Cauca, Pacífico, Alto Magdalena, Patía y Caquetá. La altura media en el Departamento es de 1693 m s.n.m. y su

perfil topográfico, con un gradiente entre 5750 m s.n.m. y el nivel del mar, le permite tener todos los pisos térmicos, desde el clima cálido húmedo al occidente en la costa Pacífica, cálido semiárido en la cuenca del Patía, templado húmedo y semihúmedo en la zona central, los climas fríos a lado y lado de la meseta de Popayán y las nieves perpetuas en el Nevado del Huila. (Universidad del Valle, 2022) (Ver Figura 1).

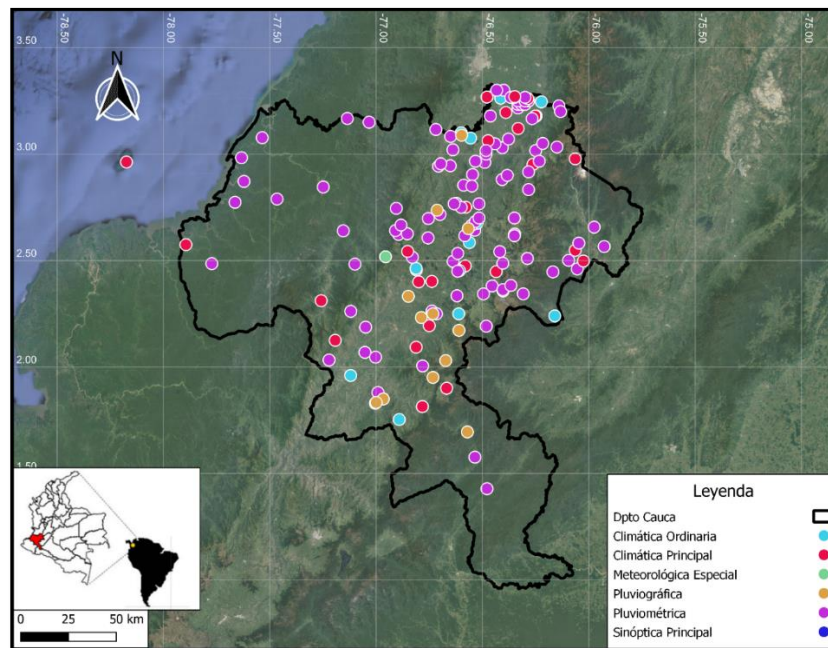
Figura 1. Ubicación de la Zona de Estudio



Nota. Imagen de elaboración propia, a partir del límite político del departamento del Cauca suministrado por el INS, el archivo *shape* de cuencas hidrográficas disponible en el portal web del Catálogo de Recursos Cartográficos Marino Costeros de Colombia, servicio del SIAM (DAMCRA-Subzonas Hidrograficas Colombia, 2022) y la información satelital de Google disponible en QGIS.

El departamento del Cauca presenta una amplia variabilidad espacial en la distribución de lluvias, registrando volúmenes anuales superiores a los 7000 mm hacia la cuenca del Pacífico, en contraste con la cuenca del Patía en el sector centro del departamento, con valores inferiores a los 1000 mm por año (Gobernación del Cauca, 2020). Exceptuando a la cuenca del Pacífico, donde el régimen lluvioso predomina durante todo el año, el régimen bimodal de lluvias predomina en el resto del departamento, con una temporada seca marcada en el periodo junio-agosto, una temporada lluviosa principal en el periodo octubre-diciembre y dos temporadas secas en el periodo enero-febrero donde se reducen las lluvias, para posteriormente incrementar en el periodo marzo-mayo (Gobernación del Cauca, 2020).

La principal fuente de información climatológica en el departamento del Cauca es la base de datos del catálogo de estaciones IDEAM, que cuenta con un total de 167 estaciones activas que incluyen 14 estaciones climáticas ordinarias, 28 climáticas principales, 1 meteorológica especial, 12 pluviográficas, 111 pluviométricas y 1 sinóptica principal, resultando en un promedio de una estación por cada 175.5 km² (IDEAM, 2022a) (Ver Figura 2).

Figura 2. *Distribución Espacial Estaciones IDEAM en el Departamento del Cauca*

Nota. Imagen de elaboración propia, a partir del límite político del departamento del Cauca suministrado por el INS, la base de datos de las estaciones IDEAM (IDEAM, 2022a) y la información satelital de Google disponible en QGIS.

6. Metodología

La estrategia de escalamiento para la obtención de datos climatológicos en alta resolución para su implementación en estudios de salud pública se llevó a cabo en tres fases: i) Selección de una metodología estadística apropiada para la reducción de escala espacial de precipitación y temperatura para el caso de estudio; ii) Aplicación de la metodología estadística seleccionada para la reducción de escala espacial de las variables de interés generando un banco de datos climatológicos del departamento del Cauca, Colombia; iii) Determinación del grado de incertidumbre de los datos escalados por medio de métricas estadísticas. Mediante la ejecución de

cada una de las fases anteriormente mencionadas, se aseguró el cumplimiento de los objetivos de la propuesta de investigación.

6.1 Fase I - Selección de una Metodología Estadística Apropriada para la Reducción de Escala Espacial de Precipitación y Temperatura para el Caso de Estudio

6.1.1 Caracterización de las Metodologías Estadísticas de Reducción de Escala de Variables Climatológicas

Se realizó una búsqueda de metodologías para la reducción de escala mediante la revisión de literatura en las bases de datos ScienceDirect y Scopus, en una ventana de tiempo de 7 años debido a que esta presentaba el mayor porcentaje de publicaciones referentes al tema; en este proceso se hizo énfasis en publicaciones de revistas de primer y segundo nivel (Q1 y Q2 en la clasificación SJR). Las palabras claves empleadas en la búsqueda fueron “Downscaling”, “Weather”, “CHIRPS”, “Meteorological Variables”, “Temperature”, “Precipitation”, “MSWX”, “WORLDCLIM” y “Kriging”. Adicionalmente se usaron operadores lógicos como “AND” y “OR” que permitieron optimizar el proceso de consulta dentro de las bases de datos.

Posteriormente, se elaboró una tabla resumen de los artículos identificados, en la que se diligenció información relevante para el análisis como: título, autor, revista, año de publicación, método aplicado, descripción del método, escala obtenida, variables climatológicas, bases de datos escaladas, zona de aplicación, métricas de desempeño, número de citas, insumos requeridos, replicabilidad, nivel de complejidad y recursos informáticos.

6.1.2 Caracterización del Sitio de Estudio e Información Relevante para el Estudio.

La caracterización de la zona de estudio inició con la identificación de los componentes geográficos, morfológicos y climatológicos del departamento del Cauca. Posteriormente, se evaluó la disponibilidad y calidad de información climatológica *in-situ* y alternativa disponible para la

zona con el fin de determinar la compatibilidad de dicha información con las diferentes metodologías identificadas en los artículos consultados. Para la recopilación de las series de datos de las variables meteorológicas de interés se consultaron tres tipos de recursos; el primero fue la información disponible en el Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM) a través de la plataforma DHIME (IDEAM, 2022b); también se consultó la información de campo provista por las estaciones meteorológicas instaladas por el INS en los municipios de Patía, Miranda y Piamonte en el marco del proyecto que se encuentra desarrollando. Por último, se consultó la información de las bases de datos CHIRPS, ERA5-Land, y MSWX, debido a que estas cuentan con la resolución temporal requerida por el INS.

Tabla 1. *Bases de Datos Climatológicas Globales Relevantes*

BASE DE DATOS	CHIRPS	ERA5-Land	MSWX
Resolución Espacial	0.05° x 0.05°	0.1° x 0.1°	0.1° x 0.1°
Variables Disponibles	Precipitación	Precipitación y Temperatura	Precipitación y Temperatura
Resolución Temporal	Diaria	Horaria	Cada 3 horas
Método de Análisis	Observaciones infrarrojas de duración de nubes frías	Reanálisis de modelos con observaciones alrededor del mundo.	Se basa en datos ERA5 corregidos y reducidos utilizando climatologías de referencia de alta resolución.

6.1.3 Selección de la Metodología de Reducción de Escala a Implementar

Finalmente, a partir de la información recopilada mediante la revisión de literatura, la caracterización de la zona de estudio y la evaluación de la información climatológica disponible, se desarrolló una mesa de trabajo con los evaluadores del anteproyecto, mediante la cual se

evaluaron varios filtros que permitieron seleccionar la metodología a implementar, como se nombran a continuación: I) Se descartaron las metodologías basadas en información de estaciones locales, debido a la baja densidad de estaciones presentes en el departamento del Cauca; II) Se dio prioridad a aquellos métodos que permitían su aplicación con diversas bases de datos, con el objeto de poder comparar la variación de los resultados en las diferentes fuentes disponibles en la zona. Por último, se priorizó para la selección el método que permitía mayor acceso a la información de los scripts y detalles del procedimiento, lo cual facilitaba su replicabilidad.

6.2 Fase II - Aplicación de la Metodología Estadística Seleccionada para la Reducción de Escala Espacial de las Variables de Interés Generando un Banco de Datos Climatológicos del Departamento del Cauca, Colombia

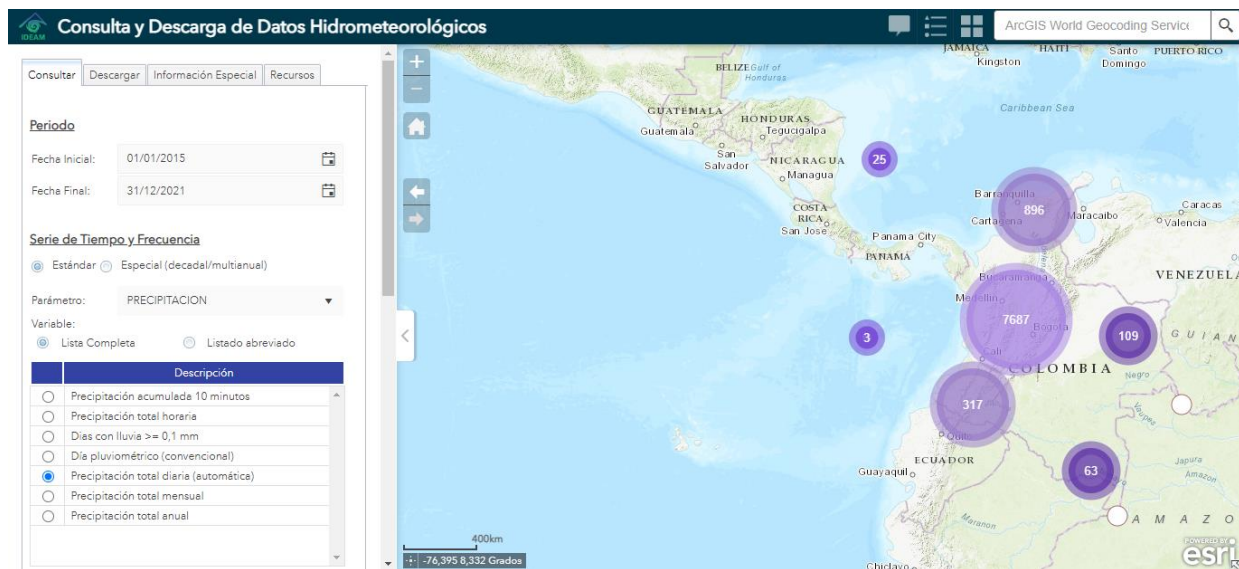
6.2.1 Procesamiento de la Información Obtenida de las Variables Meteorológicas de Interés

Descarga y Procesamiento de Información Estaciones IDEAM. Para obtener la información de las estaciones IDEAM, se utilizó el portal de consulta y descarga de datos hidrometeorológicos (Ver Figura 3); este proceso requirió la consulta y descarga de la información disponible en las estaciones locales del departamento del Cauca para cada una de las variables meteorológicas, dentro de la ventana de tiempo del 01/01/2015 al 31/12/2021 (Periodo del estudio desarrollado por el INS). Los datos, que inicialmente se descargaron del DHIME en formato “csv”, contenían los valores diarios de precipitación total, temperatura máxima, mínima y promedio, según la disponibilidad de cada estación. Esta información fue luego procesada usando Microsoft Excel, para identificar y eliminar los datos erróneos (valores que presentaban errores de digitación), y así, posteriormente, evaluar la continuidad temporal de las series de tiempo.

Debido a que los datos IDEAM se implementaron para estimar el rendimiento de las bases de datos originales y escaladas, no se realizó el proceso de llenado de datos faltantes, con el objeto

de no trasladar el error asociado a este procedimiento a las métricas estimadas para la valoración del presente estudio. Finalmente, se descartaron aquellas estaciones que tuvieron un porcentaje de datos faltantes superior al 50%, debido a la baja cantidad de estaciones disponibles para cada variable, y de aumentarse la exigencia en este criterio, se reduciría significativamente el número de estaciones disponibles para la evaluación de los resultados.

Figura 3. Portal web de consulta y descarga de datos Hidrometeorológicos del IDEAM (DHIME)



Nota: IDEAM. (2023). *Consulta y Descarga de Datos Hidrometeorológicos*. <http://dhime.ideam.gov.co/atencionciudadano/>

Descarga y Procesamiento de Información Estaciones INS. La información de las estaciones meteorológicas instaladas por el INS fue compartida por el equipo de trabajo del grupo de entomología de la dirección de investigación en salud pública, mediante una carpeta compartida en Google Drive. Los datos se entregaron en formato “xlsx”, lo cual permitió realizar el mismo proceso de validación de continuidad temporal de los datos realizado para las estaciones IDEAM.

Descarga y Procesamiento de Información CHIRPS. El proceso de descarga de los datos CHIRPS se realizó a través del portal web del National Oceanic and Atmospheric Administration

(NOAA) (Ver Figura 4). Esta plataforma permite la definición del área geográfica de interés, que para este caso fueron las que definen el recuadro que encierra al departamento del Cauca (0.9° a 3.4° latitud y -78° a -75.6° longitud). La ventana de tiempo de descarga fue la misma que se definió para los datos de las estaciones IDEAM y el formato de descarga del archivo fue tipo “nc”, que es un formato multidimensional que ofrece una serie de tiempo de precipitación para cada centro de pixel de la cuadrícula definida por las coordenadas de interés. Para este caso, los datos descargados no requirieron de algún procesamiento adicional previo al proceso de escalamiento.

Figura 4. Portal Web de Descarga de Datos CHIRPS

The screenshot shows the ERDDAP web interface. At the top, there's a blue header with the ERDDAP logo and text "Easier access to scientific data". On the right, there are links for "log in" and "English", and a note "Brought to you by NOAA NMFS SWFSC ERD". Below the header, the breadcrumb trail reads "ERDDAP > griddap > Data Access Form". The main content area displays the dataset title "CHIRPS Version 2.0, Precipitation, Global, 0.05°, Daily, 1981-present" with a link to the RSS feed. Below the title, there's information about the institution (UCSB Climate Hazards Group) and dataset ID (chirps20GlobalDailyP05). A section titled "Dimensions" allows users to select time, latitude, and longitude ranges. The "Start" and "Stop" time fields are set to "2022-08-31T00:00:00Z". The "Stride" for time is set to "1". The "Size" is "9743" and "Spacing" is "1 day 10h 47m 25s (uneven)". The "latitude (degrees_north)" range is from "-49.975" to "49.975" with a "Stride" of "1" and "Size" of "2000". The "longitude (degrees_east)" range is from "-179.975" to "179.975" with a "Stride" of "1" and "Size" of "7200". Below the dimensions, there's a "Grid Variables" section with a checkbox for "precip (Precipitation, mm/day)" which is checked. At the bottom, there's a "File type" dropdown menu set to ".htmlTable - View a UTF-8 .html web page with the data in a table. Times are ISO 8601 strings." and a "Submit" button with a note "(Please be patient. It may take a while to get the data.)".

Nota: Tomado de <https://coastwatch.pfeg.noaa.gov/erddap/griddap/chirps20GlobalDailyP05.html>

Descarga y Procesamiento de Información ERA5-Land. Para acceder a la información de la base de datos ERA5-Land se requirió la creación de una cuenta de usuario en la página web del Climate Data Store del Programa de Observación de la Tierra de la Unión Europea “Copernicus”; esto con el fin de obtener el ID y la clave para la descarga de la información. El proceso de descarga se realizó mediante la función “download_ERA”, de la librería “KrigR” de R, a través de la interfaz RStudio. Este código permite definir la variable requerida, la ventana de tiempo, la extensión

geográfica y la resolución temporal. El archivo descargado es de tipo “nc”. El Apéndice A presenta el script empleado para la descarga de cada una de las variables.

Descarga y Procesamiento de Información MSWX. Previo al proceso de descarga de la información se requirió realizar el proceso de aplicación a la licencia de los datos en la página web de GloH2O, institución encargada del desarrollo del producto MSWX. Posterior a esto, vía correo electrónico, se recibió el vínculo de acceso a la carpeta de Google Drive donde se encuentra toda la información disponible para descarga; en esta se ingresó al fichero Past, que contiene el registro histórico de las variables climatológicas. Allí se accedió la carpeta de cada una de las variables requeridas para el proyecto en la opción de registros diarios y se descendieron los datos en formato “nc” de cada uno de los días dentro de la ventana de tiempo definida en el proyecto. Como resultado del proceso anterior se organizaron carpetas con cada una de las variables descendidas las cuales contenían 2557 archivos cada uno con la información de la variable para todo el mundo. Esto requirió un procesamiento para permitir el escalamiento de la información: en primer lugar mediante el aplicativo CDO de la interfaz Linux ofrecida por el sistema operativo Ubuntu se realizó el corte de la información dentro de las coordenadas externas del departamento del Cauca mediante la línea de comando “`for i in $(ls input/*);do i=$(basename $i) ;cdo sellonlatbox,-78,-75.6,0.9,3.4 input/$i output/$i ; done`”; posteriormente se unieron todos los días para crear un único archivo con toda la información requerida para el proceso de escalamiento; este proceso se realizó mediante la línea de comando “`cdo mergetime *.nc Cauca_mswx_2015-2021.nc`”.

6.2.2 Implementación de la metodología seleccionada para la reducción de escala espacial

Método Kriging. El Kriging es un método geoestadístico, que permite la estimación de valores espaciales dentro de un área mediante el uso de medias, funciones de covarianza y variogramas de datos existentes dentro de esta. Existen tres formas principales del Kriging lineal:

Kriging simple, que parte del uso de estimadores lineales mediante una media conocida y la función de covarianza; Kriging ordinario, usado para el caso de no tener una media conocida, solo utilizando funciones de covarianza; finalmente, Kriging universal, que implementa el uso de variogramas para la construcción del estimador (Chilès & Delfiner, 2012). El método implementado en el desarrollo del proyecto parte del uso del sistema de ecuaciones del Kriging universal.

$$\sum_{\beta} \lambda_{\beta} \gamma_{\alpha\beta} + \sum_l \mu_l f_{\alpha}^l = \gamma_{\alpha 0} \quad \alpha = 1, \dots, N$$

$$\sum_{\alpha} \lambda_{\alpha} f_{\alpha}^l = f_0^l \quad l = 0, \dots, L$$

Donde, λ_{β} es el peso del conjunto de datos en el punto β , $\gamma_{\alpha\beta}$ es el variograma entre los conjuntos de datos de los puntos α y β , f_{α}^l son funciones de base conocidas del conjunto de datos en el punto α , μ_l son $L + 1$ incógnitas adicionales empleadas mediante el uso de los multiplicadores de Lagrange, $\gamma_{\alpha 0}$ es el variograma entre el conjunto de datos en el punto α y el conjunto de datos en el punto 0, λ_{α} es el peso del conjunto de datos α y f_0^l son funciones de base conocidas del conjunto de datos en el punto 0 (Chilès & Delfiner, 2012).

El variograma se puede definir matemáticamente como “la mitad de la esperanza matemática del cuadrado de la diferencia de dos valores de la variable de estudio Z , correspondiente a sitios separados en una distancia h ” (Oropeza et al., 2018).

$$\gamma(h) = \frac{1}{2} E[Z(X) - Z(X + h)]^2$$

Donde, $E[Z(x)]$ es el operador de esperanza matemática, X es un vector de coordenadas y h es la distancia que separa a dos puntos donde se mide la variable Z (Oropeza et al., 2018).

El método geoestadístico Kriging ha sido utilizado en el escalamiento de variables meteorológicas, lo que lo ha convertido en una alternativa para la obtención de información climatológica de alta resolución en diversos campos de estudio. Por ejemplo, se ha utilizado para obtener datos de temperatura del aire espacialmente continua para la evaluación del efecto integral del patrón de distribución espacial tridimensional de los edificios en el medio ambiente térmico urbano (Y. Chen et al., 2022); llenar vacíos espaciales de temperatura de la superficie de la carretera en ubicaciones no monitoreadas, para la construcción de cartografía espacial de las condiciones de la superficie de las carreteras en invierno (Wu et al., 2022); construir mapas de temperatura mínima para la evaluación de la posibilidad teórica de distribución de *Bulinus globosus* en China (X. Wang et al., 2022); y rellenar brechas espaciales de la temperatura diaria de la superficie del mar obtenida por el satélite multipropósito geoestacionario de Corea-2A (Ahn & Lee, 2022).

