

Estudio espacial de los campos de precipitación y temperatura para el departamento de Santander
usando información observacional y satelital o de reanálisis

Kevin Mauricio García Rivera y Daniel Ricardo Moreno Ortiz

Trabajo de Grado para Optar al Título de Ingeniero Civil

Director

Sandra Rocío Villamizar Amaya

Ph.D. in Environmental Systems

Codirector(es)

Kevin David Blanco Silva

Mag. en Recursos Hídricos y Saneamiento

Álvaro Javier Ávila Díaz

Ph.D. in Applied Meteorology

Universidad Industrial de Santander

Facultad de Fisicomecánicas

Escuela de Ingeniería Civil

Ingeniería Civil

Bucaramanga

2024

Dedicatoria

Kevin García

En primer lugar, a Dios por darme fuerzas para superar obstáculos y dificultades a lo largo de toda mi vida; a mi mayor inspiración y la persona a la que más admiro mi adorada madre, Marleny, que siempre me ha brindado su apoyo incondicional para poder cumplir todos mis objetivos personales y académicos.

Daniel Moreno

A toda mi familia por siempre creer en mí, en especial a mi madre Claudia y mi primo Carlos quienes son mi modelo a seguir y me dieron la motivación para no rendirme y lograr todo lo que me propongo.

A mi pareja Alejandra y mi mascota Todd por su apoyo incondicional, gracias por ser la inspiración para finalizar este proyecto.

Agradecimientos

Queremos expresar nuestro más profundo agradecimiento a nuestra directora de tesis Dra. Sandra Rocío Villamizar. Su orientación, comprensión y experiencia fueron claves en el desarrollo de este proyecto; así mismo a nuestros codirectores Mag. Kevin David Blanco Silva y PhD. Álvaro Javier Ávila Diaz quienes, con gran paciencia, dedicación y confianza, nos guiaron de manera exitosa en este complejo y gratificante camino de la investigación.

Adicionalmente, agradecemos a la Universidad Industrial de Santander por abrirnos las puertas y darnos la oportunidad de crecer tanto personalmente como académicamente. Especialmente, por su apoyo constante y confianza en nuestras habilidades para una exitosa finalización de tesis.

Tabla de Contenido

	Pág.
Introducción	14
1. Objetivos.....	17
1.1 Objetivo General.....	17
1.2 Objetivos Específicos.....	17
2. Marco de Referencia.....	18
2.1 Datos satelitales y de reanálisis.....	19
2.2 Reducción de escala (downscaling).....	21
3. Metodología.....	23
3.1 Fase I - Generación de superficies climatológicas de temperatura promedio y precipitación total a escala mensual y anual para las estaciones del IDEAM	23
3.1.1 Descarga y procesamiento de la información obtenida de temperatura y precipitación de estaciones IDEAM para la zona de estudio	23
3.1.1.1 Zona de estudio.....	23
3.1.1.2 Descarga y procesamiento de información.....	25
3.1.1.3 Descarga y selección de estaciones.....	26
3.1.1.4 Imputación de datos faltantes.....	28
3.1.2 Generación de las superficies de precipitación y temperatura con datos observacionales del IDEAM.	30
3.2 Fase II – Generación de superficies climatológicas de temperatura promedio y precipitación total a escala mensual y anual, mediante datos satelitales escalados.....	31

3.2.1 Descarga de los datos de precipitación y temperatura de base de datos satelitales o de reanálisis.	31
3.2.1.1 Descarga y procesamiento de datos CHIRPS.	31
3.2.1.2 Descarga y procesamiento de datos ERA5-Land.....	32
3.2.2 Implementación del método Kriging para la reducción de escala espacial.	33
3.2.3 Retos enfrentados.....	36
3.3 Fase III – Determinación del ajuste de los datos escalados para las estaciones seleccionadas mediante el cálculo de indicadores estadísticos.....	38
4. Resultados.....	40
4.1 Superficies climatológicas de precipitación total (mensual y anual) y temperatura media (mensual y anual) a partir de series de tiempo diarias de estaciones del IDEAM.....	41
4.2 Superficies climatológicas de precipitación total (mensual y anual) y temperatura media (mensual y anual) a partir de bases de datos escaladas.....	45
4.2.1 Bases de datos crudos de las variables meteorológicas	45
4.2.1.1 Información CHIRPS.....	45
4.2.1.2 Información ERA5-Land.	45
4.2.2 Bases de datos escalados de las variables meteorológicas.....	45
El proceso de escalamiento generó datos de precipitación y temperatura máxima y mínima en una resolución de 0.01	45
4.2.3 Superficies climatológicas generadas con datos satelitales escalados	48
4.3 Grado de incertidumbre	52
4.3.1 Precipitación	52
4.3.2 Temperatura máxima	54

4.3.3 Temperatura mínima.....	57
5. Conclusiones.....	60
6. Recomendaciones	62
Referencias Bibliográficas	63

Lista de Tablas

	Pág.
Tabla 1 <i>Selección ventana de tiempo</i>	27

Lista de Figuras

	Pág.
Figura 1 <i>Portal web consulta y descarga de datos hidrometereológicos</i>	24
Figura 2 <i>Estaciones IDEAM localizadas dentro de la zona de estudio</i>	25
Figura 3 <i>Ubicación estaciones de la ventana de tiempo</i>	28
Figura 4 <i>Portal web CHIRPS V2.0</i>	32
Figura 5 <i>Diagrama de flujo de reducción de escala</i>	36
Figura 6 <i>Promedios de precipitación mensual y anual para Santander con datos IDEAM</i>	42
Figura 7 <i>Temperatura máxima mensual y anual para Santander con datos IDEAM</i>	43
Figura 8 <i>Temperatura mínima mensual y anual para Santander con datos IDEAM</i>	44
Figura 9 <i>Precipitación total diaria CHIRPS</i>	46
Figura 10 <i>Temperatura máxima diaria ERA5-Land</i>	47
Figura 11 <i>Temperatura mínima diaria ERA5-Land</i>	47
Figura 12 <i>Precipitación mensual y anual para Santander con datos CHIRPSv2</i>	49
Figura 13 <i>Temperatura máxima mensual y anual para Santander con datos ERA5-Land</i>	50
Figura 14 <i>Temperatura mínima mensual y anual para Santander con datos ERA5-Land</i>	51
Figura 15 <i>Métricas estimadas precipitación CHIRPS</i>	53
Figura 16 <i>Distribución RMSE agosto – precipitación total diaria</i>	54
Figura 17 <i>Métricas estimadas temperatura máxima ERA5-Land</i>	56
Figura 18 <i>Distribución RMSE noviembre – temperatura máxima</i>	57
Figura 19 <i>Métricas estimadas temperatura mínima ERA5-Land</i>	59
Figura 20 <i>Distribución RMSE marzo – temperatura mínima</i>	60

Lista de Apéndices

Los apéndices están adjuntos y puede visualizarlos en la base de datos de la biblioteca UIS

Apéndice A. Glosario

Apéndice B. Base de datos estaciones IDEAM

Apéndice C. Scripts datos IDEAM

Apéndice D. Scripts llenado de datos

Apéndice E. Scripts escalamiento de datos

Apéndice F. Errores y dificultades

Apéndice G. Definición indicadores

Apéndice H. Base de datos CHIRPS

Apéndice I. Base de datos ERA5-LAND

Apéndice J. Datos escalados de precipitación total

Apéndice K. Datos escalados de temperatura máxima

Apéndice L. Datos escalados de temperatura mínima

Apéndice M. Superficies climatológicas de precipitación

Apéndice N. Superficies climatológicas temperatura

Apéndice O. Métricas estadísticas estimadas

Glosario

CHIRPSv2: son un conjunto de datos de precipitación casi global de más de 35 años. CHIRPSv2, que abarca 50°S-50°N (y todas las longitudes), CHIRPSv2 en sus siglas en inglés (Climate Hazards Group InfraRed Precipitation with Station) es una base de datos, empezando reportar información desde 1981 hasta el presente, Abarca tres tipos de información: climatologías globales, estimaciones por satélite y observaciones in situ, se obtiene un mejor producto al alcance del público de forma gratuita (Achá et al., 2022; Anagua & Villazon, 2018).

Clima: Es la integración de las fluctuaciones de los factores atmosféricos con la dinámica de los elementos bióticos, abióticos, físicos y químicos que, considerados en un periodo de tiempo prolongado, le confieren a un determinado territorio geográfico, unas características que definen su carácter ambiental (Antonio, 2018).

Downscaling: Se entiende como la derivación de información de escala local a regional (10-100 kilómetros) a partir de datos modelados u observados a mayor escala. Sus enfoques principales pueden ser reducción de escala dinámica y reducción de escala estadística. Se basa en el supuesto que el clima local es una combinación de características climáticas/atmosféricas a gran escala y condiciones locales (García Abad, 2012; Lala, 2020).

ERA5-Land: Es un conjunto de datos de reanálisis que proporciona una visión consistente de la evolución de las variables terrestres durante varias décadas con una resolución mejorada en comparación con ERA5. ERA5-Land se produjo reproduciendo el componente terrestre del reanálisis climático ERA5 del ECMWF. El reanálisis combina datos del modelo con observaciones de todo el mundo en un conjunto de datos globalmente completo y consistente utilizando las leyes de la física (Copernicus Climate Change Service, 2019a; Hortua et al., 2022).

Hidrología: Es la ciencia que estudia las aguas terrestres, su origen, movimiento y distribución en nuestro planeta, propiedades físicas y químicas, interacción en el medio ambiente físico y biológico e influencia en las actividades humanas (Burbano & Zarama, 1996; IDEAM, 2024).

Kriging: Es un método geoestadístico que supone una correlación espacial entre puntos y tienen la capacidad de proporcionar al modelador una medida de certeza o precisión de las predicciones. Al comparar con otros métodos de interpolación espacial, el Kriging se considera un método robusto, pero de alta demanda computacional ya que utiliza los datos medidos para modelar el variograma y hacer las predicciones (Bernal et al., 2021; Marin Mamani et al., 2021).

Precipitación: Es la caída de partículas de agua líquida o sólida que se originan en una nube, atraviesan la atmósfera y llegan al suelo. La cantidad de precipitación es el volumen de agua lluvia que pasa a través de una superficie en un tiempo determinado. Las formas de precipitación son: lluvia, llovizna, nieve, cinarra, nieve granulada, polvo diamante, granizo y gránulos de hielo (Carrera-Villacrés et al., 2016; IDEAM, 2024).

Reanálisis: Los reanálisis climatológicos o “reconstrucciones” del clima, son el resultado de la combinación de observaciones meteorológicas globales de diferentes fuentes, como pueden ser estaciones convencionales en tierra, marítimas y aéreas, radiosondeos, imágenes satelitales, y datos meteorológicos modelados, obtenidos mediante predicciones numéricas (*NWP, Numerical Weather Predictions*) (Siles Soria, 2022).

Temperatura: Es una magnitud física que expresa el grado de frío o calor de los cuerpos o del ambiente (ASALE & RAE, 2024).

Resumen

Título: Estudio espacial de los campos de precipitación y temperatura para el departamento de Santander usando información observacional y satelital o de reanálisis*

Autor: Kevin Mauricio García Rivera y Daniel Ricardo Moreno Ortiz**

Palabras Clave: CHIRPS, Downscaling, ERA5-Land, Kriging, Precipitación, Temperatura.