Esta metodología también ha sido implementada para la estimación de datos espaciales de variables no climatológicas. Por ejemplo, se ha implementado en la interpolación de datos dispersos de aflatoxinas para la creación de mapas de riesgo (Kerry et al., 2023); ha sido probado para la estimación de la distribución espacial de las existencias de carbono orgánico del suelo en los Vertisoles del estado de Gezira, Sudán, con el uso de covariables ambientales y datos heredados del suelo (Ahmed et al., 2022); se ha utilizado para la construcción de mapas de distribución geográfica de los valores de referencia de 25-hidroxivitamina D para personas mayores sanas en China (Yang et al., 2022); y la construcción de la distribución espacial de las reservas de carbono del suelo, nitrógeno total e índices microbianos tras un incendio forestal en el pinar rojo Mediterráneo (Babur et al., 2022).

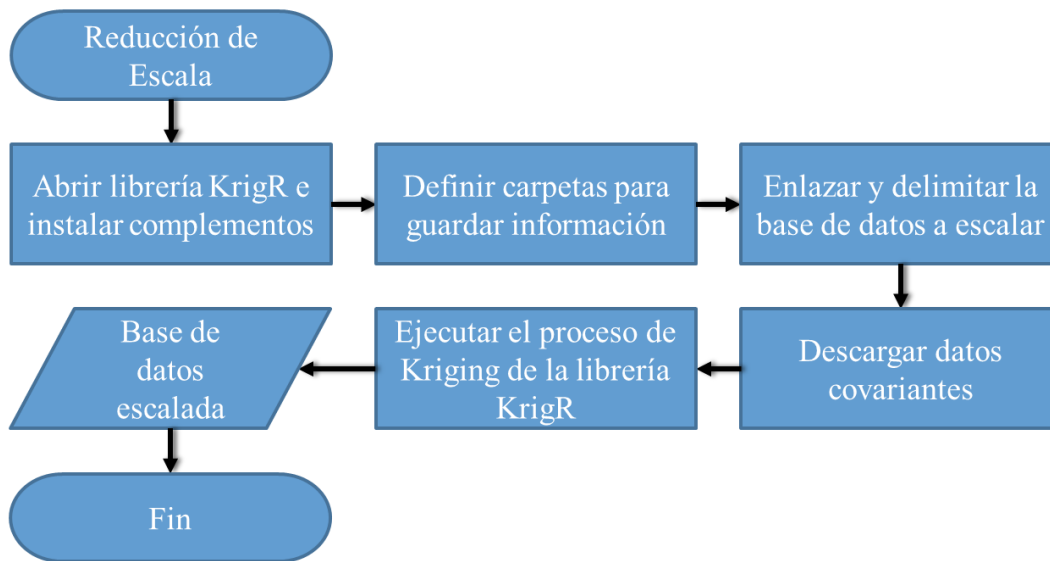
Para la implementación del método geoestadístico Kriging en la reducción de escala de las variables climatológicas de interés del proyecto se utilizó la librería KrigR de R, mediante la función krigR (multi-core Kriging), que requiere para su implementación datos de entrenamiento que se desean escalar y datos de covariables en dos resoluciones: la resolución de los datos de entrenamiento y la resolución a la cual se quiere llegar con el proceso de escalamiento. A partir de esta información, se efectúa el proceso de escalamiento mediante dos pasos; primero, la herramienta ajusta variogramas a los datos de entrenamiento y establece funciones de covarianza con los datos de las covariables en la resolución de los datos de entrenamiento, esto genera funciones que permiten describir la correlación espacial entre los datos de entrenamiento y las covariables (Kusch & Davy, 2022); posteriormente, se realiza la predicción de los valores de la variable en las nuevas ubicaciones, a partir de los datos de las covariables en la resolución objetivo (Davy & Kusch, 2021).

En el proceso de escalamiento de las bases de datos ERA5-Land y MSWX, para las variables temperatura máxima, mínima y promedio diaria, se utilizó como información covariante el modelo digital de elevación (DEM) (Resolución de 0.01°) del Servicio Geológico de Estados Unidos (USGS) (USGS.Gov | Science for a Changing World, 2022), ya que este ha sido validado para su uso en el estudio desarrollado por Kusch & Davy (2022).

Por otra parte, debido a la recomendación de Kusch & Davy (2022) de no solo utilizar el DEM para la variable precipitación total diaria, se planteó como covariante adicional el índice de vegetación de diferencia normalizada (NDVI) del estudio de mapeo y modelado de inventario global (GIMMS) (GIMMS | Global Agricultural Monitoring System, 2022). Esta decisión se soportó en la propuesta de López et al., (2018), que utilizó el NDVI en el escalamiento de la variable precipitación en la cuenca del río Magdalena.

La Figura 5 presenta un diagrama del proceso empleado para el escalamiento de los datos meteorológicos. Los scripts empleados para la reducción de escala se pueden visualizar en el Apéndice A.

Figura 5. Diagrama de Flujo Reducción de Escala



Nota: Los scripts empleados para la reducción de escala se pueden visualizar en el apéndice A.

6.3 Fase III Determinación del Grado de Incertidumbre de los Datos Escalados por Medio de Métricas Estadísticas.

La definición de las métricas de desempeño a utilizar para el proceso de evaluación de los datos obtenidos se basó en la revisión de la literatura realizada en la fase I, donde se identificaron las más utilizadas. A continuación, se detalla cada una de las métricas evaluadas en el proyecto.

Root Mean Square Error (RMSE): El RMSE es el parámetro de validación cruzada más utilizado. Los valores más bajos (tendientes a cero) indican un rendimiento más alto (Njoku et al., 2023).

$$RMSE = \sqrt{\left(\frac{1}{N}\right) \sum_{i=1}^N [P_i - Q_i]^2}$$

Mean Absolute Error (MAE): El MAE calcula el promedio de la diferencia absoluta entre el valor medido y el estimado. El mejor ajuste se evidencia con valores cercanos a cero (Njoku et al., 2023).

$$MAE = \frac{\sum_{i=1}^N |P_i - Q_i|}{N}$$

Pearson Correlation Coefficient (r): El r se utiliza para medir la correlación entre la variable observada y la estimada; este coeficiente varía entre -1 y 1. Si $r = 0$ no existe una relación lineal entre las variables; si r es 1 o -1, existe relación lineal positiva o negativa perfecta (Moriasi et al., 2007).

$$r = \frac{\left(N \sum_{i=1}^N (P_i * Q_i) - (\sum_{i=1}^N P_i)(\sum_{i=1}^N Q_i) \right)}{\sqrt{[N \sum_{i=1}^N P_i^2 - (\sum_{i=1}^N P_i)^2][N \sum_{i=1}^N Q_i^2 - (\sum_{i=1}^N Q_i)^2]}}$$

Mean Bias Error (MBE) representa el promedio aritmético de todos los errores estimados. Los valores positivos representan sobreestimación de los datos, mientras que la subestimación se representa por valores negativos (Njoku et al., 2023).

$$MBE = \frac{1}{N} \sum_{i=1}^N [P_i - Q_i]$$

En todos los casos Q_i representa los valores observados o medidos de cada variable meteorológica (precipitación total, temperatura máxima, mínima o promedio diaria) en la estación in-situ; P_i representa los valores estimados de la variable meteorológica por las bases de datos originales o escalados (CHIRPS, ERA5-Land o MSWX); y N es el número de datos evaluados. Se usó una hoja de cálculo en Microsoft Excel para la estimación de cada una de las métricas definidas. Este procedimiento se hizo tanto para los datos crudos de las bases de datos satelitales, como para los valores obtenidos en el proceso de escalamiento.

7. Resultados

7.1 Metodología Estadística Seleccionada para la Reducción de Escala Espacial

En el proceso de revisión de literatura se identificaron 23 artículos que describían metodologías para la reducción de escala. Posteriormente, se descartaron artículos que presentaban metodologías similares, priorizando el estudio con mayor número de citas. Producto de lo anterior, para la mesa de trabajo desarrollada en conjunto con los evaluadores y directores, se detallaron 17 estudios de escalamiento para datos de precipitación y/o temperatura identificando en cada uno de ellos los criterios definidos en el numeral 6.1.1 (ver Tabla 2).

Tabla 2. *Metodologías de Reducción de Escala Espacial*

Titulo	Autores	Descripción del método	Escala final	Variables Climatológicas	Base de datos escalados	Métricas
A tool for downscaling weather data from large-grid reanalysis products to finer spatial scales for distributed hydrological applications	Sen Gupta & Tarboton, 2016	Ofrece un algoritmo desarrollado para la obtención de datos climatológicos con alta resolución espacial.	100 m	Temperatura Radiación de onda corta Velocidad del viento Humedad relativa Precipitación	MERRA NOAA's Rainfall Estimation	NSE RMSE BIAS

Titulo	Autores	Descripción del método	Escala final	Variables Climatológicas	Base de datos escalados	Métricas
Comparison and assessment of spatial downscaling methods for enhancing the accuracy of satellite-based precipitation over Lake Urmia Basin	Karbalaye Ghorbanpo ur et al., 2021	Compara 5 metodologías estadísticas de reducción de escala (Support vector machine, random forest, geographically weighted regression, multiple linear regression y exponential regression).	1 km	Precipitación	Tropical Rainfall Measuring Mission (TRMM)	RMSE MAE BIAS
Spatial downscaling of radar-derived rainfall field by two-dimensional wavelet transform	Nourani et al., 2020	El método propone la reducción de escala a partir del uso del concepto de fluctuación de lluvia estandarizada, utilizando una transformada wavelet discreta bidimensional	1 km	Precipitación	NEXRAD	CC

Titulo	Autores	Descripción del método	Escala final	Variables Climatológicas	Base de datos escalados	Métricas
A downscaling-merging method for high-resolution daily precipitation estimation	Fengrui et al., 2020a	Este método se basa en cuatro pasos: I) Ampliar de escala temporal; II) Reducir la escala espacial; III) Realizar la reducción de escala temporal; IV) Integrar de los datos escalados con los datos observados en los pluviómetros.	1 km	Precipitación diaria	Tropical Rainfall Measuring Mission	r MAE RMSE BIAS POD FAR CSI
Daily gridded meteorological variables in Brazil (1980–2013)	Xavier et al., 2016	* Promedio Aritmético * Thin plate spline * Natural neighbor * Inverse distance weighting * Angular distance weighting * Ordinary point kriging	25 km	Temperatura (max, min) Radiación solar Humedad Relativa Velocidad del viento Precipitación Evapotranspiración	Estaciones meteorológicas de Brasil	R BIAS RMSE MAE CRE CSI PC

Titulo	Autores	Descripción del método	Escala final	Variables Climatológicas	Base de datos escalados	Métricas
CMIP5 downscaling and its uncertainty in China	Yue et al., 2016	Plantea una variación al método de modelado de superficie de alta resolución, mediante el uso de salida de residuos minimizados del método HASM con una ecuación lineal de mínimos cuadrados ordinarios que emplea latitud y elevación como variables independientes	1 km	Temperatura Precipitación	CMIP5	MAE MRE
Statistical downscaling of precipitation using local regression and high accuracy surface modeling method	Zhao et al., 2017	Plantea una combinación de modelo de regresión ponderada y el método de modelado de superficie de alta precisión	1 km	Precipitación	CMIP5	RMSE MRE MAE

Titulo	Autores	Descripción del método	Escala final	Variables Climatológicas	Base de datos escalados	Métricas
A spatial downscaling methodology for grace total water storage anomalies using gpm imerg precipitation estimates	Gemitzi et al., 2021	Desarrollo de un algoritmo para la reducción espacial de las lecturas de anomalía en el almacenamiento total de agua la cual está a una resolución de 1° a partir de la relación que esta tiene con información de precipitación disponible a 0.1°	10 km	Anomalías en el almacenamiento total de agua	GRACE	MAE R
Spatial downscaling of 3-hourly precipitation forecast data at river basin scale	Mendes & Maia, 2020	Plantea la disgregación de la información disponible en ERA-Interim y la interpolación de datos observados, posteriormente realiza un ajuste estadístico mediante la corrección BIAS	12.5 km	Precipitación	ERA-Interim	MTE MBE MAE RMSE

Titulo	Autores	Descripción del método	Escala final	Variables Climatológicas	Base de datos escalados	Métricas
Extreme value metastatistical analysis of remotely sensed rainfall in ungauged areas: Spatial downscaling and error modelling.	Zorzetto & Marani, 2020	Emplearon un método es escalamiento estadístico basado en la teoría de campos aleatorios.	No especif ica	Precipitación	TRMM	No especifica
Spatial Downscaling of Satellite-Based Precipitation and Its Impact on Discharge Simulations in the Magdalena River Basin in Colombia	López López et al., 2018, p.	Plantea el proceso mediante tres pasos: 1) Reducción espacial de la precipitación mensual 2) Fusión de la precipitación mensual reducida con datos observados in situ y 3) Desglose de la precipitación mensual en valores diarios	1 km	Precipitación	CMORPH PERSIANN TRMM MSWEP	CD RMSE BIAS

Titulo	Autores	Descripción del método	Escala final	Variables Climatológicas	Base de datos escalados	Métricas
Spatial Downscaling of Monthly TRMM Precipitation Based on EVI and Other Geospatial Variables over the Tibetan Plateau from 2001 to 2012	Shi & Song, 2015	El método emplea relaciones estadísticas no paramétricas entre la precipitación, índice de vegetación mejorado, la altitud, la pendiente, latitud y longitud.	1 km	Precipitación	TRMM	R2 RMSE MAE BIAS
A statistical spatial downscaling algorithm of TRMM precipitation based on NDVI and DEM in the Qaidam Basin of China	Jia et al., 2011	Método estadístico que emplea la relación entre la precipitación y factores ambientales como la topografía y la vegetación	1 km	Precipitación	TRMM	r2

Titulo	Autores	Descripción del método	Escala final	Variables Climatológicas	Base de datos escalados	Métricas
Spatial downscaling of TRMM precipitation using vegetative response on the Iberian Peninsula	Immerzeel et al., 2009	Método basado en la relación existente entre la precipitación y el índice de vegetación de diferencia normalizada	1 km	Precipitación	TRMM	RMSE
Spatial variation of materials surface temperature by regression based downscaling model's in Jaipur district, India	Vaibhav Sharma et al., 2020	Resolución de reducción de escala de regresión ajustada Regresión de Kriging Árbol de regresión	100 m	Temperatura	MODIS	RMSE
A new spatial modeling and interpolation approach for high-resolution temperature maps combining reanalysis data and ground measurements	Viggiano et al., 2019	Método de Kriging	2 km	Temperatura de aire máxima y mínima diaria	ERA-Interim	RMSE MAE

Titulo	Autores	Descripción del método	Escala final	Variables Climatológicas	Base de datos escalados	Métricas
KrigR—a tool for downloading and statistically downscaling climate reanalysis data	Kusch & Davy, 2022	Método de Kriging	1 km	Temperatura	ERA5-Land	Standard Error

Nota: La tabla completa se puede visualizar en el apéndice B.

Posterior al proceso de búsqueda de metodologías para la reducción de escala, se realizó el levantamiento de información de la zona de estudio, la cual se detalla en el capítulo 5, este proceso permitió conocer la información meteorológica disponible en la zona para el proceso de escalamiento.

Finalmente, se desarrolló la mesa de trabajo con los evaluadores del proyecto, mediante la cual se seleccionó el método geoestadístico Kriging, implementando el uso de la librería KrigR (Kusch & Davy, 2022), del software RStudio, debido a que este presentaba un alto grado de replicabilidad, un nivel de complejidad que permitía su aprendizaje dentro del tiempo establecido en el cronograma, y recursos informáticos que se ajustaban a los disponibles para el proyecto. Adicionalmente, este método no limita su uso a una base de datos de información meteorológica específica, lo cual posibilita el uso de las diferentes fuentes de datos disponibles en la zona de estudio.

7.2 Banco de Datos Climatológicos Escalados del Departamento del Cauca, Colombia

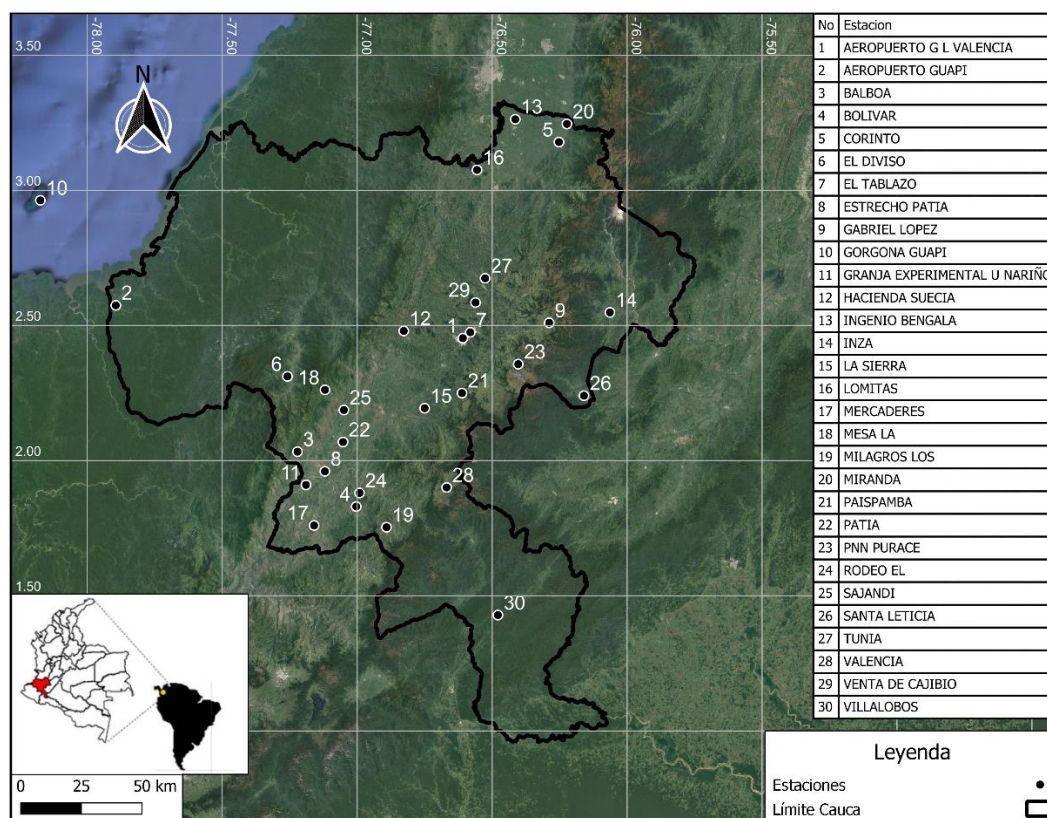
Se realizó la descarga y procesamiento de los datos climatológicos (Precipitación total, temperatura máxima, mínima y promedio diaria) de cada una de las fuentes de información

detalladas en la metodología, dentro de la ventana de tiempo requerida por el INS desde el 01/01/2015 hasta el 31/12/2021.

7.2.1 Base de Datos Cruda de las Variables Meteorológicas de Interés

Información Estaciones IDEAM. En el proceso de descarga de información de las estaciones IDEAM, se identificaron 30 estaciones con datos disponibles dentro del periodo de tiempo del proyecto (2015–2021), de las cuales 16 contaban con datos de precipitación, 18 tenían información de temperatura máxima y mínima y 1 estación contaba con datos de temperatura promedio. A continuación, en la Figura 6 y en la Tabla 3 se presenta la ubicación de las estaciones y el resumen del análisis preliminar de los datos, respectivamente.

Figura 6. Estaciones IDEAM con Datos Disponibles



Nota. Imagen de elaboración propia, a partir las coordenadas de las estaciones tomadas del IDEAM (IDEAM, 2022b) y el límite político del departamento del Cauca suministrado por el INS.