Descripción: Las variables climatológicas precipitación y temperatura son fundamentales para estudios aplicados de ingeniería, ambientales, salud pública, su calidad depende de la existencia de una red de estaciones *in-situ* de alta densidad y mantenimiento frecuente. Alternativamente, existen bases de datos satelitales y de reanálisis que pueden suplir la necesidad de este tipo de datos en zonas con poca densidad de estaciones o con disponibilidad de datos para ventanas de tiempo insuficientes; su uso para estudios regionales o locales requiere un proceso de escalamiento (*downscaling*). Este estudio utilizó datos CHIRPSv2 (precipitación) y ERA5-Land (temperatura máxima y mínima) en la ventana de tiempo 1990-2010 para generar una base de datos diarios en una resolución adecuada para estudios regionales a locales (0.01°, ~1000 m) en el departamento de Santander (Colombia). Los datos descargados fueron escalados usando el método Kriging universal con el modelo digital de elevación de la zona de estudio como covariante. Con esta información se generaron superficies que muestran la distribución espacial de los promedios anuales y mensuales de precipitación y temperatura en el Departamento, evidenciando una clara concordancia con los promedios publicados por el IDEAM, pero con la ventaja de una mayor resolución espacial. Se evaluó el desempeño de los datos escalados realizando una comparación punto-píxel con los datos observacionales de las estaciones IDEAM. En general, se obtuvieron errores aceptables, aunque fueron mucho mayores para la precipitación. Adicionalmente, para esta variable los errores fueron mayores durante los meses más lluviosos del año (marzo-mayo y septiembre-noviembre). Para el caso de la temperatura, las métricas estadísticas son mucho mejores al hacer la comparación punto-píxel entre datos escalados y observados vs. datos crudos y observados. Para la precipitación las métricas de error son similares en los dos casos lo cual es un resultado satisfactorio teniendo en cuenta que se obtiene una mejor resolución y el grado de incertidumbre se mantiene. Se incluye en los apéndices de este trabajo las bases de datos consolidadas, las superficies generadas y los códigos implementados de tal manera que es factible reproducir el trabajo realizado o utilizar los recursos para generar datos similares en otras zonas de interés.

* Trabajo de Grado

** Facultad de Fisicomecánicas. Escuela de Ingeniería Civil. Ingeniería Civil. Director: Sandra Rocío Villamizar. PhD in Environmental Systems. Codirectores: Álvaro Javier Ávila Díaz. PhD in Applied Meteorology. Kevin David Blanco Silva. Mag. en Recursos Hídricos y Saneamiento.

Abstract

Title: Spatial study of precipitation and temperature fields for the department of Santander using observational and satellite or reanalysis data. *

Author(s): Kevin Mauricio García Rivera and Daniel Ricardo Moreno Ortiz **

Key Words: Precipitation, Temperature, CHIRPS, ERA5-Land, Downscaling, Kriging.

Description: Climatological variables such as precipitation and temperature are fundamental for applied engineering, environmental and public health studies, and their quality depends on the existence of a high density and frequently maintained network of in-situ stations. Alternatively, there are satellite and reanalysis databases that can supply the need for this type of data in areas with low station density or with data availability for insufficient time windows; their use for regional or local studies requires a downscaling process. This study used CHIRPSv2 (precipitation) and ERA5-Land (maximum and minimum temperature) data in the 1990-2010 time window to generate a daily database at a resolution suitable for regional to local studies (0.01° , ~ 1000 m) in the department of Santander (Colombia). The downloaded data were scaled using the universal Kriging method with the digital elevation model of the study area as covariant. With this information, surfaces were generated showing the spatial distribution of the annual and monthly averages of precipitation and temperature in the Department, showing a clear agreement with the averages published by IDEAM, but with the advantage of a higher spatial resolution. The performance of the scaled data was evaluated by performing a point-pixel comparison with observational data from IDEAM stations. In general, acceptable errors were obtained, although they were much higher for precipitation. Additionally, for this variable, errors were higher during the rainiest months of the year (March-May and September-November). In the case of temperature, the statistical metrics are much better when making the point-pixel comparison between scaled and observed data vs. raw and observed data. For precipitation the error metrics are similar in both cases which is a satisfactory result considering that a better resolution is obtained, and the degree of uncertainty is maintained. The appendices of this work include the consolidated databases, the generated surfaces, and the implemented codes in such a way that it is feasible to reproduce the work done or to use the resources to generate similar data in other areas of interest.

* Degree Work

** School of Physicomechanics. School of Civil Engineering. Civil Engineering. Director: Sandra Rocío Villamizar. PhD in Environmental Systems. Co-Directors: Álvaro Javier Ávila Díaz. PhD in Applied Meteorology. Kevin David Blanco Silva. Mag. in Water Resources and Sanitation.

Introducción

Las variables precipitación y temperatura son un insumo primario de estudios de clima (Eguiguren-Velepucha et al., 2016; Fan & van den Dool, 2008a), ecológicos (Meier & Leuschner, 2008; Z. Wu et al., 2011), ambientales (Punia, 2021; Yan & Baas, 2015), de salud pública (Blanco, 2022; Li et al., 2020; Menebo, 2020; Park et al., 2015; Tamerius et al., 2016) e hidrológicos (Miller et al., 2003; Vano et al., 2012), entre otros. En la ingeniería civil, las aplicaciones de esta información son diversas; por ejemplo, sirven para generar curvas de intensidad-duración-frecuencia (IDF), claves para el dimensionamiento de estructuras (Cheng & AghaKouchak, 2014; Sun et al., 2019); permiten comprender problemas relacionados con el tránsito y transporte (Liu et al., 2023); permiten informar acciones relacionadas con la gestión de la construcción (Jarkas & Bitar, 2012); y juegan un papel crítico para los estudios de gestión del riesgo de desastres (Jiménez-Ramos et al., 2023; Masello et al., 2023).