Tabla 3. Datos Estaciones IDEAM

No	Estación	Latitud	Longitud	P	T max	T min	T prom
				% faltante			
1	AEROPUERTO G L VALENCIA [26035030]	2.4529	-76.6088	*	21.74%	21.90%	*
2	AEROPUERTO GUAPI - AUT [53045040]	2.5744	-77.8948	52.01%	20.88%	19.28%	*
3	BALBOA - AUT [52015050]	2.0328	-77.2217	39.19%	*	*	*
4	BOLIVAR [52025010]	2.4529	-76.6088	*	17.25%	43.76%	*
5	CORINTO - AUT [2604500039]	2.5744	-77.8948	54.67%	51.11%	51.11%	51.51%
6	EL DIVISO - AUT [53075020]	2.0328	-77.2217	50.37%	*	*	*
7	EL TABLAZO - AUT [26015010]	1.8299	-77.0040	28.16%	*	*	*
8	ESTRECHO PATIA - AUT [52025080]	3.1786	-76.2528	26.98%	*	*	*
9	GABRIEL LOPEZ [26025070]	2.3114	-77.2588	*	42.12%	42.71%	*
10	GORGONA GUAPI [57025020]	2.4748	-76.5813	29.96%	54.60%	69.69%	*
11	GRANJA EXPERIMENTAL U NARIÑO [52025060]	1.9600	-77.1200	*	72.59%	80.95%	*
12	HACIENDA SUECIA - AUT [26035100]	2.5098	-76.2891	98.28%	*	*	*
13	INGENIO BENGALA [26045010]	2.9629	-78.1744	*	2.74%	2.46%	*
14	INZA - AUT [21055070]	1.9100	-77.1900	0.00%	*	*	*
15	LA SIERRA - AUT [52025090]	2.4796	-76.8278	49.98%	12.83%	11.89%	*
16	LOMITAS [26025110]	3.2628	-76.4145	*	0.35%	0.27%	*
17	MERCADERES [52025030]	2.5482	-76.0639	*	91.98%	91.40%	*
18	MESA LA [52010160]	2.1938	-76.7503	22.80%	*	*	*
19	MILAGROS LOS [52025020]	3.0754	-76.5563	*	9.97%	6.10%	*
20	MIRANDA [26065020]	1.7600	-77.1600	*	6.18%	21.55%	*
21	PAISPAMBA [52015010]	2.2616	-77.1196	*	0.35%	1.53%	*
22	PATIA [52010180]	1.7533	-76.8917	6.92%	*	*	*
23	PNN PURACE - AUT [26015030]	3.2457	-76.2231	33.44%	*	*	*

No	Estación	Latitud	Longitud	P	T max	T min	T prom
				% faltante			
24	RODEO EL [52020050]	2.2497	-76.6119	6.37%	*	*	*
25	SAJANDI [52010080]	2.0686	-77.0529	38.91%	*	*	*
26	SANTA LETICIA [21055030]	2.3568	-76.4035	*	50.45%	18.65%	*
27	TUNIA [26025100]	1.8802	-76.9916	*	8.06%	7.94%	*
28	VALENCIA [44015030]	2.1871	-77.0497	*	0.27%	0.59%	*
29	VENTA DE CAJIBIO [26025090]	2.2400	-76.1600	*	2.82%	3.56%	*
30	VILLALOBOS [44010130]	2.6745	-76.5263	7.39%	*	*	*

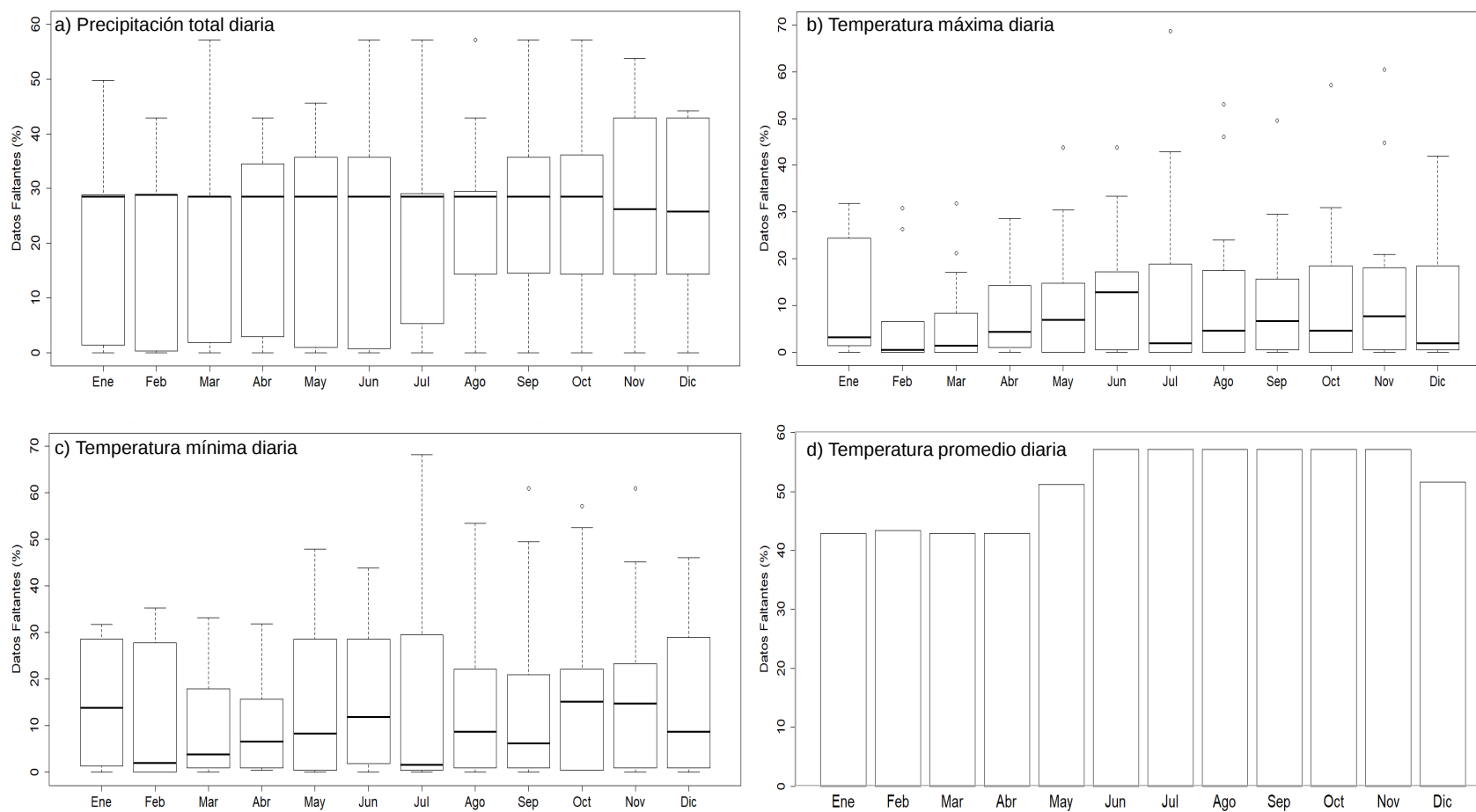
Nota. La tabla completa se puede visualizar en el apéndice C.

Tras el procesamiento preliminar de los datos de las estaciones IDEAM, se descartaron 4 estaciones con datos de precipitación, 5 estaciones con datos de temperatura máxima y 4 estaciones con datos de temperatura mínima, debido a que presentaban un porcentaje de datos faltantes superior al 50% dentro del periodo de estudio. Para el caso de la estación que contenía información de temperatura promedio, a pesar de contar con un porcentaje superior al 50%, se decidió conservar para poder valorar los resultados de esta variable, debido a la ausencia de más estaciones. Adicionalmente, se identifica que existe una baja densidad espacial de estaciones (Figura 6), y que gran parte de las existentes no cuentan con una continuidad estable de datos (Tabla 3). Las estaciones seleccionadas fueron empleadas para validar estadísticamente, usando las métricas de desempeño descritas en la sección 6.3, los resultados obtenidos en el proceso de escalamiento de bases de datos ERA5-Land, CHIRPS y MSWX.

Después del proceso de filtrado, se contó con un total de 12 estaciones con datos de precipitación, 13 estaciones con datos de temperatura máxima, 14 estaciones con datos de temperatura mínima y 1 estación con datos de temperatura promedio. Para el caso de las variables precipitación total, temperatura máxima y mínima diaria la mayor parte de las estaciones presentan

un porcentaje de datos faltantes inferior al 30% de las diferentes variables estudiadas (Ver Figura 7). Por otra parte, para el caso de la estación de temperatura promedio diaria, todos los meses presentan un porcentaje de datos faltantes superior al 40% y superior al 50% entre los meses de mayo a diciembre (Ver Figura 7).

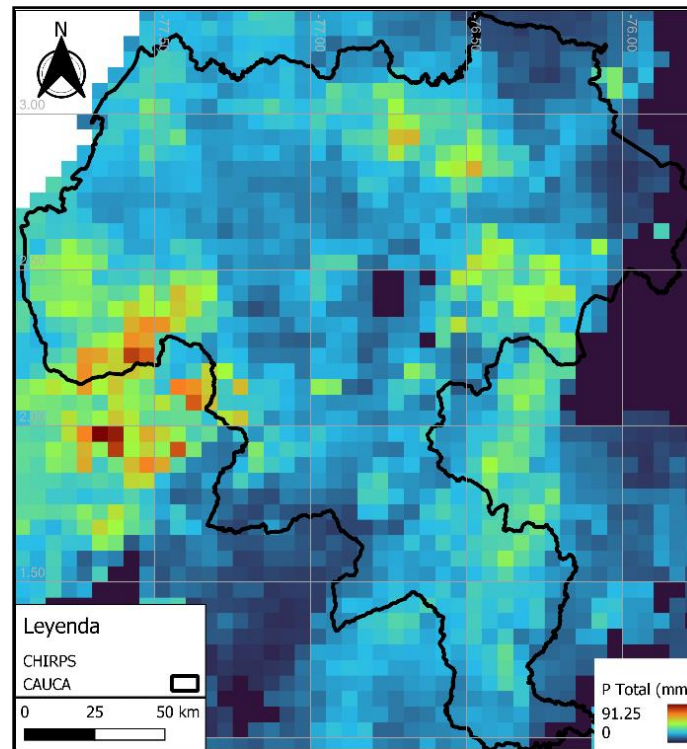
Información Estaciones INS. El INS suministró la información de las estaciones meteorológicas Davis Vantage instaladas en los municipios de Patía, Miranda y Piamonte, las cuales fueron descartadas debido a que solo contaba con un 4% de datos dentro del periodo de estudio (Ver apéndice C).

Figura 7. *Porcentaje de Datos Faltantes Estaciones IDEAM Seleccionadas*

Nota. Porcentaje de datos faltantes estaciones IDEAM seleccionadas para las variables (a) Precipitación Total Diaria (b) Temperatura Máxima Diaria (c) Temperatura Mínima Diaria (d) Temperatura Promedio Diaria.

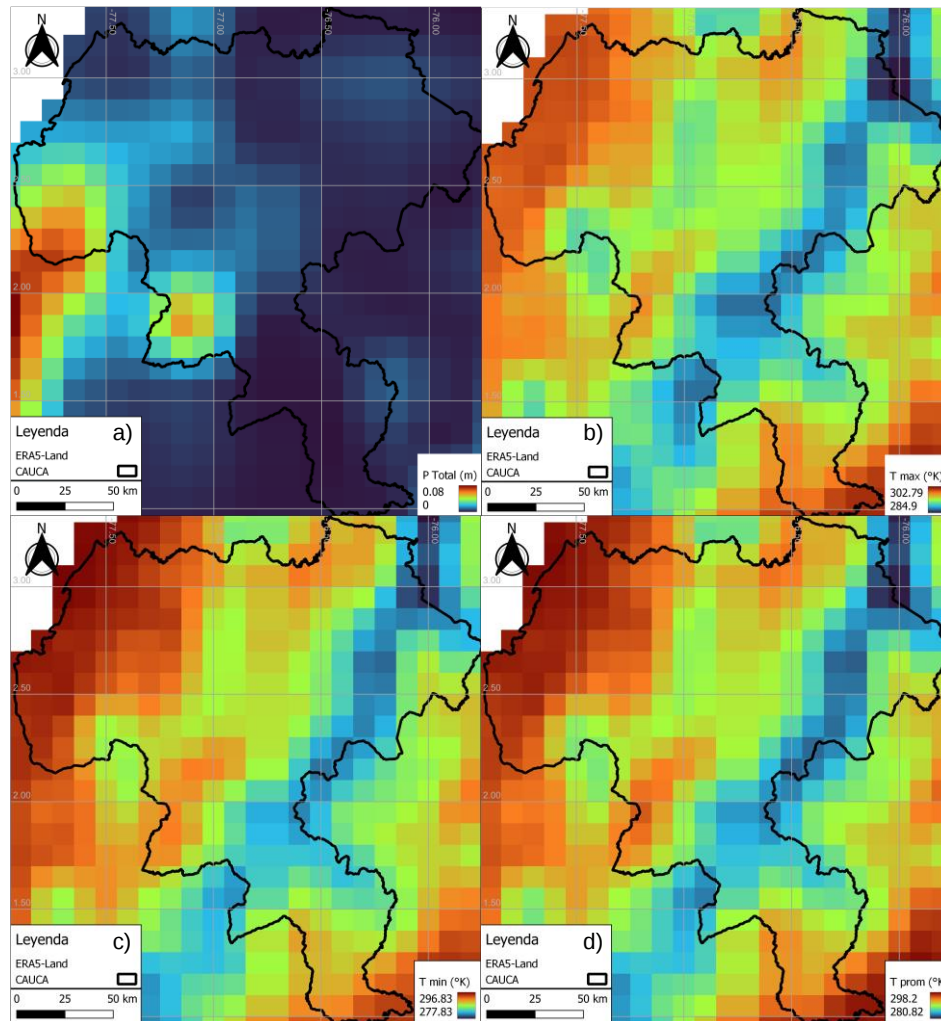
Información CHIRPS. En la Figura 8, se presenta un ejemplo de la información obtenida mediante la base de datos CHIRPS; se logró la descarga de datos de precipitación total diaria con una continuidad del 100% del periodo de estudio y una resolución espacial de 0.05° (5 km). La información completa de esta base de datos se encuentra en el apéndice D.

Figura 8. *Precipitación Total Diaria CHIRPS*



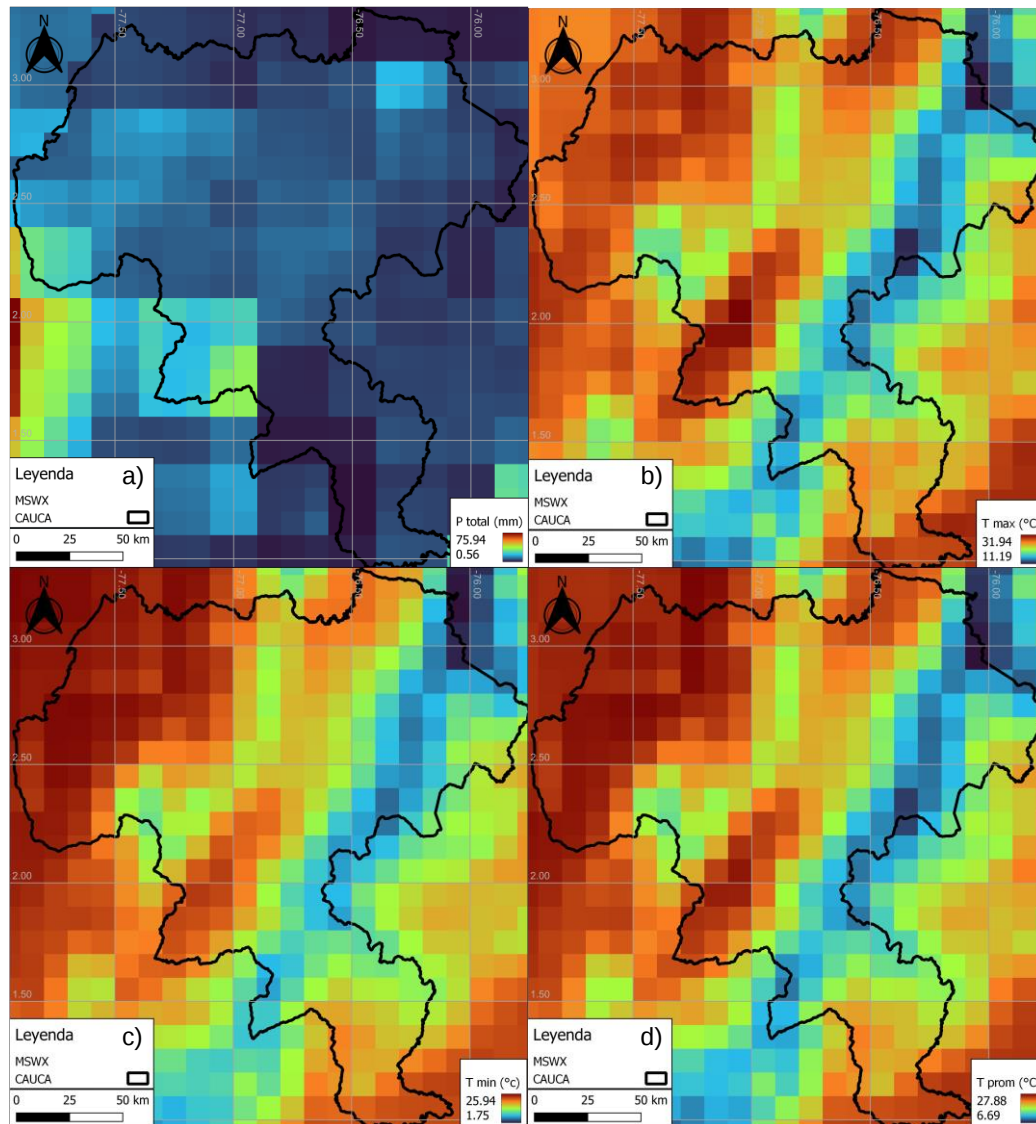
Nota. Ejemplo de datos de precipitación total diaria (mm) - 17/01/2015, resolución espacial 0.05° (5 km), fuente CHIRPS.

Información ERA5-Land. En la Figura 9 se presenta una muestra de la información obtenida mediante la base de datos ERA5-Land, en esta fuente de información se obtuvieron datos de las variables precipitación total, temperatura máxima, mínima y promedio diaria, para toda la ventana de tiempo del proyecto y una resolución de 0.1° (10 km). La información completa de esta base de datos se puede encuentra en el apéndice E.

Figura 9. Variables Disponibles ERA5-Land

Nota. Ejemplo de datos de (a) precipitación total diaria (m); (b) temperatura máxima (°K); (c) temperatura mínima (°K); (d) temperatura promedio (°K) - 17/01/2015, resolución espacial 0.1° (10 km), fuente ERA5-Land.

Información MSWX. En la Figura 10 se presenta una muestra de la información obtenida en la base de datos MSWX, la cual, permitió descargar datos de las variables precipitación total, temperatura máxima, mínima y promedio diaria, para el total de la ventana de tiempo del proyecto, la resolución espacial disponible en esta base de datos es de 0.1° (10 km). La información completa de esta base de datos se encuentra en el apéndice F.

Figura 10. Variables Disponibles MSWX

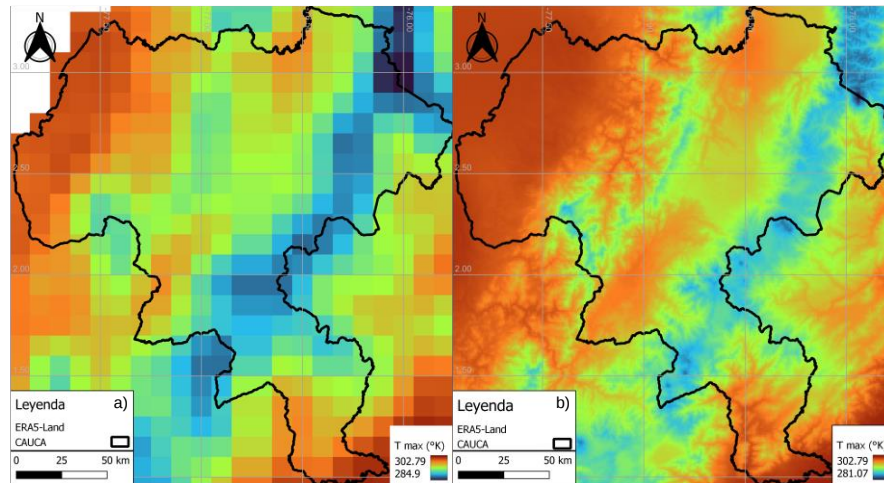
Nota. Ejemplo de datos de (a) precipitación total diaria (mm); (b) temperatura máxima (°C); (c) temperatura mínima (°C); (d) temperatura promedio (°C) - 17/01/2015, resolución espacial 0.1° (10 km), fuente MSWX.