En Colombia, el monitoreo de las variables meteorológicas está a cargo del Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM). Esta entidad ofrece mapas de la distribución espacial de precipitación y temperatura (promedios mensuales y anuales) usando datos climatológicos relevantes para el período 1981-2010 (IDEAM, 2024) aunque recientemente la entidad publicó los datos de variables climatológicas para el período 1991-2020. Sin embargo, las superficies generadas tienen una escala media (desde 1:500.000 hasta 1:100.000) que no permite apreciar las variaciones locales de relevancia para estudios específicos. Esto es especialmente problemático teniendo en cuenta que Colombia presenta una variabilidad climática regida por la zona de confluencia intertropical, con una fuerte dinámica de los océanos pacífico y atlántico, junto con las cuencas del Amazonas y el Orinoco (Urrea et al., 2019). Adicionalmente, la

climatología del departamento de Santander, que es el área de interés de este trabajo, está determinada por un gradiente topográfico fuerte que se extiende entre menos de 100 m.s.n.m. en su parte occidental hasta los 4500 m.s.n.m. en su parte oriental (TessaDem & OpenStreetMap, 2023) lo que sugiere que existe una variación importante en los patrones espaciales de precipitación y temperatura.

Para Santander, se tiene registro de 299 estaciones instaladas por el IDEAM. Se cuenta con información de precipitación en 254 estaciones y de temperatura en 64, es decir, aproximadamente, se tiene una estación de precipitación por cada 120 km² y una estación de temperatura por cada 954 km² de territorio en el departamento. Aunque estas densidades calculadas pueden parecer altas (Gubler et al., 2017), la red tiene problemas de continuidad en el tiempo haciendo que, en realidad, a la fecha, se cuente con 93 estaciones activas de precipitación y 28 estaciones activas de temperatura (densidades reales de 0.0032 estaciones/km² y 0.00055 estaciones/km², respectivamente). La falta de información de alta resolución en fuentes oficiales hace que los estudios específicos en el Departamento obtengan información de diversas fuentes y hagan sus propios procesamientos de información para generar estas superficies y otra información requerida. Esto hace que se genere una incertidumbre que se escala a los resultados obtenidos en dichos proyectos y dificulta la integración de resultados entre estudios realizados en la misma zona por diferentes autores.

Los datos satelitales y de reanálisis se constituyen en una alternativa valiosa ya que proveen información continua para ventanas de tiempo climatológicamente significativas para cualquier lugar del planeta (usando la base de datos adecuada). Sin embargo, requieren de un proceso de escalamiento (*downscaling*) para lograr alcanzar resoluciones espaciales relevantes para estudios locales y regionales (Sun et al., 2019). Recientemente, (Blanco, 2022) aplicó un método

geoestadístico de escalamiento (Kusch & Davy, 2022a) para generar datos de temperatura y precipitación de alta resolución espacial, necesarios para un estudio de salud pública en el departamento del Cauca.

Partiendo de esta experiencia previa, se planteó la generación de una base de datos climatológicamente robusta para Santander que sirva para el desarrollo de diferentes proyectos de investigación y consultoría que se lleven a cabo en el Departamento. Mediante el método de escalamiento geoestadístico Kriging, este estudio generó series de tiempo diarias de precipitación y temperatura con una resolución espacial de 0.01 grados (1 km aproximadamente). Con los datos diarios se generaron superficies mensuales y anuales de valores promedio de precipitación y temperatura, permitiendo tener una mejor comprensión de la distribución espacial de estas variables en el Departamento de Santander. Los resultados de este proyecto permitirán enriquecer y consolidar el uso de datos de alta resolución espacial y continuidad temporal en otras regiones de Colombia, fundamentalmente, en aquellas donde el sistema de datos in-situ sea insuficiente.^{††}

^{††} El Apéndice A de este documento presenta un glosario con términos relevantes para la comprensión del trabajo.

1. Objetivos

1.1 Objetivo General

Comparar espacialmente los campos de precipitación y temperatura para el departamento de Santander usando información observacional y satelital o de reanálisis.

1.2 Objetivos Específicos

Generar superficies climatológicas de precipitación total (mensual y anual) y temperatura media (mensual y anual) a partir de series de tiempo diarias de estaciones del IDEAM que tengan una disponibilidad temporal climatológicamente relevante.

Generar superficies climatológicas de precipitación total (mensual y anual) y temperatura media (mensual y anual) de alta resolución espacial, a partir de datos satelitales o de reanálisis escalados por un método estadístico.

Determinar el ajuste entre las superficies generadas mediante la comparación punto-píxel a partir del cálculo de métricas estadísticas de desempeño.

2. Marco de Referencia

La climatología de un área de estudio se puede caracterizar por la distribución espacial de variables clave como la temperatura, la presión barométrica, la velocidad del viento y la precipitación. Esta distribución espacial parte de puntos de información conocida de la variable de interés y, tradicionalmente, la generación de isopletas o líneas de igual valor de dicha variable para un período de tiempo determinado (Jarraud & Bokova, 2012). Más recientemente, con el avance de las técnicas de interpolación espacial, es posible generar superficies que ofrecen un valor de la variable de interés en el espacio continuo (Berndt & Haberlandt, 2018; Hadi & Tombul, 2018; Ozturk & Kilic, 2016).

Entre los diferentes métodos estadísticos para este tipo de aplicaciones se pueden destacar el “gridpoint statistical interpolation” (GSI) y el “spectral nudging” (SN), los cuales, por ejemplo, fueron usados para evaluar dos conjuntos de datos de reanálisis regional chino (CNRR) de alta resolución (18 km) entre 1998 y 2009, (Lu et al., 2021). Así mismo, otros tres métodos de interpolación pueden ser; distancia inversa ponderada, Spline y Kriging, a los cuales, por ejemplo, se les evaluó su rendimiento en un Sistema de Información Geográfica (SIG) para estudiar la variabilidad espacial de las variables climáticas (temperatura máxima del aire, temperatura mínima del aire y precipitaciones estacionales y anuales) en Nebraska. Adicionalmente, es importante recordar que no existe un único método de interpolación preferido y la selección del método de interpolación generalmente se basa en los datos disponibles, el nivel deseado de precisión y los recursos disponibles (Berndt & Haberlandt, 2018).