7.2.2 Base de Datos Escalada de las Variables Meteorológicas de Interés

Temperatura Máxima. En las Figuras 11 y 12 se presenta un ejemplo de los resultados obtenidos en el proceso de escalamiento de la variable temperatura máxima diaria, utilizando el DEM como covariante, lo que permitió pasar de la resolución de 0.1° (10 km) de las bases de datos

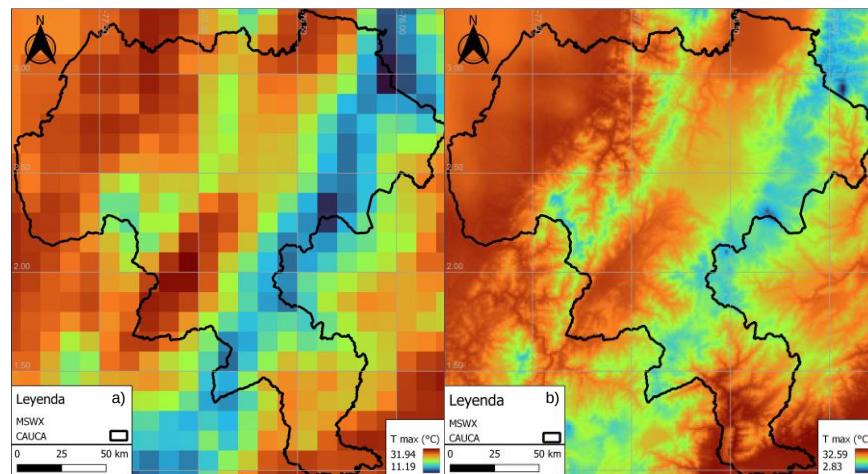
MSWX y ERA5-Land a 0.01° (1 km). La información completa de esta base de datos escaladas se encuentra en el apéndice G.

Figura 11. *Temperatura Máxima Diaria Escalada ERA5-Land*



Nota. Ejemplo de datos de temperatura máxima diaria ($^\circ\text{K}$) (a) Sin escalar (0.1°); (b) Escalados (0.01°) - 17/01/2015, base de datos ERA5-Land.

Figura 12. *Temperatura Máxima Diaria Escalada MSWX*

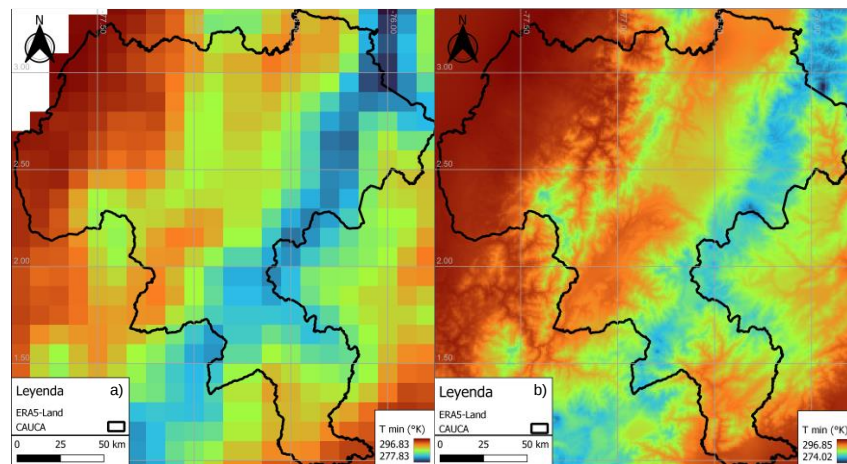


Nota. Ejemplo de datos de temperatura máxima diaria ($^\circ\text{C}$) (a) Sin escalar (0.1°); (b) Escalados (0.01°) - 17/01/2015, base de datos MSWX.

Temperatura Mínima. En las Figuras 13 y 14 se presenta un ejemplo de los resultados obtenidos en el proceso de escalamiento de la variable temperatura mínima diaria, para esto se

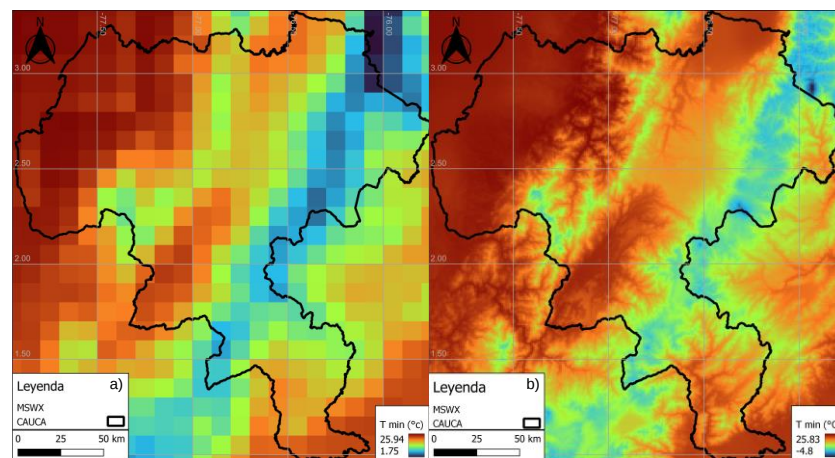
empleó como covariante el DEM, con este procedimiento se logró pasar de un pixel de 0.1° (10 km) de las bases de datos MSWX y ERA5-Land a 0.01° (1 km). La información completa de esta base de datos se encuentra en el apéndice H.

Figura 13. *Temperatura Mínima Diaria Escalada ERA5-Land*



Nota. Ejemplo de datos de temperatura mínima diaria ($^\circ\text{K}$) (a) Sin escalar (0.1°); (b) Escalados (0.01°) - 17/01/2015, base de datos ERA5-Land.

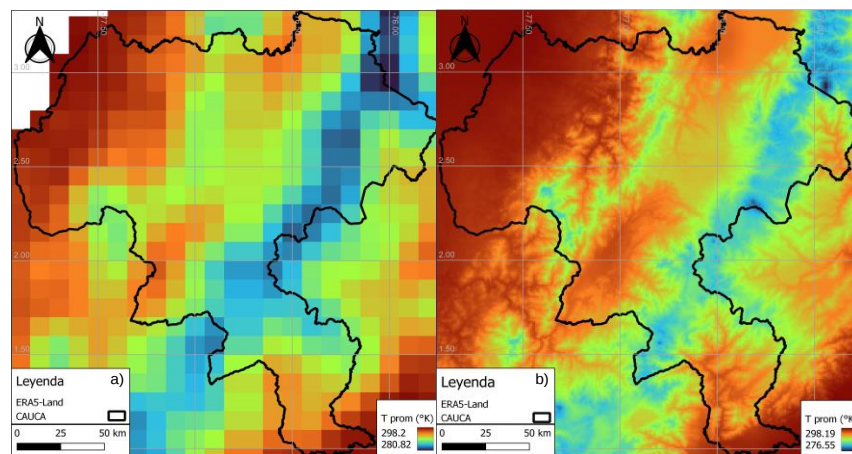
Figura 14. *Temperatura Mínima Diaria Escalada MSWX*



Nota. Ejemplo de datos de temperatura mínima diaria ($^\circ\text{C}$) (a) Sin escalar (0.1°); (b) Escalados (0.01°) - 17/01/2015, base de datos MSWX.

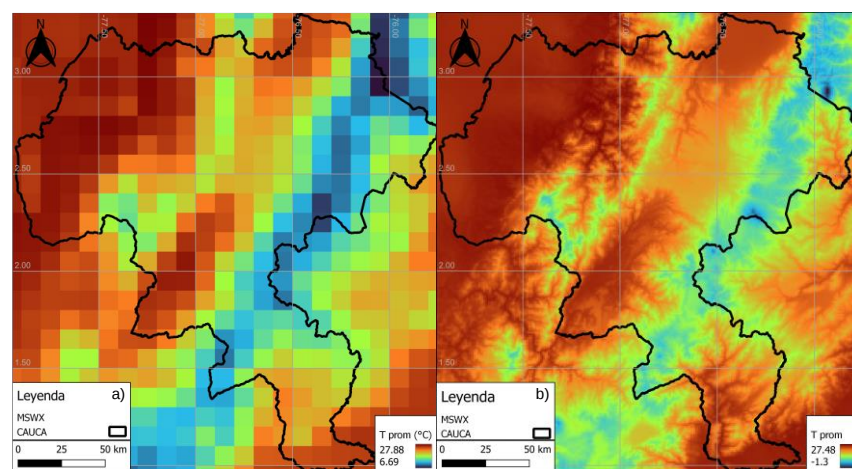
Temperatura Promedio. En las figuras 15 y 16 se presenta un ejemplo de los resultados obtenidos en el proceso de escalamiento de la variable temperatura promedio diaria, empleando como covariante el DEM, se logró pasar de la resolución de 0.1° (10 km) de las bases de datos MSWX y ERA5-Land a 0.01° (1 km). La información completa de esta base de datos se encuentra en el apéndice I.

Figura 15. *Temperatura Promedio Diaria Escalada ERA5-Land*



Nota. Ejemplo de datos de temperatura promedio diaria ($^\circ\text{K}$) (a) Sin escalar (0.1°); (b) Escalados (0.01°) - 17/01/2015, base de datos ERA5-Land.

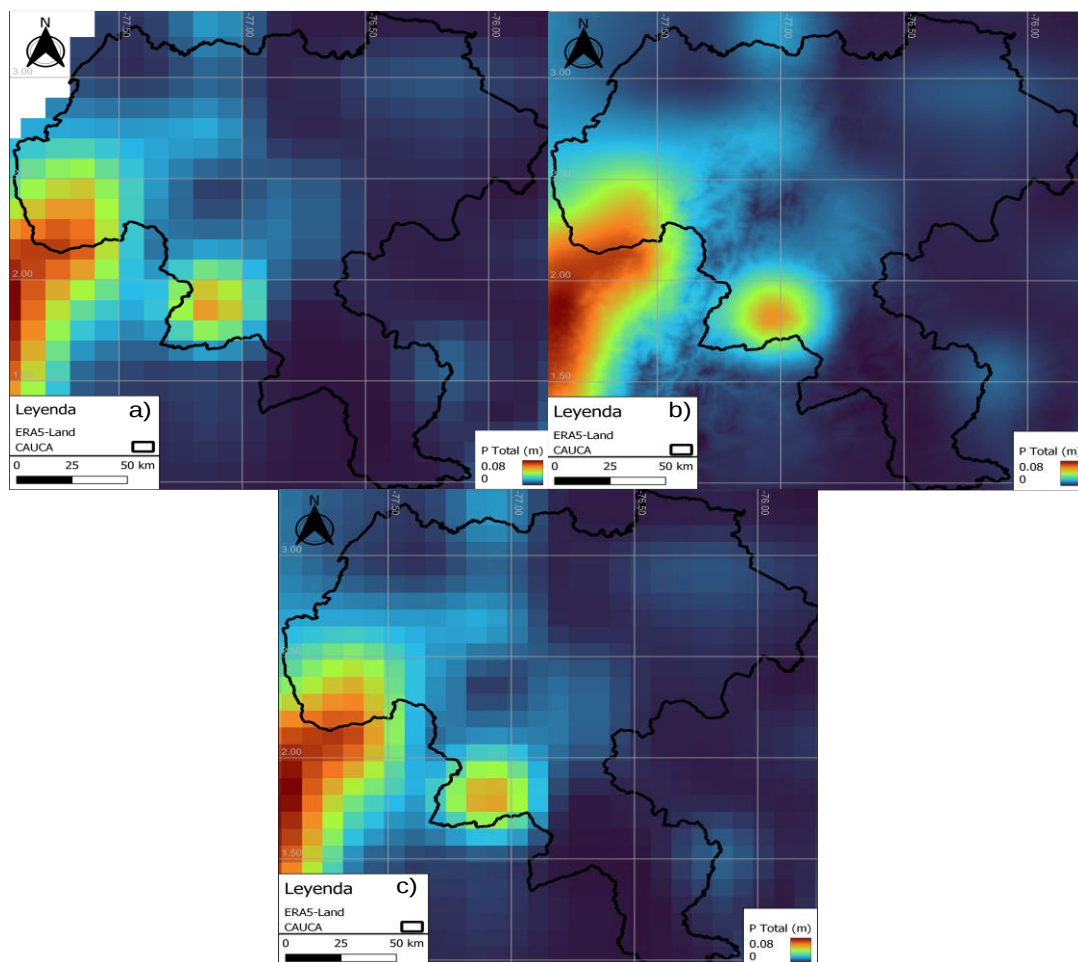
Figura 16. *Temperatura Promedio Diaria Escalada MSWX*



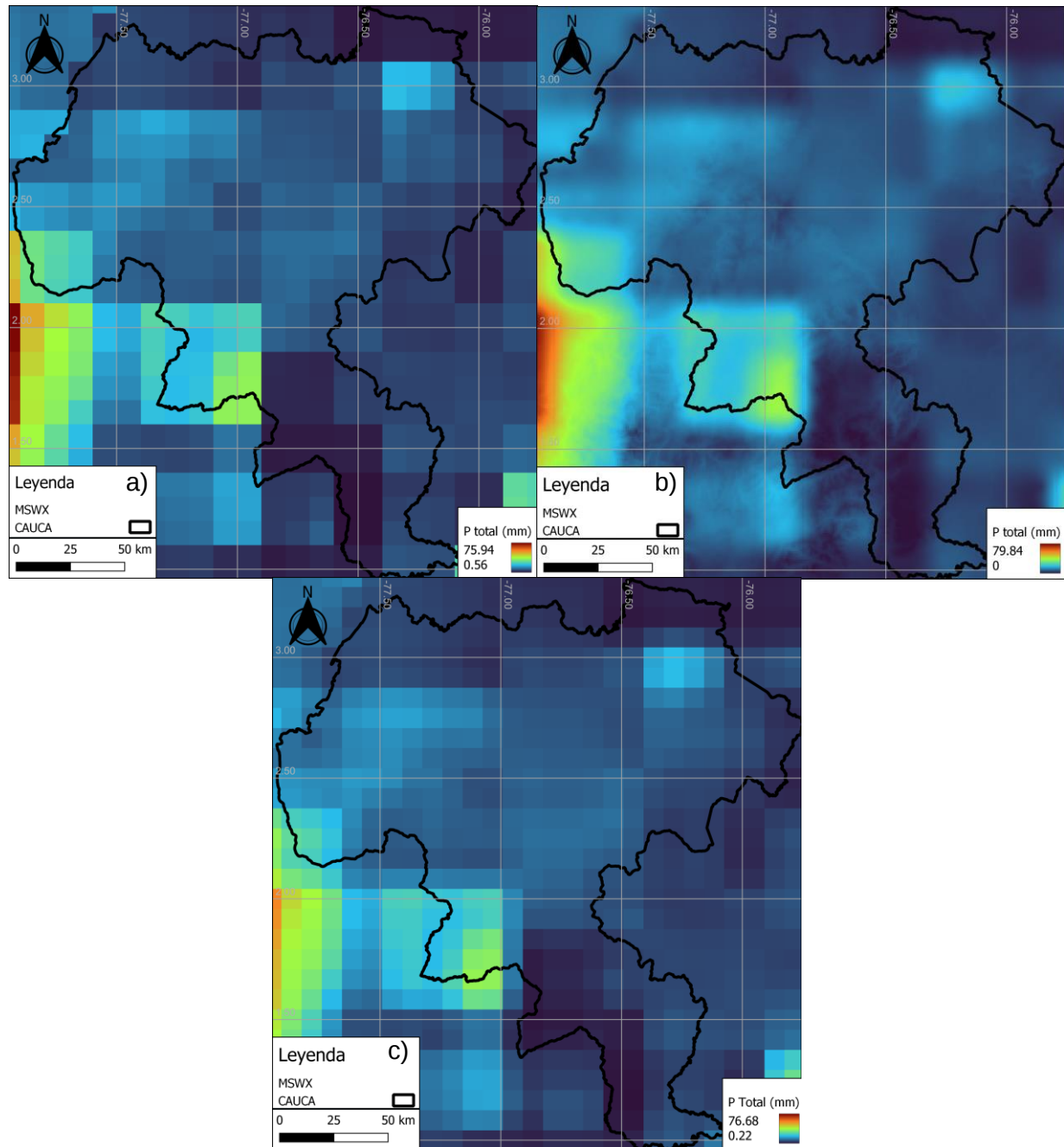
Nota. Ejemplo de datos de temperatura promedio diaria ($^\circ\text{C}$) (a) Sin escalar (0.1°); (b) Escalados (0.01°) - 17/01/2015, base de datos MSWX.

Precipitación. En las Figuras 17, 18 y 19 se presenta un ejemplo de los resultados obtenidos en el proceso de escalamiento de la variable precipitación total diaria, este proceso se realizó con dos covariantes: I) El DEM que permitió pasar de la resolución de 0.1° (10 km) a 0.01° (1 km) en las bases de datos MSWX y ERA5-Land y de 0.05° (5 km) a 0.01° (1 km) en la base de datos CHIRPS; II) El NDVI con la cual se logró una reducción de 0.1° (10 km) a 0.08° (8 km) en las bases de datos MSWX y ERA5-Land. La información completa de esta base de datos se encuentra en el apéndice J.

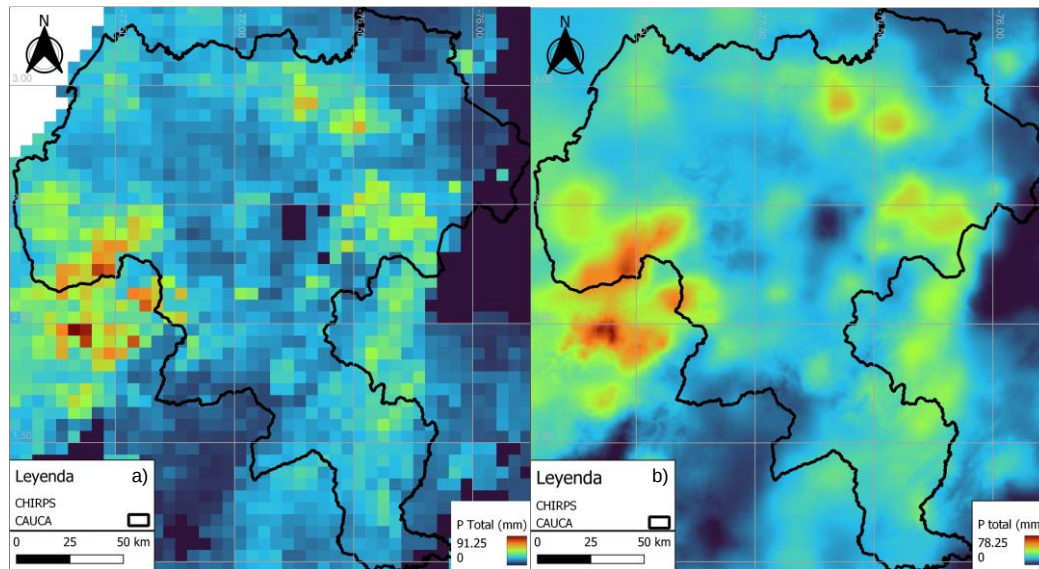
Figura 17. *Precipitación Total Diaria Escalada ERA5-Land*



Nota. Ejemplo de dato de precipitación total diaria (m) (a) Sin escalar (0.1°); (b) Escalados DEM (0.01°); (c) Escalados NDVI (0.08°) - 17/01/2015, base de datos ERA5-Land.

Figura 18. *Precipitación Total Diaria Escalada MSWX*

Nota. Ejemplo de datos de precipitación total diaria (mm) (a) Sin escalar (0.1°); (b) Escalados DEM (0.01°); (c) Escalados NDVI (0.08°) - 17/01/2015, base de datos MSWX.

Figura 19. *Precipitación Total Diaria Escalada CHIRPS*

Nota. Ejemplo de datos de precipitación total diaria (mm) (a) Sin escalar (0.1°); (b) Escalados DEM (0.01°) - 17/01/2015, base de datos CHIRPS.

7.3 Grado de Incertidumbre

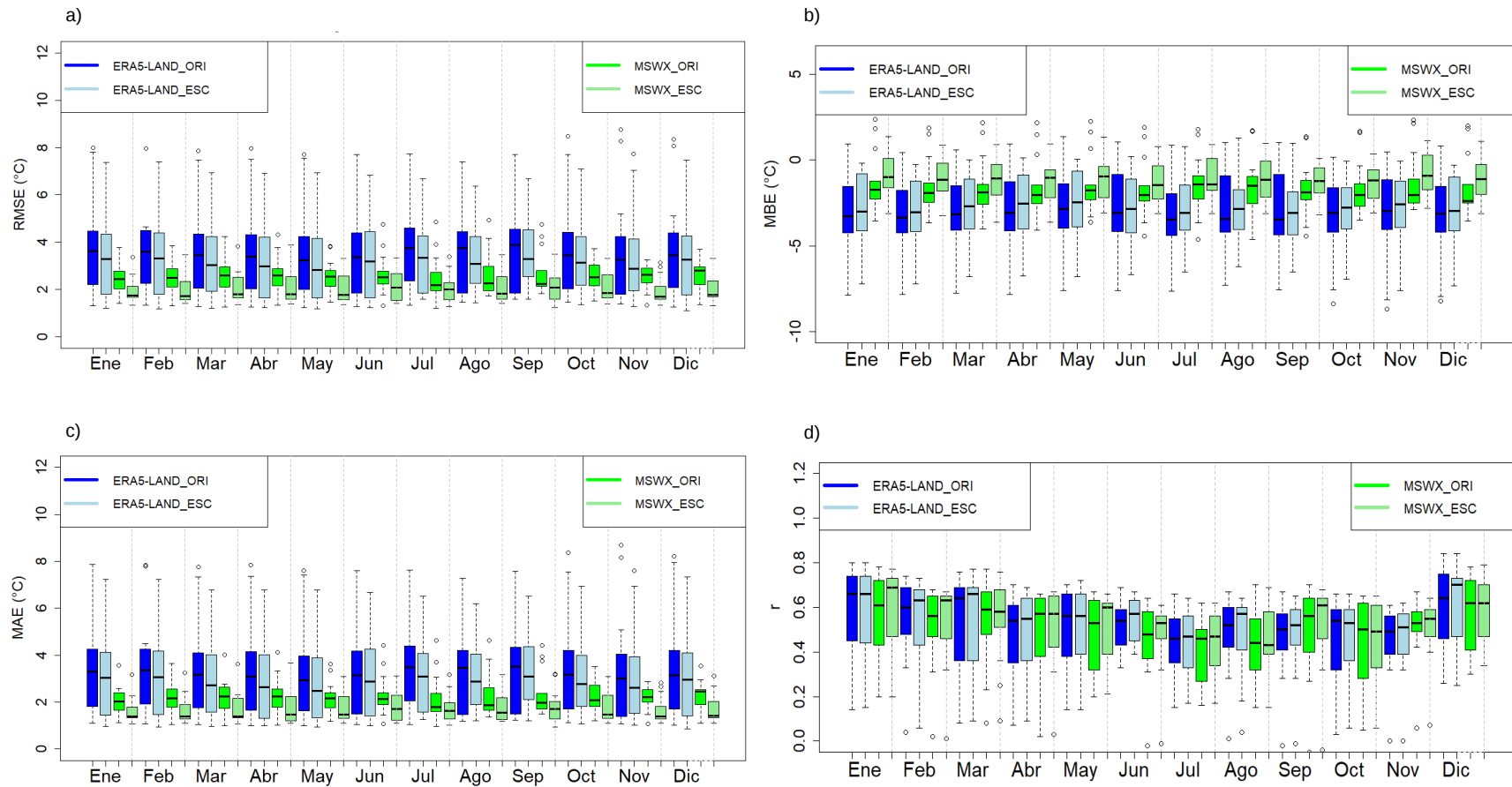
7.3.1 Temperatura Máxima

En la Figura 20 se presentan los resultados de las métricas de desempeño estimadas para la variable temperatura máxima diaria de las bases de datos MSWX y ERA5-Land originales y escaladas. La Figura 20a, permite apreciar que en las dos bases de datos se presentó una tendencia a subestimar los valores de temperatura máxima en la mayoría de las estaciones, debido a que en la mayor parte del año al menos el 75% de las estaciones presentó un MBE inferior a 0, sin embargo, este comportamiento se evidencia en mayor medida para la base de datos ERA5-Land. Sumado a lo anterior, la Figura 20d muestra que en ambas bases de datos existe una correlación positiva con valores entre 0.4 y 0.7 en la mayoría de las estaciones y meses del año. Por otra parte, en las Figuras 20a y 20c se puede observar que la base de datos MSWX presentó un mejor rendimiento en las métricas RMSE ($<3^\circ\text{C}$) y MAE ($<2.5^\circ\text{C}$), lo cual nos indica, que esta base de

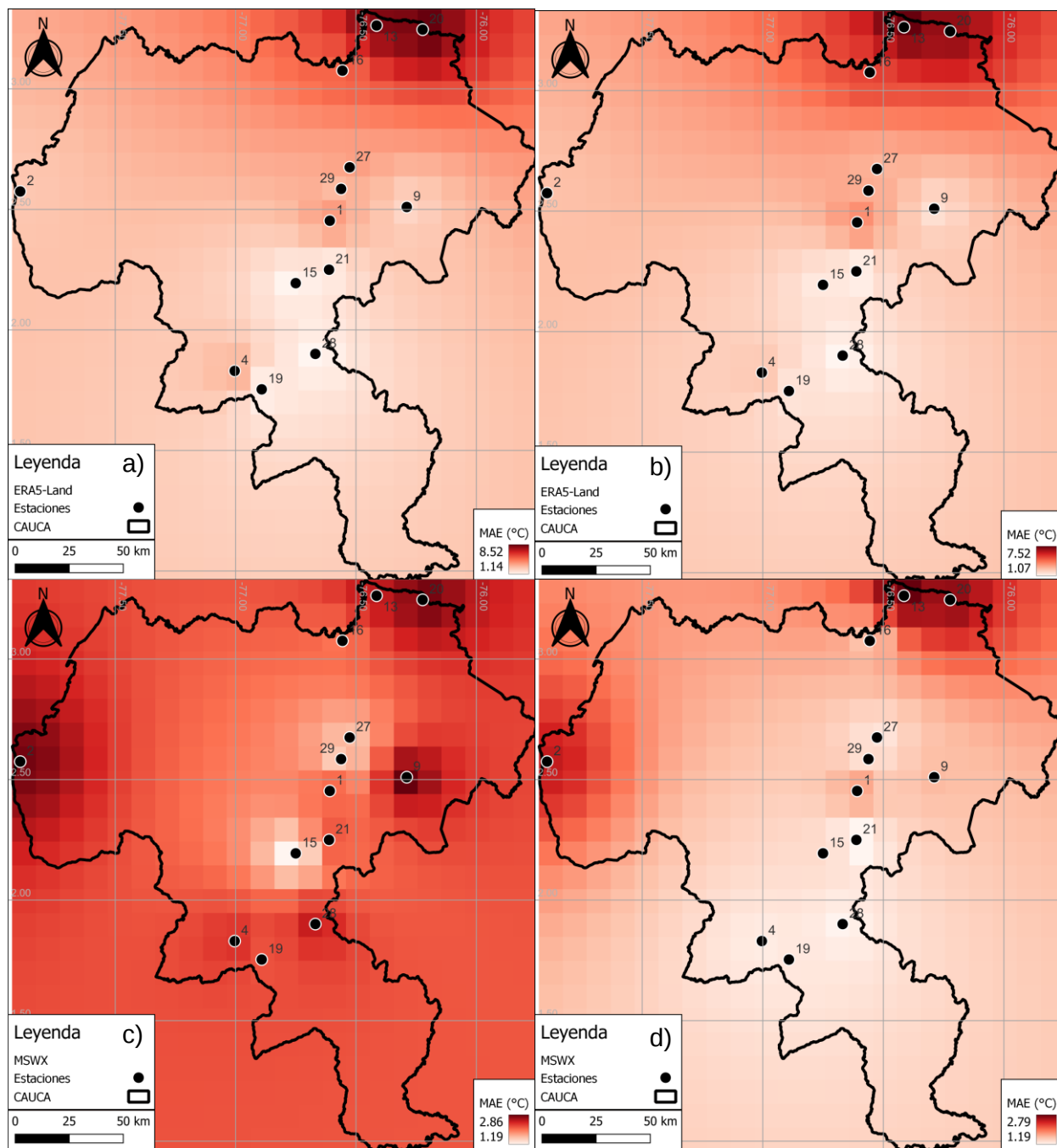
datos tiene una mayor capacidad de estimación para esta variable. Adicionalmente, se puede observar que existe un rendimiento similar de las métricas estimadas entre el periodo de lluvias y la temporada seca marcada característica del departamento.

En contraste con lo anterior, las dos bases de datos escaladas obtuvieron una mejora en el rendimiento de todas las métricas respecto a las obtenidas por los datos originales. Sumado a esto, a pesar de presentarse un comportamiento similar en el rendimiento de las bases de datos a lo largo del año, es válido resaltar que la base de datos MSWX presentó un mejor rendimiento para el mes de enero, contrario a lo evidenciado en el mes de octubre donde se percibió el peor comportamiento. Por su parte, la base de datos ERA5-Land presento en el mes de mayo su mejor comportamiento y en julio el más bajo respecto al resto del año.

Como se aprecia en la Figura 20c, el mes de noviembre se aprecia un mayor número de datos atípicos de la métrica MAE, motivo por el cual, mediante un proceso de interpolación se elaboró un mapa de distribución de esta métrica en el departamento (ver Figura 21), lo que permitió identificar que para la base de datos ERA5-Land, tanto en los datos originales como en los escalados, se presentó un bajo rendimiento al norte del departamento en las estaciones INGENIO BENGALA y MIRANDA, por su parte la base de datos MSWX, persistió el bajo rendimiento al norte del departamento, aunque en esta base de datos también se puede observar que al occidente del departamento en la estación AEROPUERTO GUAPI se presenta un bajo rendimiento con respecto a lo mostrado en el centro del departamento donde se localizan la mayor parte de las estaciones.

Figura 20. Métricas Temperatura Máxima Diaria

Nota. Métricas de desempeño temperatura máxima diaria (a) RMSE; (b) MBE; (c) MAE (d) r .

Figura 21. Distribución MAE Noviembre - Temperatura Máxima Diaria

Nota. MAE temperatura máxima diaria - noviembre (a) ERA5-Land Original; (b) ERA5-Land Escalado; (c) MSWX Original; (d) MSWX Escalado.

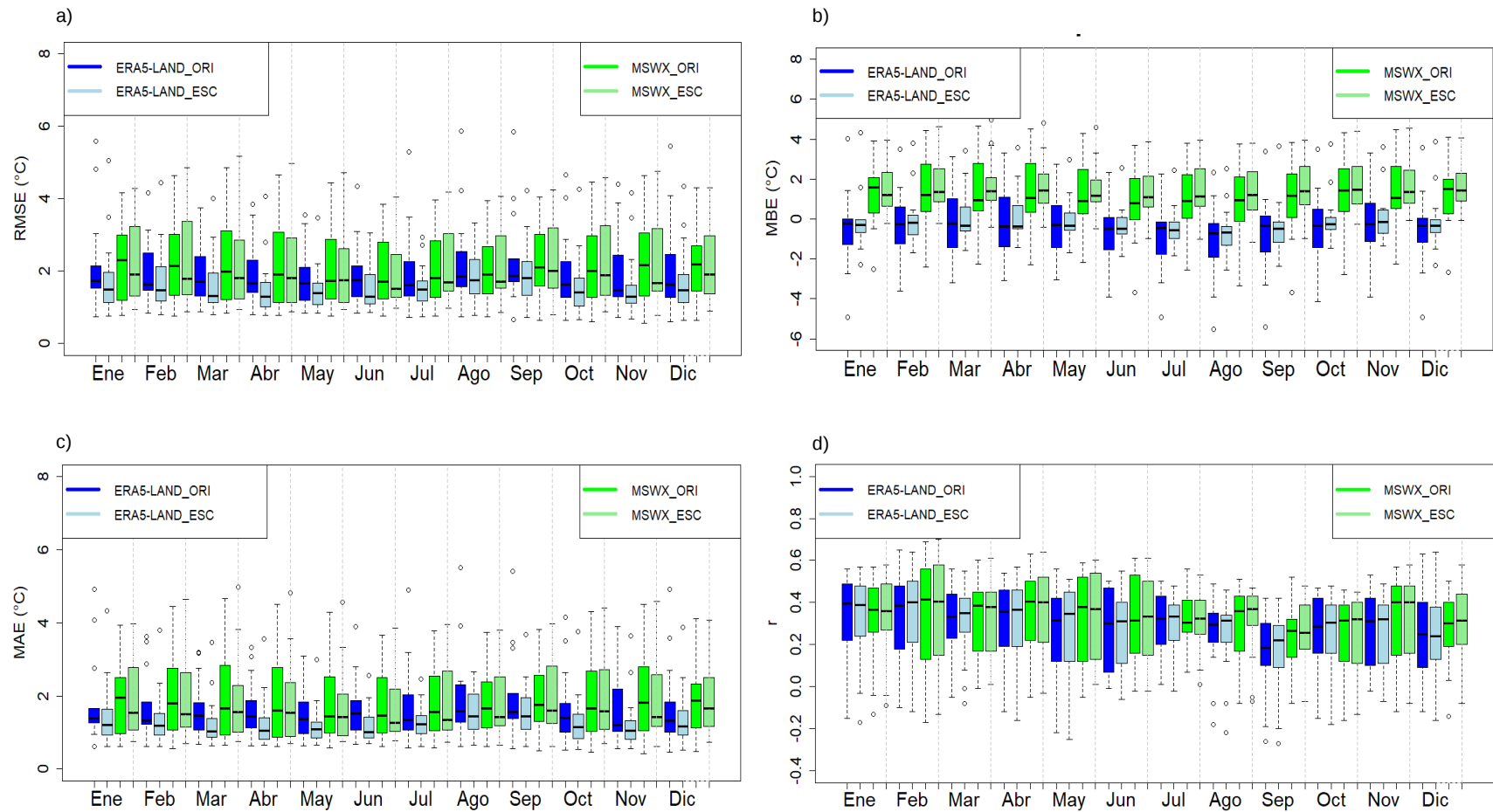
7.3.2 Temperatura Mínima

En la Figura 22 se presentan los resultados de las métricas de desempeño estimadas para la variable temperatura mínima diaria de las bases de datos MSWX y ERA5-Land originales y escaladas. En la Figura 22b se puede observar que la base de datos ERA5-Land presentó en la mayor parte del año un predominio de valores de MBE inferiores a 0, lo cual refleja una tendencia a subestimar en la mayoría de las estaciones, mientras que en la base de datos MSWX se observaron en su mayoría valores de MBE superiores a 0.5 °C, denotando una clara tendencia a sobreestimar, sumado a lo anterior, en la Figura 22d se aprecia que en ambas bases de datos existe una correlación positiva con valores entre 0.1 a 0.5 en la mayor parte de las estaciones. En contraste con lo anterior, a diferencia de la variable temperatura máxima diaria, en las Figuras 22a y 22c se evidencia que el mejor rendimiento en las métricas RMSE (<2 °C) y MAE (<1.8 °C) se percibió en la base de datos ERA5-Land, aunque, al igual que en la variable temperatura máxima, también se evidenció un rendimiento similar en el periodo seco y la temporada de lluvia.

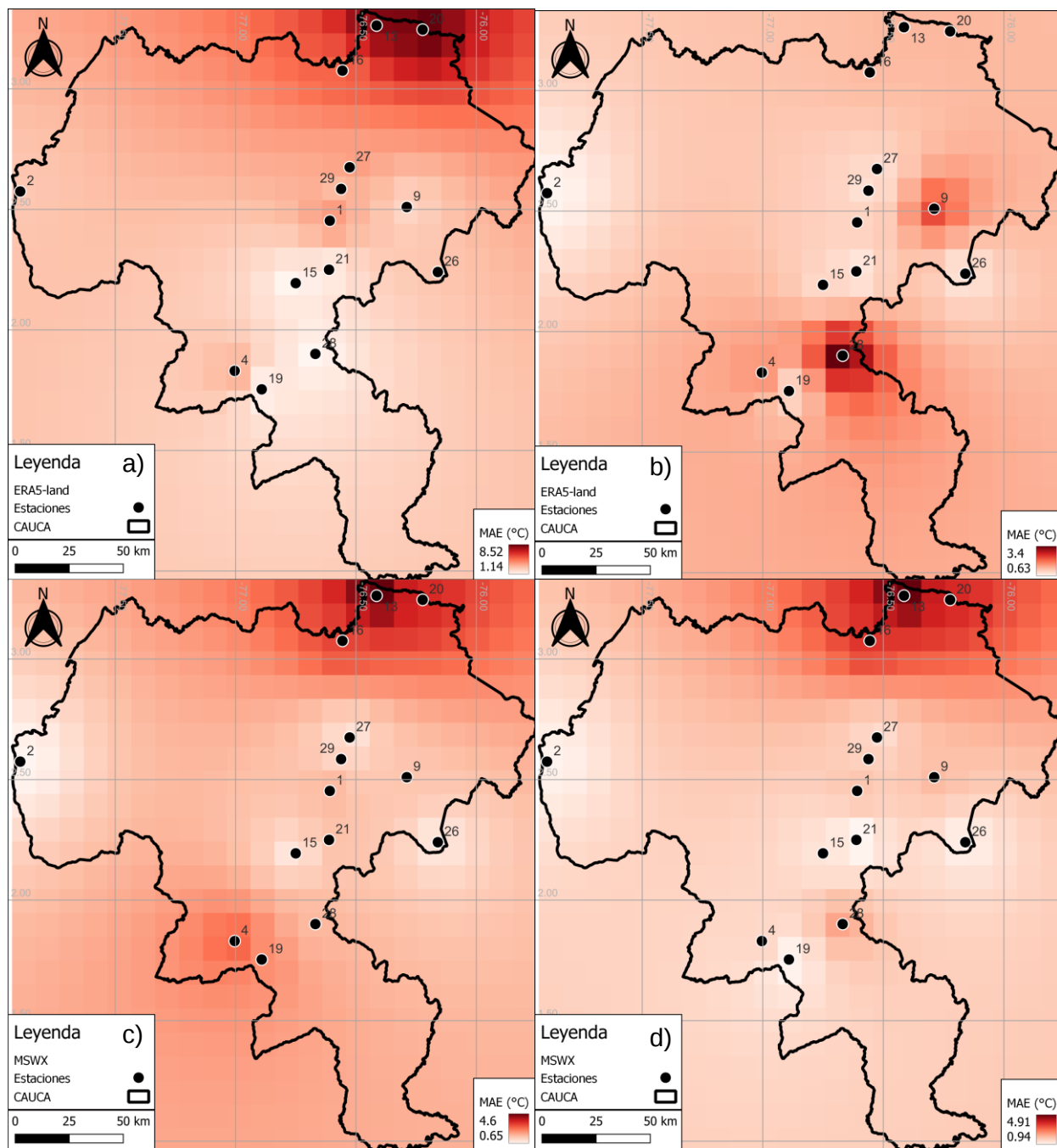
Adicionalmente, en ambas bases de datos escaladas se obtuvo una mejora en el rendimiento de las métricas calculadas respecto a los resultados obtenidos por los datos originales, aunque esta tendencia se aprecia en mayor medida para el banco de datos ERA5-Land. Sumado a lo anterior, se percibió que en ambas bases de datos se percibe un rendimiento homogéneo a lo largo del año, aunque en el mes de mayo se aprecia un mejor rendimiento en ambas, contrario a los meses de agosto y septiembre donde se observa el peor rendimiento del año.

En la Figura 22c se puede observar que el mes de marzo presenta un mayor número de datos atípicos de la métrica MAE, debido a esto, se elaboró un mapa de distribución de esta métrica en el departamento mediante un proceso de interpolación (ver Figura 23), lo cual, permitió identificar que en la base de datos ERA5-Land original, existe un bajo rendimiento al norte en la

estación MIRANDA y al centro del departamento en las estaciones BOLIVAR y VALENCIA, por su parte, la base de datos escalados presento una mejora significativa, aunque persistió el bajo rendimiento al centro del departamento en la estación VALENCIA. Por otra parte, en el MSWX al igual que en la variable temperatura máxima, persistió el bajo rendimiento al norte del departamento, en las estaciones INGENIO BENGALA, LOMITAS y MIRANDA.

Figura 22. Métricas Temperatura Mínima Diaria

Nota. Métricas de desempeño Temperatura Mínima Diaria (a) RMSE; (b) MBE; (c) MAE (d) r.

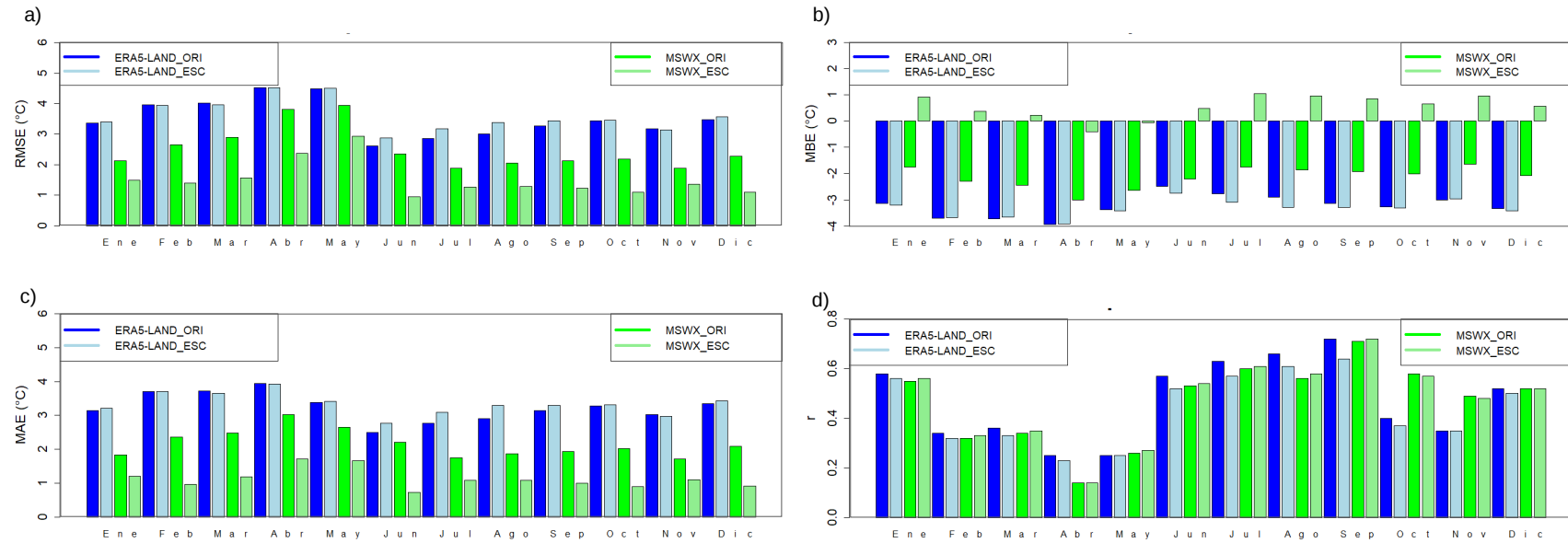
Figura 23. Distribución MAE marzo - Temperatura Mínima Diaria

Nota. MAE Temperatura mínima diaria - marzo (a) ERA5-Land Original; (b) ERA5-Land Escalado; (c) MSWX Original; (d) MSWX Escalado.