La calidad de estas superficies interpoladas depende de una red de estaciones robusta. Teniendo en cuenta que el montaje y mantenimiento de una red meteorológica robusta en un país

como Colombia es inviable desde los aspectos técnicos, económicos y operativos, surge la alternativa de usar datos satelitales y de reanálisis que permitan solventar los problemas de continuidad temporal y espacial de información. Estos datos ofrecen una continuidad temporal significativa climatológicamente en una grilla regular, que debe ser escalada (*downscaled*) para que pueda ser usada en estudios locales. Las subsecciones siguientes ofrecen información sobre los datos satelitales y de reanálisis, y del proceso de escalamiento de datos climatológicos para mejorar su resolución espacial.

2.1 Datos satelitales y de reanálisis

Existen datos meteorológicos generados por sensores remotos y de reanálisis que complementan la información in-situ, permitiendo la obtención de información para la generación de las superficies climatológicas. Entre ellas se destacan CHIRPS v2.0, con una resolución espacial de 0.05° (~4500m) y temporal diaria (desde 1981) (ERDDAP, 2023); MSWX (Multi-Source Weather), con datos de 0.1° (~9000m) cada 3 horas (MSWX-GloH20, 2021); NOAA GHCN_CAMS, con datos de alta resolución 0.5° (~45000m) cada mes desde 1948 hasta el presente (NOAA, 2021); y ERA5-Land, producto de reanálisis con alta resolución (0.1° , ~9000m) cada hora y evolución de variables meteorológicas durante varias décadas, entre otros (Beck et al., 2022; Blanco, 2022; Copernicus Climate Change Service, 2019a; Fan & van den Dool, 2008b; Funk et al., 2015).

Este tipo de información complementa las mediciones de superficie tradicionales y ofrece una cobertura sin precedentes a escala mundial. Por ejemplo, en el estado de Santa Catarina (Brasil) se usaron datos de reanálisis climático y teledetección para desarrollar un modelado predictivo en apoyo al programa de vigilancia de floraciones de algas nocivas (Vianna et al.,

2023). Por otro lado, en la cuenca del río Kilombero (Tanzania) se presentaron conflictos socioeconómicos por un conocimiento inadecuado en la disponibilidad real del agua; por tal motivo, se analizaron diversas variables climatológicas, y con datos satelitales se mejoró la comprensión espacio-temporal de la zona (Sigalla et al., 2023). Así mismo, en Marruecos, país con un terreno complejo y regiones escasamente pobladas, debido a que su clima es altamente vulnerable a inundaciones por las altas precipitaciones, la disponibilidad limitada de datos fue el principal obstáculo a lo que se recurrió a la evaluación de 9 conjuntos de datos basados en información satelital y de reanálisis para una comprensión más integral de este fenómeno (Tuel & El Moçayd, 2023). Aunque para el caso de la precipitación se ha evidenciado que la información satelital no es mejor que la información *in-situ*, se encontró que la base de datos CHIRPS de alta resolución puede ser una alternativa promisoría en sitios con escasa información (Duan et al., 2019).

En Colombia, la base de datos CHIRPS v2.0 para el caso de información diaria de precipitación ha sido evaluada en el departamento de Antioquia, permitiendo validar un buen funcionamiento en la mayoría de las condiciones climáticas de este departamento, incluyendo los periodos más intensos del ENOS (López-Bermeo et al., 2022). Sin embargo, se encontró que su desempeño fue mejor en la región Andina que en la región Pacífico de nuestro país (Ocampo-Marulanda et al., 2022). Más recientemente se encontró que la base de datos CHIRPS v2.0 tuvo una mejor ejecución que otros productos de precipitación (CMORPHv1, ERA5, ERA5-Land, IMERG, MSWEPv2.8, PERSIANN-CCS, y PERSIANN-CDR) para todas las elevaciones y escalas de tiempo analizadas. Sin embargo, la mejor ejecución de este tipo de información se dio en la escala de tiempo mensual y en las zonas de bajas elevaciones del país (menor que 1000 m.s.n.m.) (Valencia et al., 2023).

2.2 Reducción de escala (downscaling)

El *downscaling* consiste en el proceso de reducción de escala mediante el cual se obtienen datos de alta resolución a partir de un producto de resolución menor (Río San José, 2010). Las metodologías de reducción de escala se clasifican en estadísticas y computacionales (o de *machine learning*). Los diversos métodos estadísticos (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) en general implican procesos de interpolación y extrapolación permitiendo la construcción de cuadrículas a partir de estaciones locales (Blanco, 2022); Sin embargo, se debe resaltar que los resultados son altamente influenciados por la densidad de la red de observación y las escalas tanto espacial como temporal en que se trabaja (Serrano-Notivoli & Tejedor, 2021). Por ejemplo, se realizó un estudio que divide a China en ocho zonas de vegetación basadas en los diferentes tipos de esta. Dentro de cada zona de vegetación, se empleó un enfoque de árboles aleatorios extremos ajustado estacionalmente para reducir espacialmente los datos de precipitación mensual del satélite GPM (medición de precipitación global) y para validar la efectividad de este método, lo compararon con la interpolación Kriging y los métodos tradicionales de reducción de escala global. Los resultados indican que el método de predicción de área dividida supera a otros enfoques, lo que da como resultado un conjunto de datos de precipitación que no solo logra una alta precisión, sino que también ofrece una resolución espacial más fina en comparación con el conjunto de datos de precipitación GPM original.

Por otro lado, los métodos computacionales 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. Aunque existe una tendencia creciente en su uso, puesto que se mejora

considerablemente el escalamiento, estos procesos requieren un alto consumo de recursos informáticos y alta experticia (Khashei & Bijari, 2010).