7.3.3 *Temperatura Promedio*

En la Figura 24 se presentan los resultados de las métricas de desempeño estimadas para la variable temperatura promedio diaria de las bases de datos MSWX y ERA5-Land originales y escaladas. Tal y como se aprecia en la Figura 24b, en ambas bases de datos originales se presentó una tendencia a subestimar los valores de temperatura promedio en la estación CORINTO (MBE $<-1^{\circ}\text{C}$), por otra parte, a pesar de que en la base de datos escalada del ERA5-Land persistió esta tendencia, para los datos escalados del MSWX predominó la tendencia a la sobreestimación en la mayor parte del año. Adicionalmente, la Figura 24d muestra que en las dos bases de datos existe una correlación positiva con valores entre 0.3 y 0.7 en la mayor parte del año. Sumado a lo anterior, se evidenció un rendimiento similar en el periodo seco y la temporada de lluvia, con un comportamiento ligeramente superior para el periodo seco en la base de datos ERA5-Land y en el periodo de lluvia para la base de datos MSWX. Del mismo modo, las figuras 24a y 24c permiten inferir que la base de datos MSWX, tanto en datos originales como escalados, presentó un mejor rendimiento en la estimación de datos de temperatura promedio en todo el año, debido a que presentó valores inferiores en las métricas RMSE ($<4^{\circ}\text{C}$) y MAE (2°C).

Además, en todas las métricas presentadas en la Figura se observó a lo largo del año, una mejora en el rendimiento de los datos escalados respecto a los datos originales de la base de datos MSWX, mientras que en los datos escalados del ERA5-Land solo se percibió una leve mejora en los meses de febrero, marzo, y noviembre. Sumado a lo anterior, en ambas bases de datos se apreció un mejor comportamiento en el segundo semestre del año, aunque, en el mes de junio se apreciaron los mejores resultados en las métricas MAE y RMSE, mientras que el peor rendimiento se presentó en el mes de abril.

Figura 24. Métricas Temperatura Promedio Diaria

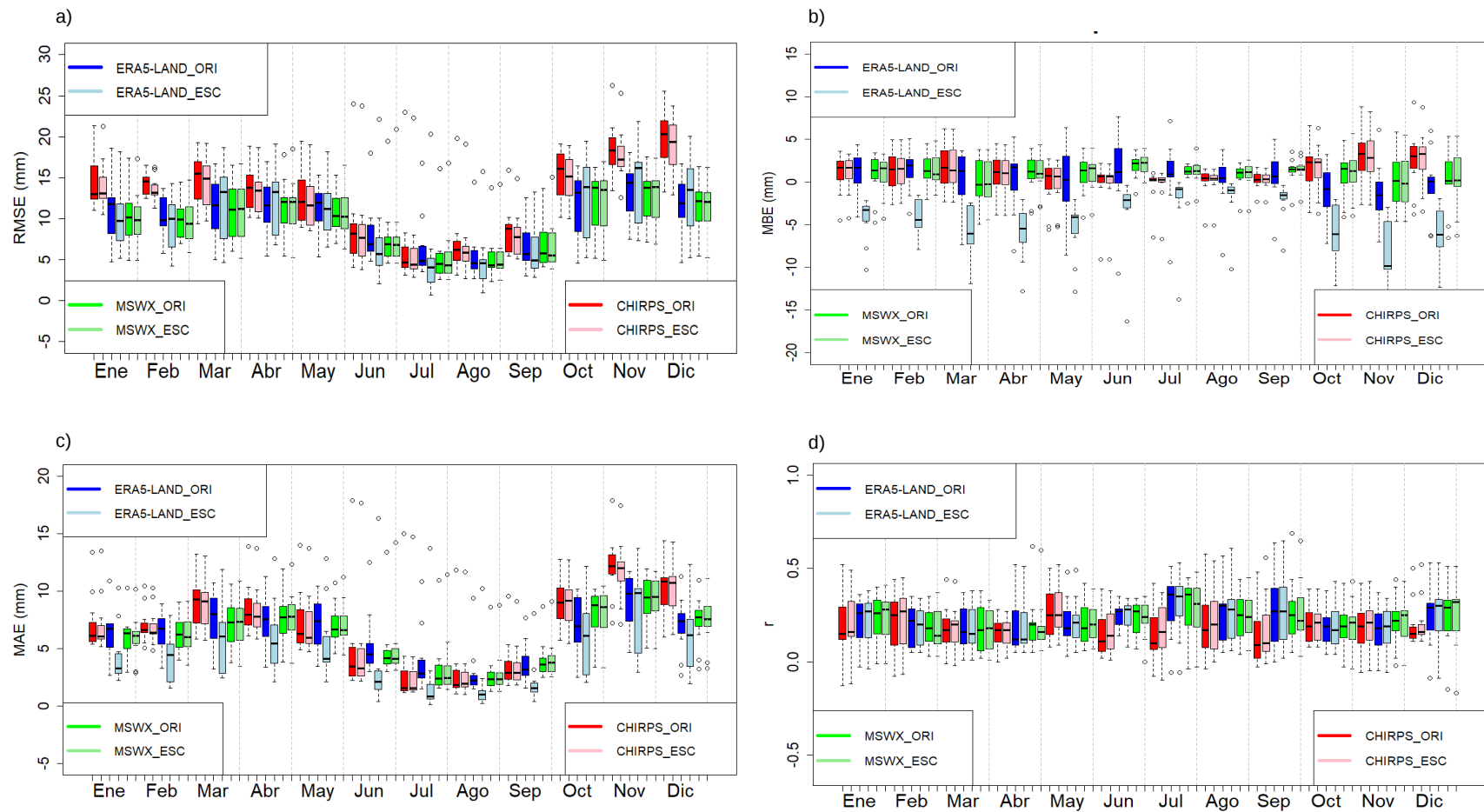
Nota. Métricas de desempeño temperatura promedio diaria (a) RMSE; (b) MBE; (c) MAE (d) r.

7.3.4 Precipitación

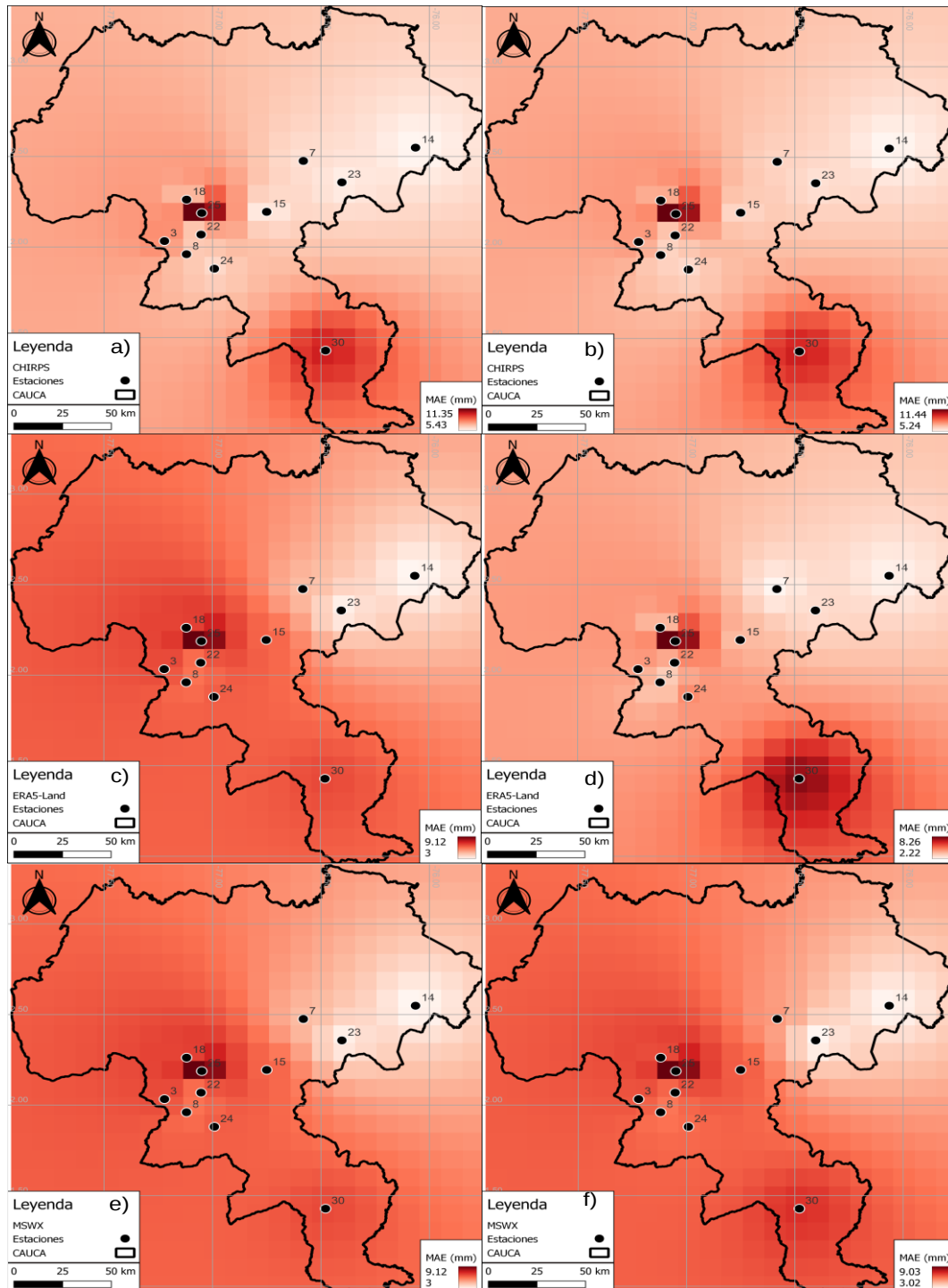
En la Figura 25 se presentan los resultados de las métricas de desempeño estimadas para la variable temperatura promedio diaria de las bases de datos CHIRPS, MSWX y ERA5-Land originales y escaladas con la covariante DEM. De la Figura 25b podemos inferir que las bases de datos CHIRPS y MSWX presentaron valores de MBE superiores a 0 en la mayor parte del año, lo cual, refleja una tendencia a sobreestimar los valores de precipitación total diaria en la mayoría de las estaciones, por otra parte, los datos originales y escalados del ERA5-Land, mostraron en su mayoría valores inferiores a 0, reflejando un comportamiento de subestimación de esta variable en gran parte de las estaciones. Adicional a lo anterior, en la Figura 25d se evidencia que las tres bases de datos presentan en la mayoría de las estaciones una correlación lineal positiva con valores inferiores a 0.5. Por otra parte, en las Figuras 25a y 25c se puede apreciar que el mejor rendimiento tanto en los datos originales como en los escalados, se presentó en la base de datos ERA5-Land, mientras que, los datos del MSWX presentaron el peor comportamiento en el periodo seco y los datos CHIRPS en la temporada de lluvia.

En contraste con lo anterior, en la mayor parte del año todas las bases de datos escaladas presentaron una mejora en las métricas estimadas, respecto a los valores obtenidos en los datos originales. Adicionalmente, a diferencia de las variables de temperatura, en todas las bases de datos de precipitación se evidenció un mejor comportamiento de las métricas de desempeño en la temporada seca ($MAE < 5mm$) respecto al periodo de lluvias ($MAE > 8mm$), lo cual representa una marcada tendencia a reducir su capacidad de estimación en esta época del año. Sumado a lo anterior, el mes de agosto registró el mejor comportamiento en todas las bases de datos (originales y escaladas), mientras que, en el mes de noviembre se observaron los peores resultados.

En base a lo observado en la Figura 25c, en el mes de enero se aprecia la presencia de un mayor número de datos atípicos de la métrica MAE, debido a esto, se elaboró un mapa de distribución de esta métrica en el departamento mediante un proceso de interpolación (ver Figura 26), lo anterior, permitió identificar que en la base de datos CHIRPS tanto para los datos originales y escalados existe un bajo rendimiento en el centro del departamento en la estación SAJANDI y al sur en la estación VILLALOBOS. Por su parte en la base de datos ERA5-Land, persiste el bajo rendimiento en estas estaciones, aunque, de una forma más marcada para el caso de los datos escalados. Finalmente, el MSWX, al igual que las otras dos bases de datos, presenta un bajo rendimiento en estas estaciones.

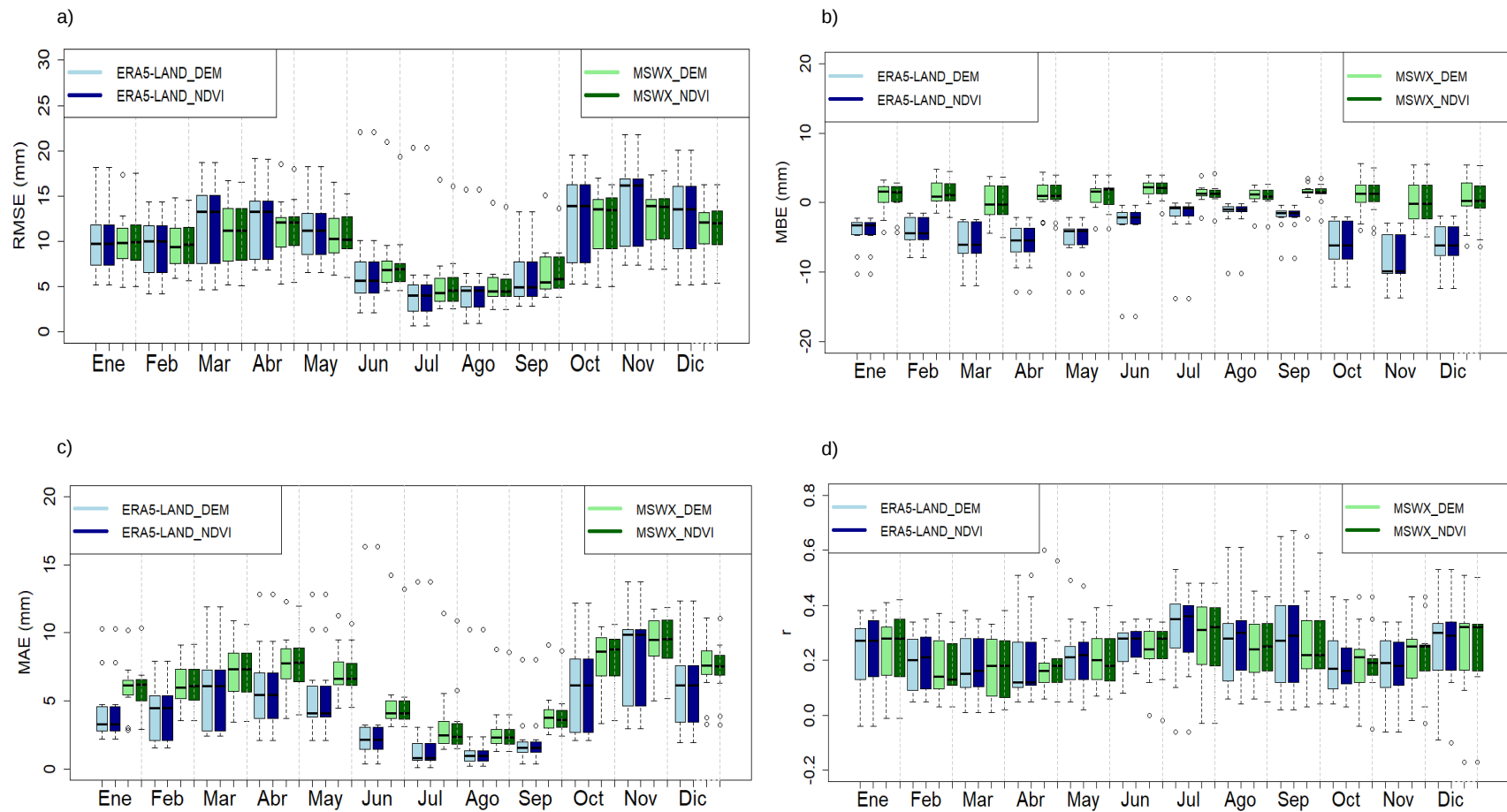
Figura 25. Métricas Precipitación Total Diaria

Nota. Métricas de desempeño precipitación total diaria (a) RMSE; (b) MBE; (c) MAE (d) r .

Figura 26. Distribución MAE enero - Precipitación Total Diaria

Nota. MAE Precipitación Total Diaria - Enero (a) CHIRPS Original; (b) CHIRPS Original; (c) ERA5-Land Original; (d) ERA5-Land Escalado; (e) MSWX Original; (f) MSWX Escalado.

El proceso de escalamiento de la variable precipitación implementando como covariante el NDVI, se realizó con las bases de datos ERA5-Land y MSWX debido a que la resolución disponible de esta covariante permitía llegar hasta un pixel de 0.08° (8 km), lo que obligó a descargar su uso para la base de datos CHIRPS, la cual cuenta con una resolución disponible de 0.05° (5 km). Las métricas de los datos escalados mediante el uso de la covariante NDEVI no mostraron una variación clara respecto a las métricas de desempeño del proceso de escalamiento usando como covariante el DEM (Figura 27).

Figura 27. Métricas Precipitación Total Diaria Escalada (DEM y NDVI)

Nota. Métricas de desempeño Precipitación Total Diaria Escalada (DEM y NDVI) (a) RMSE; (b) MBE; (c) MAE (d) r.

8. Discusión

8.1 Datos Sin Escalar

Previo a la evaluación de las bases de datos escaladas obtenidas en el presente estudio, es de gran interés conocer el rendimiento presentado por las fuentes de datos utilizadas en el proceso de escalamiento (CHIRPS, ERA5-Land y MSWX) para la estimación de las variables precipitación total, temperatura máxima, mínima y promedio diaria, debido a que no se cuenta con un precedente que valide de estas fuentes en el departamento del Cauca, Colombia.

Para el caso de la base de datos CHIRPS, la variable precipitación total diaria en la mayoría de las estaciones registró una correlación positiva con valores inferiores a 0.40, una tendencia a sobreestimar ($MBE > 0$ mm) y un RMSE inferior a 20 mm. Estos valores se encuentran cercanos a los obtenidos por López-Bermeo et al. (2022) en el departamento de Antioquia, Colombia, donde también se obtuvo una correlación (r) positiva, para este caso entre 0.13 – 0.29 en la mayoría de las estaciones evaluadas, sumado a esto, también se presentó en la mayoría de las estaciones una tendencia a sobreestimar ($MBE > 6$ mm) y un RMSE entre 10 mm – 20 mm. Adicionalmente, en la República Islámica de Pakistán (Nadeem et al., 2022), también se obtuvo una correlación positiva entre 0.10 y 0.40, sin embargo se percibió una tendencia a subestimar ($MBE < 0$ mm). Por otra parte, en el departamento de Nariño, ubicado al sur del departamento del cauca, también se evidenció la tendencia a sobreestimar ($MBE > 0$), en este caso la evaluación se realizó a una escala anual, la cual, también evidenció una correlación positiva pero con valores entre 0.60 – 0.90 en la mayoría de las estaciones evaluadas (Ocampo-Marulanda et al., 2022). En este orden de ideas, se puede inferir que en el departamento del Cauca esta base de datos, presentó un comportamiento similar al identificado en otras zonas del país.

Los resultados obtenidos en la base de datos ERA5-Land en el departamento del Cauca, muestran para el caso de la variable precipitación total diaria una correlación positiva con valores entre 0.1 – 0.4, una tendencia sobreestimar ($MBE > 0$ mm) y un RMSE entre 5 mm – 10 mm. Comportamiento contrario al encontrado en el estudio desarrollado por Nadeem et al., (2022) en la República Islámica de Pakistán, donde, se presentó una correlación positiva entre 0.1 – 0.3, pero, predominó la tendencia a la subestimación ($MBE < 0$ mm). En Colombia, se tiene el precedente de la evaluación de la precipitación mensual producto del ERA5 en la cuenca del río Sinu (Vegadurán et al., 2021), en el departamento de Córdoba, para este caso los datos presentaron una tendencia a sobreestimar ($MBE > 0$ mm), una correlación lineal positiva con valores entre 0.3 – 0.7 en la mayor parte del año, en este estudio también se evidenció un mejor rendimiento en los meses de junio a septiembre (Temporada de baja precipitación), respecto a los meses de octubre a diciembre (Temporada de alta precipitación).

Por otra parte, para las variables temperatura máxima, mínima y promedio diaria de la base de datos ERA5-Land presentaron una correlación positiva entre 0.2 – 0.6, una tendencia a subestimar, un RMSE inferior a 4°C y un MAE inferior a 4°C en la mayoría de las estaciones del departamento del Cauca. En Colombia no se cuenta con estudios de validación de esta base de datos para estas variables, sin embargo, en los glaciales de los Andes Peruanos (Bonshoms et al., 2022), se han evaluado los datos de temperatura media del aire a una escala horaria ofrecida por el ERA5, donde se obtuvo una correlación positiva entre 0.80 – 0.90, una tendencia sobreestimar en la temporada de lluvia y a subestimar en el periodo seco, sumado a lo anterior se aprecia un mejor comportamiento de las métricas RMSE ($< 2^{\circ}\text{C}$) y MAE ($< 1.8^{\circ}\text{C}$), respecto a los resultados obtenidos en el departamento del Cauca.

Por último, la base de datos MSWX en el departamento del Cauca presentó una correlación positiva entre 0.10 – 0.50, una tendencia a sobreestimar y un RMSE entre 5 mm – 15 mm para la variable precipitación total diaria. Por otra parte, para las variables temperatura máxima, mínima y promedio diaria se presentaron correlaciones positivas entre 0.2 – 0.6, una tendencia a subestimar para el caso de la temperatura máxima y promedio y a sobrestimar en la temperatura mínima, adicionalmente, presentaron un RMSE inferior a 3 °C. En Colombia, no se cuentan con estudios de evaluación de esta base de datos, sin embargo, esta ha sido evaluada en un estudio desarrollado en la cuenca del río Karasu (Hafizi & Sorman, 2022), en Turquía, encontrando que para las variables precipitación, temperatura máxima, mínima y promedio diaria existe una correlación positiva de 0.59, 0.99, 0.96 y 1.00 respectivamente, por otra parte, se percibió una tendencia de estimación similar a la presentada en el departamento del Cauca para las variables de precipitación, temperatura máxima y mínima.

8.2 Datos Escalados

Los resultados del presente estudio, evidenciaron un mejor comportamiento de las métricas de desempeño estimadas a partir de las bases de datos escaladas, respecto a los resultados obtenidos de los datos originales, lo que implica, que el proceso de escalamiento mejoró la capacidad de estimación de las variables de estudio en el departamento del Cauca. Sin embargo, es importante contextualizar las métricas estimadas a partir de los datos escalados, en relación a los resultados obtenidos en otros estudios de escalamiento.

La base de datos escalados del ERA5-Land presentó el mejor rendimiento de la variable precipitación diaria, con un RMSE entre 5mm – 15 mm, una correlación positiva entre 0.1 – 0.5 y una tendencia a subestimar. Contrario a lo obtenido en el estudio desarrollado por López et al., (2018) para la reducción de escala de la variable precipitación mensual en la cuenca del río

Magdalena, Colombia, que permitió obtener datos de precipitación total mensual a una resolución espacial de 1 km, en este estudio se presentó una tendencia a sobrestimar los datos, sin embargo, para el caso de la correlación se asemeja el comportamiento con valores entre 0.2 a 0.6.

Los resultados de escalamiento de la precipitación diaria en el departamento del Cauca, presentan un menor rendimiento respecto a los datos obtenidos mediante la herramienta desarrollada por Sen Gupta & Tarboton (2016), en la cual, los datos escalados en Estados Unidos a una resolución de 100 m obtuvieron un valor de RMSE igual a 3.73 mm. Sumado a lo anterior, los datos de precipitación total diaria escalados en China a una resolución de 1 km por Fengrui et al. (2020b), también mostraron un mejor rendimiento con valores de RMSE entre 3.59 mm y 6.41 mm y una correlación positiva con valores entre 0.6 a 0.9. Por otra parte, los datos escalados de precipitación diaria a una resolución de 25 km obtenidos en Brasil a partir de estaciones locales, en el estudio desarrollado por Xavier et al. (2016) arrojaron valores de correlación entre 0.36 a 0.72 y RMSE entre 6.8 mm a 13 mm, los cuales, se asemeja al rendimiento obtenido por la base de datos escaladas del ERA5-Land en el departamento del Cauca.

Contrario a la tendencia de estimación de precipitación diaria del banco de datos escalados del ERA5-Land en el departamento del Cauca, las bases de datos escaladas de CHIRPS y el MSWX presentaron una tendencia a sobrestimar. Esto concuerda con el comportamiento evidenciado en el estudio desarrollado por Karbalaye Ghorbanpour et al., (2021), mediante el cual se evaluaron cinco metodologías para la reducción de escala (Support vector machine, random forest, geographically weighted regression, multiple linear regression y exponential regresion) de la variable precipitación total anual, dando como resultado una resolución espacial de 1 Kilometro, se obtuvo una tendencia a la sobreestimación ($MBE > 0$). Adicional a lo anterior, el banco de datos de precipitación anual escalados, en el estudio desarrollado por Jia et al., (2011), ubicado en la

cuenca del río Qaidam, en China, mediante el cual se logró una resolución espacial de 1 kilómetro para esta variable y cuyo comportamiento también presentó una marcada tendencia a sobreestimar. Por otra parte, el proceso de escalamiento realizado por Shi & Song, (2015), que permitió obtener datos de precipitación total mensual a una escala espacial de 1 kilómetro, en la meseta Tibetana, mostró también una tendencia a sobrestimar. Para finalizar la discusión de los resultados obtenidos en la variable precipitación, se puede tomar como referencia los datos obtenidos por Lu et al., (2022), donde mediante el método Kriging se obtuvo una resolución de 1 km para la variable precipitación horaria en las montañas Tianshan, en China, donde a diferencia del presente estudio, se apreció una reducción en la capacidad para la estimación entre los datos originales (RMSE 0.3 mm - 0.5 mm) respecto a los datos escalados (RMSE 0.4 mm - 0.6 mm).

Por otra parte, para el caso de la variable temperatura máxima diaria, el banco de datos escalados del MSWX presentó el mejor rendimiento con un RMSE entre 1 °C y 3 °C, una correlación positiva entre 0.4 – 0.7 y una marcada tendencia a subestimar. En la variable temperatura mínima diaria, el banco de datos escalados del ERA5-Land mostró el mejor rendimiento, con un RMSE entre 1 °C y 2 °C, una correlación positiva entre 0.1 – 0.5 y una tendencia a subestimar. Finalmente, para la variable temperatura promedio diaria, el banco de datos del MSWX contó con el mejor rendimiento, con un RMSE entre 1 °C y 2 °C, una correlación positiva entre 0.1 – 0.7 y una tendencia a sobrestimar. En contraste con lo anterior, el banco de datos escalados por Sen Gupta & Tarboton (2016), en la cual, se obtuvo una resolución de 100 m, presentó un valor de RMSE igual a 3.98 °C, 4.72 °C, 4.20 °C, para las variables temperatura promedió, mínima, máxima diaria respectivamente, ante esto, se puede apreciar que en términos generales los valores escalados en el departamento del Cauca. Por otra parte, es válido resaltar el estudio realizado por Vaibhav Sharma et al., (2020), evaluó tres métodos (Resolution adjusted

regression downscaling, Kriging Regression y Regression Tree) mediante los cuales se obtuvieron una resolución espacial de 100 m, los cuales obtuvieron un rendimiento superior a los resultados obtenidos en el presente estudios con valores de RMSE y MAE inferiores a 1°C.

Sumado a lo anterior, el estudio desarrollado por Viggiano et al. (2019), en el cual se obtienen valores escalados de temperatura a una resolución de 2 km en el norte de Italia, presentó valores de RMSE entre 1.00 °C a 1.24 °C para la variable temperatura máxima diaria y RMSE entre 1.54 °C a 1.86 °C para la variable temperatura mínima diaria, respecto a esto, se puede apreciar que estos resultados presentan un mejor rendimiento frente a los datos escalados de las bases de datos MSWX y ERA5-Land. Por otra parte, de Carvalho Alves et al., (2022) empleó el método Kriging universal para la construcción de superficies de temperatura promedio mensual con una resolución de 1 km, en el estado de Minas Gerais, en Brasil, este proceso arrojó como resultado valores de RMSE entre 0.7 a 1.2 °C, los cuales se encuentran cercanos a los datos obtenidos en la reducción de escala de la base de datos MSWX.

8.3 Beneficios Proyecto INS

Finalmente, a partir de los estudios anteriormente presentados podemos inferir que el rendimiento de los datos obtenidos en el presente estudio, para las variables temperatura máxima, mínima y promedio, se encuentran cercanos a las métricas conseguidas por otros autores, lo cual permite validar las bases de datos escaladas. Por otra parte, para el caso de la variable precipitación existen metodologías con un mejor rendimiento, sin embargo, la mejora en cuanto a la capacidad de estimación obtenida en el proceso de escalamiento, permite admitir como válida el uso de esta información. Adicionalmente, las bases de datos escaladas en el presente estudio, fueron empleadas en el proyecto desarrollado por el INS para la construcción del sistema de alerta temprana para el seguimiento del riesgo por dengue en los municipios de Patía, Piamonte y

Miranda, del departamento del Cauca, permitiendo una mejora del 19% del R^2 en los modelos de regresión. Sumado a esto, la base de datos escaladas de temperatura máxima, mínima y promedio diaria permitieron diferenciar la variación de estas entre los barrios que conformaban las cabeceras municipales (Lesmes et al., 2022). Lo que permitió una mejora en el rendimiento de los modelos MAXENT para la presencia de mosquitos *Aedes Aegypti* en los municipios de Patía, Miranda y Piamonte (Méndez-Cardona et al., 2022).

9. Conclusiones

Los resultados obtenidos en el proceso de escalamiento, validan la selección de la metodología kriging, mediante el uso de la librería KrigR del software RStudio, debido a que este permitió la reducción de escala de las variables precipitación total, temperatura máxima, mínima y promedio diaria, a una resolución de 0.01° (1 km).

La base de datos ERA5-Land sin escalar presentó un mejor rendimiento en las variables precipitación y temperatura mínima diaria, mientras que, por su parte la base de datos MSWX mostró un mejor rendimiento en las variables temperatura máxima y promedio diario. Adicionalmente, para la variable precipitación se obtuvo un mejor rendimiento para el periodo seco frente a la época de lluvia, diferente, al caso de las variables temperatura máxima, mínima y promedio diaria, en las que no se presentaron diferencias significativas en el rendimiento de los datos entre estos dos periodos.

Se obtuvo una reducción general en el error asociado a los datos estimados en el proceso de reducción de escala, respecto a los valores de las bases de datos originales. Sumado a esto, la base de datos escalados del ERA5-Land, presentó el mejor rendimiento para en las variables

precipitación y temperatura mínima diaria, mientras que, para las variables temperatura máxima y promedio diaria, se obtuvieron mejores resultados con la base de datos escalados del MSWX.

Contrario a lo que se esperaba, la variación en el proceso de escalamiento de la variable precipitación mediante el uso de la covariante DEM y el NDVI, no presentó una diferencia significativa en los resultados obtenidos, solo en el caso del MSWX se apreció una leve mejora mediante el uso del NDVI, sin embargo, esta no compensa la limitación respecto a la resolución obtenida mediante esta covariante de 0.08° (8 km).

10. Recomendaciones

Con el fin de poder mejorar los resultados obtenidos en el presente estudio, mediante el uso de la metodología kriging, se podría plantear la búsqueda de nuevas fuentes de covariantes que permitan optimizar el proceso de escalamiento, como, por ejemplo, una base de datos de DEM, que permita reducir la escala espacial a un valor inferior a 0.01° (1 km).

Si bien, se presentó una mejora en el rendimiento de los datos obtenidos en el escalamiento de la variable precipitación, respecto a los datos originales, se podría plantear la búsqueda de otra covariante que permita mejorar las métricas obtenidas, a tal punto, de llevarlas a los valores cercanos encontrados en los estudios consultados, o, de ser posible la implementación otra metodología para el escalamiento de esta variable.

Con el objeto de enriquecer la validación de la metodología empleada, su implementación en otras regiones del país, permitirá conocer el comportamiento de esta, en las diferentes condiciones climáticas con las que cuenta el territorio nacional, lo cual permitirá la

consolidación de un método que permite la obtención de datos de alta resolución espacial y continuidad temporal en aquellas regiones donde el sistema de datos in-situ es insuficiente.

Referencias Bibliográficas

- Acosta Cardona, L. A. (2015). Evaluación de factores ambientales y climáticos como elementos de riesgo asociados con la transmisión del dengue y la Leishmaniasis a diferentes escalas temporales y espaciales en Colombia. <https://repositorio.unal.edu.co/handle/unal/55890>
- Ahmed, I. S., Hassan, F. A., Sulieman, M. M., Keshavarzi, A., Elmobarak, A. A., Yousif, K. M., & Brevik, E. C. (2022). Using environmental covariates to predict soil organic carbon stocks in Vertisols of Sudan. *Geoderma Regional*, 31. Scopus. <https://doi.org/10.1016/j.geodrs.2022.e00578>
- Ahn, J., & Lee, Y. (2022). Spatial Gap-Filling of GK2A Daily Sea Surface Temperature (SST) around the Korean Peninsula Using Meteorological Data and Regression Residual Kriging (RRK). *Remote Sensing*, 14(20). Scopus. <https://doi.org/10.3390/rs14205265>
- Awada, H., Di Prima, S., Sirca, C., Giadrossic, F., Marras, S., Spano, D., & Pirastru, M. (2021). Daily actual evapotranspiration estimation in a mediterranean ecosystem from landsat observations using SEBAL approach. *Forests*, 12, 189. <https://doi.org/10.3390/f12020189>
- Babur, E., Dindaroglu, T., Danish, S., Häggblom, M. M., Ozlu, E., Gozukara, G., & Uslu, O. S. (2022). Spatial responses of soil carbon stocks, total nitrogen, and microbial indices to post-wildfire in the Mediterranean red pine forest. *Journal of Environmental Management*, 320. Scopus. <https://doi.org/10.1016/j.jenvman.2022.115939>
- Beck, H. E., Dijk, A. I. J. M. van, Larraondo, P. R., McVicar, T. R., Pan, M., Dutra, E., & Miralles, D. G. (2021). MSWX: Global 3-hourly 0.1° bias-corrected meteorological data including near real-time updates and forecast ensembles. *Bulletin of the American Meteorological Society*, 1(aop), 1–55. <https://doi.org/10.1175/BAMS-D-21-0145.1>

- Beck, H. E., Wood, E. F., Pan, M., Fisher, C. K., Miralles, D. G., Dijk, A. I. J. M. van, McVicar, T. R., & Adler, R. F. (2019). MSWEP V2 Global 3-Hourly 0.1° Precipitation: Methodology and Quantitative Assessment. *Bulletin of the American Meteorological Society*, 100(3), 473–500. <https://doi.org/10.1175/BAMS-D-17-0138.1>
- Bonshoms, M., Ubeda, J., Liguori, G., Körner, P., Navarro, Á., & Cruz, R. (2022). Validation of ERA5-Land temperature and relative humidity on four Peruvian glaciers using on-glacier observations. *Journal of Mountain Science*, 19(7), 1849–1873. Scopus. <https://doi.org/10.1007/s11629-022-7388-4>
- Brunetti, M., Maugeri, M., Nanni, T., Simolo, C., & Spinoni, J. (2014). High-resolution temperature climatology for Italy: Interpolation method intercomparison. *International Journal of Climatology*, 34(4), 1278–1296. <https://doi.org/10.1002/joc.3764>
- Ceccato, P., Ramirez, B., Manyangadze, T., Gwakisa, P., & Thomson, M. C. (2018). Data and tools to integrate climate and environmental information into public health. *Infectious Diseases of Poverty*, 7(1), 126. <https://doi.org/10.1186/s40249-018-0501-9>
- Chen, F., Gao, Y., Wang, Y., & Li, X. (2020a). A downscaling-merging method for high-resolution daily precipitation estimation. *Journal of Hydrology*, 581, 124414. <https://doi.org/10.1016/j.jhydrol.2019.124414>
- Chen, F., Gao, Y., Wang, Y., & Li, X. (2020b). A downscaling-merging method for high-resolution daily precipitation estimation. *Journal of Hydrology*, 581, 124414. <https://doi.org/10.1016/j.jhydrol.2019.124414>
- Chen, Y., Shan, B., Yu, X., Zhang, Q., & Ren, Q. (2022). Comprehensive effect of the three-dimensional spatial distribution pattern of buildings on the urban thermal environment. *Urban Climate*, 46. Scopus. <https://doi.org/10.1016/j.uclim.2022.101324>

- Chilès, J.-P., & Delfiner, P. (2012). *Geostatistics: Modeling Spatial Uncertainty: Second Edition* (p. 699). Scopus. <https://doi.org/10.1002/9781118136188>
- Condom, T., Martínez, R., Pabón, J. D., Costa, F., Pineda, L., Nieto, J. J., López, F., & Villacis, M. (2020). Climatological and Hydrological Observations for the South American Andes: In situ Stations, Satellite, and Reanalysis Data Sets. *Frontiers in Earth Science*, 8, 92. <https://doi.org/10.3389/feart.2020.00092>
- Copernicus Climate Change Service. (2019). ERA5-Land hourly data from 2001 to present [Data set]. ECMWF. <https://doi.org/10.24381/CDS.E2161BAC>
- DAMCRA-Subzonas Hidrograficas Colombia. (2022, December 13). GeoNode SIAM. http://192.168.3.87/layers/geonode%3Aszh_sep_26_2012
- Davy, R., & Kusch, E. (2021). Reconciling high resolution climate datasets using KrigR. *Environmental Research Letters*, 16(12), 124040. <https://doi.org/10.1088/1748-9326/ac39bf>
- de Carvalho Alves, M., Sanches, L., & de Carvalho, L. G. (2022). Geostatistical surfaces of climatological normals of mean air temperature in Minas Gerais. *Environmental Monitoring and Assessment*, 194(7). Scopus. <https://doi.org/10.1007/s10661-022-10162-0>
- Eguiguren-Velepucha, P. A., Chamba, J. A. M., Mendoza, N. A. A., Ojeda-Luna, T. L., Samaniego-Rojas, N. S., Furniss, M. J., Howe, C., & Mendoza, Z. H. A. (2016). Tropical ecosystems vulnerability to climate change in southern Ecuador: *Tropical Conservation Science*. <https://doi.org/10.1177/1940082916668007>
- Fall, P., Diouf, I., Deme, A., & Sene, D. (2022). Assessment of Climate-Driven Variations in Malaria Transmission in Senegal Using the VECTRI Model. *Atmosphere*, 13(3). Scopus. <https://doi.org/10.3390/atmos13030418>

- Fan, Y., & van den Dool, H. (2008). A global monthly land surface air temperature analysis for 1948-present. *Journal of Geophysical Research Atmospheres*, 113(1). Scopus. <https://doi.org/10.1029/2007JD008470>
- Funk, C., Peterson, P., Landsfeld, M., Pedreros, D., Verdin, J., Shukla, S., Husak, G., Rowland, J., Harrison, L., Hoell, A., & Michaelsen, J. (2015). The climate hazards infrared precipitation with stations—A new environmental record for monitoring extremes. *Scientific Data*, 2, 150066. <http://dx.doi.org/10.1038/sdata.2015.66>
- Gemitzi, A., Koutsias, N., & Lakshmi, V. (2021). A spatial downscaling methodology for grace total water storage anomalies using gpm imerg precipitation estimates. *Remote Sensing*, 13(24). Scopus. <https://doi.org/10.3390/rs13245149>
- Gil-Guirado, S., Pérez-Morales, A., Pino, D., Peña, J. C., & Martínez, F. L. (2022). Flood impact on the Spanish Mediterranean coast since 1960 based on the prevailing synoptic patterns. *Science of The Total Environment*, 807, 150777. <https://doi.org/10.1016/j.scitotenv.2021.150777>
- GIMMS | Global Agricultural Monitoring System. (2022). <https://glam1.gsfc.nasa.gov/>
- Giraldo-Osorio, J. D., Trujillo-Osorio, D. E., & Baez-Villanueva, O. M. (2022). Analysis of ENSO-Driven Variability, and Long-Term Changes, of Extreme Precipitation Indices in Colombia, Using the Satellite Rainfall Estimates CHIRPS. *Water (Switzerland)*, 14(11). Scopus. <https://doi.org/10.3390/w14111733>
- GISTEMP Team. (2022, December 11). NASA GISS Surface Temperature Analysis: NOAA Physical Sciences Laboratory NASA GISS Surface Temperature Analysis. <https://www.psl.noaa.gov/data/gridded/data.gistemp.html>

- Gobernación del Cauca. (2020). PÉRFIL DEPARTAMENTO DEL CAUCA. chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.cauca.gov.co/Dependencias/OficinaAsesoradePlaneacion/InformacioneIndicadores/Perfil%20Departamento%20del%20Cauca.pdf
- Hafizi, H., & Sorman, A. A. (2022). Integrating Meteorological Forcing from Ground Observations and MSWX Dataset for Streamflow Prediction under Multiple Parameterization Scenarios. *Water* (Switzerland), 14(17). Scopus. <https://doi.org/10.3390/w14172721>
- Hofstra, N., Haylock, M., New, M., Jones, P., & Frei, C. (2008). Comparison of six methods for the interpolation of daily, European climate data. *Journal of Geophysical Research: Atmospheres*, 113(D21). <https://doi.org/10.1029/2008JD010100>
- IDEAM. (2018). Metodología de la operación estadística variables meteorológicas. 02, 113.
- IDEAM. (2022a). Catálogo Nacional de Estaciones del IDEAM | Datos Abiertos Colombia. <https://www.datos.gov.co/widgets/hp9r-jxuu>
- IDEAM. (2022b, January 23). Consulta y Descarga de Datos Hidrometeorológicos. <http://dhime.ideam.gov.co/atencionciudadano/>
- IDEAM. (2023). Fenomeno Niño y Niña—IDEAM. <http://www.ideam.gov.co/web/siac/ninoynina>
- Immerzeel, W. W., Rutten, M. M., & Droogers, P. (2009). Spatial downscaling of TRMM precipitation using vegetative response on the Iberian Peninsula. *Remote Sensing of Environment*, 113(2), 362–370. <https://doi.org/10.1016/j.rse.2008.10.004>
- INS. (2020). Informe técnico de avance o final de programas y proyectos de CTel.