Aunque existe incertidumbre en la eficiencia del escalamiento, especialmente para la precipitación (Dibike et al., 2008; Lanzante et al., 2018; Souvignet & Heinrich, 2011; H. Wu et al., 2016), la reducción de escala ha permitido avanzar en la comprensión de procesos a escalas regional y local (Blanco, 2022; F. Chen et al., 2020; Vargas Godoy & Markonis, 2023). Recientemente, se están usando los datos escalados para evaluar el impacto del cambio climático en el planeta (Araya-Osses et al., 2020; Hundedcha et al., 2016, 2016; Sunyer et al., 2015; Werner & Cannon, 2016).

En Colombia se han usado datos escalados para diversos estudios. Para la cuenca del río Magdalena, se desarrolló un nuevo procedimiento de reducción de escala espacial de precipitación y se aplicó a cuatro conjuntos de datos de precipitación, teniendo como objetivo analizar qué tan sensibles son las simulaciones de caudal a la resolución espacial, encontrando que el modelo GWR (Geographically Weighted Regression) y la pendiente tienen una mínima contribución en la reducción de escala a diferencia de los predictores de vegetación y elevación (López López et al., 2018). En el valle de la cordillera central de los andes colombianos, se utilizó un método de reducción de escala dinámico para mejorar la representación de la precipitación en la base de datos ERA-Interim, encontrando que el método WRF (Weather Research and Forecasting) reduce la sobreestimación de la precipitación en la costa pacífica (Posada-Marín et al., 2019). Recientemente, en el departamento del Cauca se usó la metodología Kriging para hacer un escalamiento de datos de precipitación y temperatura que permitieron evaluar la incidencia del dengue en esta región. Como en la mayoría de los estudios, se evidenció que el ajuste del

escalamiento fue mejor para los datos de temperatura que para los datos de precipitación (Blanco, 2022).

3. Metodología

La comparación de las superficies climatológicas escaladas a partir de datos satelitales frente a las generadas con los datos observacionales obtenidos con medición *in-situ* de las estaciones del IDEAM se llevó a cabo mediante tres fases: i) Generación de superficies climatológicas de temperatura promedio y precipitación total a escala mensual y anual, para las estaciones del IDEAM con extensión temporal climatológicamente significativa; ii) Generación de superficies climatológicas de temperatura promedio y precipitación total a escala mensual y anual, mediante datos satelitales escalados con extensión temporal climatológicamente significativa; iii) Determinación del ajuste de los datos escalados para las estaciones seleccionadas mediante el cálculo de indicadores estadísticos. Las siguientes subsecciones describen la estrategia metodológica para el desarrollo de cada una de las fases.

3.1 Fase I - Generación de superficies climatológicas de temperatura promedio y precipitación total a escala mensual y anual para las estaciones del IDEAM

3.1.1 Descarga y procesamiento de la información obtenida de temperatura y precipitación de estaciones IDEAM para la zona de estudio

3.1.1.1 Zona de estudio. La zona de estudio está delimitada por las coordenadas 5.7° a 8.1° latitud norte y -74.5° a -72.5° longitud oeste, que enmarcan en su totalidad el departamento de Santander. Se utilizó la plataforma DHIME (figura 1) del IDEAM (dhime, 2024) para

identificar las estaciones existentes dentro del recuadro de la zona de estudio. Dentro de esta área el IDEAM reporta 218 estaciones pluviométricas, 21 estaciones pluviográficas, 46 estaciones climatológicas ordinarias, 43 estaciones climatológicas principales, y 150 estaciones limnimétricas, entre otras (Figura 2). El apéndice B, presenta la tabla principal con todas las estaciones existentes en el área de estudio, mostrando elementos descriptivos reportados por la base de datos DHIME.

Figura 1

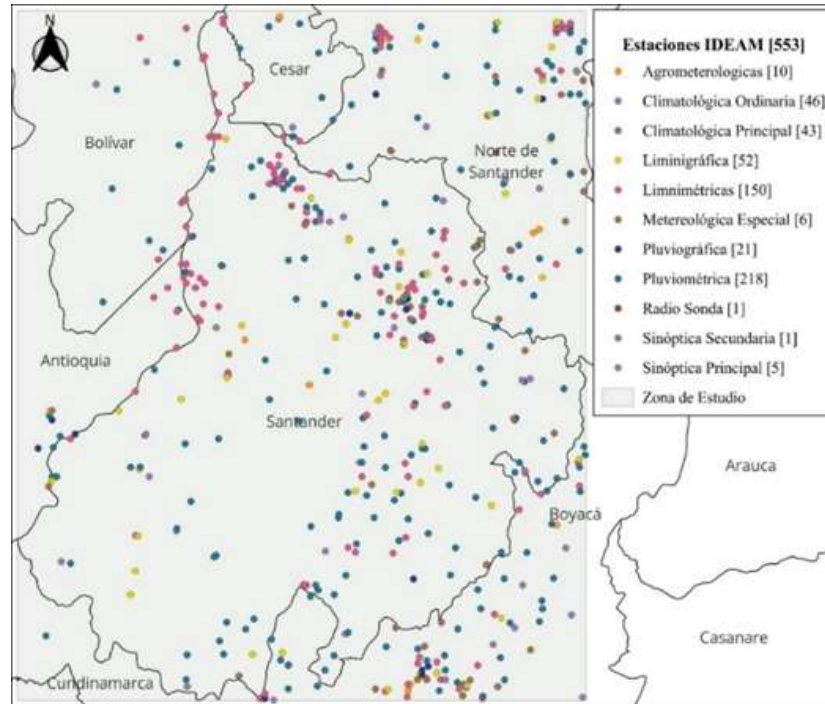
Portal web consulta y descarga de datos hidrometeorológicos



Nota. El gráfico representa la plataforma HDIME; el portal web Consulta y descarga de datos hidrometeorológicos del IDEAM. Tomado de (dhime, 2024).