- INS. (2022). Boletín Epidemiológico. <https://www.ins.gov.co/buscador-eventos/Paginas/Vista-Boletin-Epidemilogico.aspx>
- Jia, S., Zhu, W., Lü, A., & Yan, T. (2011). A statistical spatial downscaling algorithm of TRMM precipitation based on NDVI and DEM in the Qaidam Basin of China. *Remote Sensing of Environment*, 115(12), 3069–3079. <https://doi.org/10.1016/j.rse.2011.06.009>
- Karbalaye Ghorbanpour, A., Hessels, T., Moghim, S., & Afshar, A. (2021). Comparison and assessment of spatial downscaling methods for enhancing the accuracy of satellite-based precipitation over Lake Urmia Basin. *Journal of Hydrology*, 596, 126055. <https://doi.org/10.1016/j.jhydrol.2021.126055>
- Kerry, R., Ingram, B., Orellana, M., Ortiz, B. V., & Scully, B. (2023). Development of a method to assess the risk of aflatoxin contamination of corn within counties in southern Georgia, USA using remotely sensed data. *Smart Agricultural Technology*, 3. Scopus. <https://doi.org/10.1016/j.atech.2022.100124>
- Khashei, M., & Bijari, M. (2010). An artificial neural network (p,d,q) model for timeseries forecasting. *Expert Systems with Applications*, 11.
- Kusch, E., & Davy, R. (2022). KrigR—a tool for downloading and statistically downscaling climate reanalysis data. *Environmental Research Letters*, 17(2), 024005. <https://doi.org/10.1088/1748-9326/ac48b3>
- Lesmes, M., Marceló, C., Ávila, Á., & Santamaría, E. (2022). Sistema de alerta temprana para el seguimiento del riesgo por dengue en municipios prioritarios del departamento del Cauca. Universidad de Ciencias Aplicadas y Ambientales.
- Li, B. Q., Hou, B. D., Dong, X. Y., & Wang, H. J. (2017). A multifactor analysis of the spatial distribution of annual mean extreme precipitation-Taking the Yellow River Basin as an

- example. IOP Conference Series: Earth and Environmental Science, 82, 012025.
<https://doi.org/10.1088/1755-1315/82/1/012025>
- Li, L., Hao, D., Li, X., Chen, M., Zhou, Y., Jurgens, D., Asrar, G., & Sapkota, A. (2022). Satellite-based phenology products and in-situ pollen dynamics: A comparative assessment. *Environmental Research*, 204, 111937. <https://doi.org/10.1016/j.envres.2021.111937>
- López López, P., Immerzeel, W. W., Rodríguez Sandoval, E. A., Sterk, G., & Schellekens, J. (2018). Spatial Downscaling of Satellite-Based Precipitation and Its Impact on Discharge Simulations in the Magdalena River Basin in Colombia. *Frontiers in Earth Science*, 6. <https://www.frontiersin.org/article/10.3389/feart.2018.00068>
- López-Bermeo, C., Montoya, R. D., Caro-Lopera, F. J., & Díaz-García, J. A. (2022). Validation of the accuracy of the CHIRPS precipitation dataset at representing climate variability in a tropical mountainous region of South America. *Physics and Chemistry of the Earth*, 127. Scopus. <https://doi.org/10.1016/j.pce.2022.103184>
- Louis, V. R., Phalkey, R., Horstick, O., Ratanawong, P., Wilder-Smith, A., Tozan, Y., & Dambach, P. (2014). Modeling tools for dengue risk mapping—A systematic review. *International Journal of Health Geographics*, 13(1), 50. <https://doi.org/10.1186/1476-072X-13-50>
- Lu, X.-Y., Chen, Y.-Y., Tang, G.-Q., Wang, X.-Q., Liu, Y., & Wei, M. (2022). Quantitative estimation of hourly precipitation in the Tianshan Mountains based on area-to-point kriging downscaling and satellite-gauge data merging. *Journal of Mountain Science*, 19(1), 58–72. Scopus. <https://doi.org/10.1007/s11629-021-6901-5>
- Ma, X. (2022). Improving the accuracy and spatial resolution of precipitable water vapor dataset using a neural network-based downscaling method. *Atmospheric Environment*, 12.

- Macharia, D., Fankhauser, K., Selker, J. S., Neff, J. C., & Thomas, E. A. (2022). Validation and Intercomparison of Satellite-Based Rainfall Products over Africa with TAHMO In Situ Rainfall Observations. *Journal of Hydrometeorology*, 23(7), 1131–1154. Scopus. <https://doi.org/10.1175/JHM-D-21-0161.1>
- Makinoshima, F., Oishi, Y., Yamazaki, T., Furumura, T., & Imamura, F. (2021). Early forecasting of tsunami inundation from tsunami and geodetic observation data with convolutional neural networks. *Nature Communications*, 12(1), 1–10. <https://doi.org/10.1038/s41467-021-22348-0>
- Mendes, J., & Maia, R. (2020). Spatial downscaling of 3-hourly precipitation forecast data at river basin scale. *Meteorology and Atmospheric Physics*, 132(1), 143–158. <https://doi.org/10.1007/s00703-019-00678-5>
- Méndez-Cardona, S., Marceló, C., Morales, C., Lesmes, M. C., Fuya, P., Cadena, H., Ávila-Díaz, A., Piamba, A., & Santamaría, E. (2022). Mosquitos adultos intradomiciliados en municipios de alto riesgo para dengue en Cauca, Colombia.
- Młyński, D., Operacz, A., & Wałęga, A. (2020). Sensitivity of methods for calculating environmental flows based on hydrological characteristics of watercourses regarding the hydropower potential of rivers. *Journal of Cleaner Production*, 250.
- 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. Scopus.
- MSWEP - GloH2O. (2020, April 13). <http://www.gloh2o.org/mswep/>
- MSWX - GloH2O. (2021, February 18). <http://www.gloh2o.org/mswx/>

- Muñoz, E., Poveda, G., Arbeláez, M. P., & Vélez, I. D. (2021). Spatiotemporal dynamics of dengue in Colombia in relation to the combined effects of local climate and ENSO. *Acta Tropica*, 224, 106136. <https://doi.org/10.1016/j.actatropica.2021.106136>
- Nadeem, M. U., Ghanim, A. A. J., Anjum, M. N., Shangguan, D., Rasool, G., Irfan, M., Niazi, U. M., & Hassan, S. (2022). Multiscale Ground Validation of Satellite and Reanalysis Precipitation Products over Diverse Climatic and Topographic Conditions. *Remote Sensing*, 14(18). Scopus. <https://doi.org/10.3390/rs14184680>
- Njoku, E. A., Akpan, P. E., Effiong, A. E., & Babatunde, I. O. (2023). The effects of station density in geostatistical prediction of air temperatures in Sweden: A comparison of two interpolation techniques. *Resources, Environment and Sustainability*, 11. Scopus. <https://doi.org/10.1016/j.resenv.2022.100092>
- NOAA. (2021). All Gridded Datasets: NOAA Physical Sciences Laboratory. <https://www.psl.noaa.gov/data/gridded/>
- Nourani, V., Farshbaf, A., & Adarsh, S. (2020). Spatial downscaling of radar-derived rainfall field by two-dimensional wavelet transform. *Hydrology Research*, 51(3), 456–469. Scopus. <https://doi.org/10.2166/nh.2020.165>
- Ocampo-Marulanda, C., Fernández-Álvarez, C., Cerón, W. L., Canchala, T., Carvajal-Escobar, Y., & Alfonso-Morales, W. (2022). A spatiotemporal assessment of the high-resolution CHIRPS rainfall dataset in southwestern Colombia using combined principal component analysis. *Ain Shams Engineering Journal*, 13(5). Scopus. <https://doi.org/10.1016/j.asej.2022.101739>
- OPS. (2021, October 10). Cambio Climático y Salud. <https://www.paho.org/es/temas/cambio-climatico-salud>

- Oropeza, J. A. P., Vélez, E. P., & Vélez, O. P. (2018). Ajuste de un variograma esférico de la precipitación anual de las normales climatológicas 1951-2010—Adjustment of a spherical variogram of the annual precipitation of the Climatological Normals 1951-2010. *Tecnología y ciencias del agua*, 9(6), Article 6. <https://doi.org/10.24850/j-tyca-2018-06-10>
- Ortega, E. (2016). *Sistemas de información Geográfica. Teoría y práctica*. <https://bibliotecavirtual.uis.edu.co:3806/a/67546/sistemas-de-informacion-geografica.-teoria-y-practica>
- Padilla, J. C., Lizarazo, F. E., Murillo, O. L., Mendigaña, F. A., Pachón, E., & Vera, M. J. (2017). Epidemiología de las principales enfermedades transmitidas por vectores en Colombia, 1990-2016. *Biomédica*, 37, 27–40. <https://doi.org/10.7705/biomedica.v37i0.3769>
- Padilla, J. C., Rojas, D. P., & Saénz-Gomez, R. (2012). Dengue en Colombia—Epidemiología de la reemergencia a la hiperendemia.
- PAHO. (2019, October 16). Third Global Conference on Health and Climate Change. <https://www.paho.org/en/news/16-10-2019-third-global-conference-health-and-climate-change>
- Paredes-Trejo, F. J., Barbosa, H. A., & Lakshmi Kumar, T. V. (2017). Validating CHIRPS-based satellite precipitation estimates in Northeast Brazil. *Journal of Arid Environments*, 139, 26–40. <https://doi.org/10.1016/j.jaridenv.2016.12.009>
- Pérez Navarro, A. (2011). *Introducción a los sistemas de información geográfica y geotelemática*. <https://bibliotecavirtual.uis.edu.co:3806/a/20241/introduccion-a-los-sistemas-de-informacion-geografica-y-geotelematica>

- Retalis, A., Tymvios, F., Katsanos, D., & Michaelides, S. (2017). Downscaling CHIRPS precipitation data: An artificial neural network modelling approach. *International Journal of Remote Sensing*, 38(13), 3943–3959. <https://doi.org/10.1080/01431161.2017.1312031>
- Reyes, A. J. R., & Instituto Nacional de Salud. (2019). Informe de evento dengue, Colombia, 2018. 04, 21.
- Río San José, J. del. (2010). Tratamiento de datos espaciales en hidrología. Bubok Publishing S.L. https://bibliotecavirtual.uis.edu.co:4259/es/lc/uis/titulos/55605?as_all=hidrolog%C3%A1Da&as_all_op=unaccent__icontains&prev=as
- Rojas, M. N. R. (2016). Usos de herramientas geoespaciales en la detección de áreas con riesgo epidemiológico a partir de variables biofísicas y casos de dengue en Jujuy—Argentina. <https://rdu.unc.edu.ar/handle/11086/2841>
- Rotela, C. H. (2019). Desarrollo de Modelos e Indicadores Remotos de Riesgo Epidemiológico de Dengue en Argentina. <https://rdu.unc.edu.ar/handle/11086/11609>
- Saragih, N. F., Sitepu, S., Simanungkalit, G. T., Sinambela, M., Rajagukguk, E., Larosa, F. G., & Jaya, I. K. (2022). Validation of CHIRPS Estimation Rainfall Data Using Numerical Accuracy Test with Precipitation Observation Data. 1083(1). Scopus. <https://doi.org/10.1088/1755-1315/1083/1/012095>
- Sen Gupta, A., & Tarboton, D. G. (2016). A tool for downscaling weather data from large-grid reanalysis products to finer spatial scales for distributed hydrological applications. *Environmental Modelling & Software: With Environment Data News*, 84, 50–69. <https://doi.org/10.1016/j.envsoft.2016.06.014>

- Serifi, A., Günther, T., & Ban, N. (2021). Spatio-Temporal Downscaling of Climate Data Using Convolutional and Error-Predicting Neural Networks. *Frontiers in Climate*, 3, 26. <https://doi.org/10.3389/fclim.2021.656479>
- Serrano-Notivolí, R., & Tejedor, E. (2021). From rain to data: A review of the creation of monthly and daily station-based gridded precipitation datasets. *Wiley Interdisciplinary Reviews: Water*. <https://wires.onlinelibrary.wiley.com/doi/full/10.1002/wat2.1555>
- Shi, Y., & Song, L. (2015). Spatial Downscaling of Monthly TRMM Precipitation Based on EVI and Other Geospatial Variables over the Tibetan Plateau from 2001 to 2012. *Mountain Research and Development*, 35(2), 180–194. Scopus. <https://doi.org/10.1659/MRD-JOURNAL-D-14-00119.1>
- UNESCO. (2011). *Manual sobre Sistemas de Alerta Temprana*.
- UNESCO. (2020, March 6). *World Water Development Report 2020 – Water and Climate Change*. UNESCO. <https://en.unesco.org/themes/water-security/wwap/wwdr/2020>
- Universidad del Valle. (2022). *Prevención de Violencia en Colombia*. chrome-extension://efaidnbmninnibpcajpcglclefindmkaj/https://prevencionviolencia.univalle.edu.co/observatorios/cauca/departamental/archivos/perfil_cauca.pdf
- Urrea, V., Ochoa, A., & Mesa, O. (2019). Seasonality of Rainfall in Colombia. *Water Resources Research*, 55(5), 4149–4162. <https://doi.org/10.1029/2018WR023316>
- USGCRP. (2017). *Climate Science Special Report* (pp. 1–470). U.S. Global Change Research Program, Washington, DC. <https://science2017.globalchange.gov/chapter/executive-summary/>
- USGS.gov | Science for a changing world. (2022, December 13). <https://www.usgs.gov/>

- Vaibhav Sharma, K., Khandelwal, S., & Kaul, N. (2020). Spatial variation of materials surface temperature by regression based downscaling model's in Jaipur district, India. *Materials Today: Proceedings*, 28, 1825–1832. <https://doi.org/10.1016/j.matpr.2020.05.286>
- Vega-durán, J., Escalante-castro, B., Canales, F. A., Acuña, G. J., & Kaźmierczak, B. (2021). Evaluation of areal monthly average precipitation estimates from merra2 and era5 reanalysis in a colombian caribbean basin. *Atmosphere*, 12(11). Scopus. <https://doi.org/10.3390/atmos12111430>
- Viggiano, M., Busetto, L., Cimini, D., Di Paola, F., Gerald, E., Ranghetti, L., Ricciardelli, E., & Romano, F. (2019). A new spatial modeling and interpolation approach for high-resolution temperature maps combining reanalysis data and ground measurements. *Agricultural and Forest Meteorology*, 276–277, 107590. <https://doi.org/10.1016/j.agrformet.2019.05.021>
- Wang, C.-L., Zhao, N., Yue, T.-X., Zhao, M.-W., & Chen, C. (2015). Change trend of monthly precipitation in China with an improved surface modeling method. *Environmental Earth Sciences*, 74(8), 6459–6469. <https://doi.org/10.1007/s12665-014-4012-0>
- Wang, X., Juma, S., Li, W., Suleman, M., Muhsin, M. A., He, J., He, M., Xu, D., Zhang, J., Bergquist, R., & Yang, K. (2022). Potential risk of colonization of *Bulinus globosus* in the mainland of China under climate change. *Infectious Diseases of Poverty*, 11(1). Scopus. <https://doi.org/10.1186/s40249-022-00980-2>
- Wiwoho, B. S., Astuti, I. S., Alfarizi, I. A. G., & Sucahyo, H. R. (2021). Validation of Three Daily Satellite Rainfall Products in a Humid Tropic Watershed, Brantas, Indonesia: Implications to Land Characteristics and Hydrological Modelling. *Hydrology*, 8(4), Article 4. <https://doi.org/10.3390/hydrology8040154>

- WorldClim. (2022, November 12). Global climate and weather data—WorldClim 1 documentation. WorldClim. https://worldclim.org/data/index.html#google_vignette
- Wu, M., Kwon, T. J., & Fu, L. (2022). Spatial Mapping of Winter Road Surface Conditions via Hybrid Geostatistical Techniques. *Journal of Cold Regions Engineering*, 36(4). Scopus. [https://doi.org/10.1061/\(ASCE\)CR.1943-5495.0000286](https://doi.org/10.1061/(ASCE)CR.1943-5495.0000286)
- Xavier, A. C., King, C. W., & Scanlon, B. R. (2016). Daily gridded meteorological variables in Brazil (1980–2013). *International Journal of Climatology*, 36(6), 2644–2659. <https://doi.org/10.1002/joc.4518>
- Xie, P., Chen, M., Yang, S., Yatagai, A., Hayasaka, T., Fukushima, Y., & Liu, C. (2007). A Gauge-Based Analysis of Daily Precipitation over East Asia. *Journal of Hydrometeorology*, 8(3), 607–626. <https://doi.org/10.1175/JHM583.1>
- Yang, W., Ge, M., Wang, Y., Pang, X., & Wang, C. (2022). Spatial distribution differences of 25-hydroxyvitamin D in healthy elderly people under the influence of geographical environmental factors. *Scientific Reports*, 12(1). Scopus. <https://doi.org/10.1038/s41598-022-17198-9>
- Yue, T., Zhao, N., Fan, Z., Li, J., Chen, C., Lu, Y., Wang, C., Xu, B., & Wilson, J. (2016). CMIP5 downscaling and its uncertainty in China. *Global and Planetary Change*, 146, 30–37. <https://doi.org/10.1016/j.gloplacha.2016.09.003>
- Zhang, Y., Wu, C., Yeh, P. J.-F., Li, J., Hu, B. X., Feng, P., & Jun, C. (2022). Evaluation and comparison of precipitation estimates and hydrologic utility of CHIRPS, TRMM 3B42 V7 and PERSIANN-CDR products in various climate regimes. *Atmospheric Research*, 265, 105881. <https://doi.org/10.1016/j.atmosres.2021.105881>

- Zhao, N., & Jiao, Y. (2021). A New HASM-Based Downscaling Method for High-Resolution Precipitation Estimates. Remote Sensing, 13(14), Article 14. <https://doi.org/10.3390/rs13142693>
- Zhao, N., Yue, T., Zhou, X., Zhao, M., Liu, Y., Du, Z., & Zhang, L. (2017). Statistical downscaling of precipitation using local regression and high accuracy surface modeling method. Theoretical and Applied Climatology, 129(1), 281–292. <https://doi.org/10.1007/s00704-016-1776-z>
- Zorzetto, E., & Marani, M. (2020). Extreme value metastatistical analysis of remotely sensed rainfall in ungauged areas: Spatial downscaling and error modelling. Advances in Water Resources, 135, 103483. <https://doi.org/10.1016/j.advwatres.2019.103483>