Figura 2

Estaciones IDEAM localizadas dentro de la zona de estudio



Nota. El gráfico representa todas las estaciones IDEAM existentes en la zona de estudio, pero solo las climatológicas, pluviográficas y pluviométricas fueron usadas en el proyecto.

3.1.1.2 Descarga y procesamiento de información. La descarga de datos de precipitación y temperatura a través del portal DHIME del IDEAM consideró la ventana de tiempo desde la fecha de instalación de las estaciones (algunas desde hace más de 70 años) hasta el 01/01/2023. La información fue descargada del portal en formato “.csv”, en grupos de 10 estaciones para las variables temperatura máxima, temperatura mínima, temperatura promedio y precipitación total diaria. Dichos archivos fueron procesados mediante el lenguaje de programación R, utilizando la interfaz RStudio, inicialmente, para descomprimir y separar la información de cada estación en archivos “.csv” (Código disponible en el Apéndice C).

Se descargaron datos de 430 estaciones de precipitación y 65 estaciones de temperatura máxima y mínima. Se usó una versión modificada de la librería “IdeamR” desarrollada por el ingeniero Geomar Perales para la gestión de datos de las estaciones del IDEAM (Apaico, 2021/2023). El código adaptado se usó para analizar los atributos más importantes de cada estación como su identificación, ubicación, ventana temporal, disponibilidad de datos, y el valor mínimo, máximo y promedio anual; esto permitió generar tablas resumen (Apéndice B) con la información clave de todas las estaciones. El código modificado está disponible en el Apéndice C.

3.1.1.3 Descarga y selección de estaciones. Debido a que los datos IDEAM son el patrón para estimar el rendimiento de las bases de datos satelitales originales y escaladas, se requirió establecer un criterio de selección de estaciones que cumplieran con un porcentaje de disponibilidad de datos representativo y una ventana temporal mayor a 20 años. Para esto se plantearon distintos criterios de clasificación tales como el 50%, 60%, 70% y 80% de disponibilidad de datos y una ventana temporal mayor a 20 o 30 años. El análisis de las diferentes combinaciones permitió definir que se trabajaría con estaciones con porcentaje de disponibilidad de datos mayor al 80% y con más de 20 años de registro (Tabla 1). Esta decisión buscó que los resultados del estudio fueran climatológicamente significativos y que la calidad de los datos se afectara mínimamente por procesos de imputación de datos faltantes.

Tabla 1*Selección ventana de tiempo*

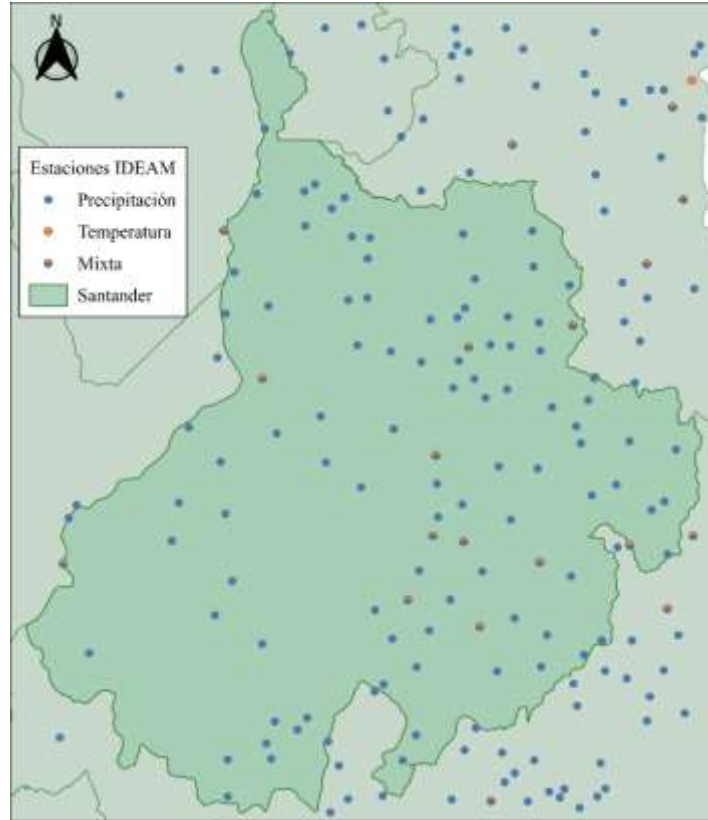
Rango de Fechas		Número de Estaciones con +20años y >80%		
Inicial	Final	Temperatura Min	Temperatura Max	Precipitación
1/01/1984	1/01/2004	19	19	165
1/01/1986	1/01/2006	20	20	180
1/01/1987	1/01/2007	21	20	186
1/01/1990	1/01/2010	21	20	186

Nota. Esta tabla muestra la cantidad de estaciones habilitadas dependiendo de la ventana de tiempo escogida con los criterios de >20 años reportados y >80% de datos disponibles.

Se cotejaron los periodos de tiempo de las estaciones seleccionadas tras el filtro de porcentaje de disponibilidad de datos y número de años disponibles mediante un diagrama de Gantt en busca de una ventana de tiempo de al menos 20 años que permitiera acoger la mayor cantidad de estaciones. En este proceso se lograron identificar 4 intervalos tentativos 1984-2003, 1986-2005, 1987-2006 y 1990-2009. Se escogió la ventana de tiempo más actual (1990-2009) que, adicionalmente, contó con la mayor cantidad de estaciones. El Apéndice B muestra el análisis realizado para la selección de la ventana de tiempo y la Figura 3 muestra las estaciones seleccionadas para el estudio de acuerdo con los criterios de disponibilidad, extensión y ventana de tiempo seleccionada.

Figura 3

Ubicación estaciones de la ventana de tiempo



Nota. El gráfico representa todas las estaciones IDEAM existentes en la zona de estudio, pero solo las climatológicas, pluviográficas y pluviométricas fueron usadas en el proyecto.

3.1.1.4 Imputación de datos faltantes. Para la correcta imputación de los valores faltantes en las series de tiempo de las estaciones seleccionadas, primero se realizó un análisis mediante un código en R que permite contar el número de días consecutivos con datos faltantes dentro de la serie de tiempo estudiada. Esto permitió realizar una clasificación del estado de los vacíos y así identificar las estaciones que tuvieran una gran cantidad de días consecutivos con datos faltantes, dado que la imputación de estos datos resulta inviable.

Las series de tiempo de temperatura máxima y mínima se imputaron de manera automática mediante el método Spline únicamente para vacíos entre 1 y 5 días consecutivos. Para las series de precipitación se propuso una imputación automática mediante el método Spline para los vacíos de 1 día y una imputación manual para los vacíos de 2 a 5 días consecutivos.

Para determinar el método de imputación automático a utilizar se planteó realizar la imputación de una estación que no presentara vacíos, pero eliminando algunos datos conocidos para así poder medir la afinidad de los valores a imputar mediante tres métodos distintos: Media Móvil, Interpolación Spline y Relleno hacia adelante. Los resultados de las diferentes imputaciones fueron comparadas frente a la serie original mediante el error de raíz cuadrada media (RMSE) y el error porcentual absoluto medio (MAPE), siendo los valores del RMSE muy similares para los tres métodos, mientras que los valores del MAPE demostraron que la técnica de interpolación Spline funciona mucho mejor para la imputación de datos de precipitación y temperatura. Este método se basa en la interpolación de funciones polinómicas por tramos (Apéndice D). Vale la pena aclarar que, debido a las propiedades de los polinomios y las series de tiempo de precipitación diaria, es posible llegar a obtener valores negativos en la imputación, lo cual es físicamente imposible, se requirió delimitar los valores negativos a ceros dentro de las series de precipitación diaria imputadas.

La imputación manual de las estaciones de precipitación se realizó mediante el método de interpolación de distancia inversa ponderada (IDW). Para esto se realizó una clasificación entre las estaciones disponibles para seleccionar las que servirán como auxiliares en la estimación de los datos; las estaciones seleccionadas fueron aquellas cuyas series anuales de precipitación total no tuvieran una diferencia mayor al 10%, además de que se encontraran en un rango de más o menos 1000 metros de altitud de diferencia. Para estas estaciones se calculó la precipitación acumulada

promedio para los rangos de fechas a imputar ponderadas mediante la distancia inversa entre las coordenadas de las estaciones comparadas. Los códigos relacionados con los análisis descritos están disponibles en el Apéndice D.

Finalmente, el escalamiento temporal de las variables precipitación y temperatura, a nivel mensual y anual se realizó a partir de las siguientes condiciones: para el cálculo de la temperatura máxima y mínima promedio mensual solo se tuvieron en cuenta los meses con disponibilidad de más de 25 días (~80% de los días del mes); para el cálculo de las temperaturas promedio anuales solo se usaron estaciones con reportes de promedios mensuales completos, es decir, todos los meses disponibles con más de 25 días con datos; para el cálculo de totales mensuales y anuales de precipitación solo se tuvieron en cuenta los meses y años con la totalidad de los datos registrados para dicho periodo.

3.1.2 Generación de las superficies de precipitación y temperatura con datos observacionales del IDEAM.

Para la generación de superficies con datos observacionales del IDEAM se utilizó el software QGIS haciendo uso de la herramienta “Interpolación IDW”. Es importante resaltar que las superficies se generaron usando los datos mensuales y anuales calculados según el procesamiento indicado en la sección anterior, y a partir de las estaciones disponibles en el área de estudio clasificadas por los procesos de filtrado y selección descritos anteriormente.

3.2 Fase II – Generación de superficies climatológicas de temperatura promedio y precipitación total a escala mensual y anual, mediante datos satelitales escalados

Estudios previos realizados por (Blanco, 2022) y (Valencia et al., 2023) compararon, mediante medidores estadísticos, la afinidad de distintas bases de datos satelitales como CHIRPS, CMORPH, ERA5, ERA5-LAND, MSWX, IMERG, con datos observacionales tradicionales. Estos estudios concluyeron que las bases de datos que han presentado mejores resultados han sido CHIRPSv2 para precipitación y ERA5-LAND para temperatura. Estos resultados son consistentes con lo que sugiere la literatura; CHIRPSv2 tiene un buen desempeño en la región (López-Bermeo et al., 2022) y ERA5-LAND, junto con los productos de Copernicus EU, son referentes internacionales en cuanto este tipo de información y de usual comparación (Lu et al., 2021).

3.2.1 Descarga de los datos de precipitación y temperatura de base de datos satelitales o de reanálisis.

3.2.1.1 Descarga y procesamiento de datos CHIRPS. Mediante la plataforma online del National Oceanic and Atmospheric Administration (NOAA) se accedió a los datos globales de precipitación con resolución espacial de 0.05° (~4500 m) y se descargaron mediante la definición del área geográfica de interés (Copernicus Climate Change Service, 2019b). La ventana de tiempo de descarga fue la misma que se definió para los datos de las estaciones IDEAM (1990 a 2009) y el formato de descarga del archivo fue tipo netCDF (“.nc”). Este es un formato multidimensional que, para el caso de los datos CHIRPS, ofrece una serie de tiempo de precipitación para cada centro de píxel de la cuadrícula definida por las coordenadas de interés (Figura 4). Para este caso, los datos descargados no requirieron de algún procesamiento adicional previo al proceso de escalamiento.

Figura 4

Portal web CHIRPS V2.0

The screenshot shows the ERDDAP Data Access Form for CHIRPS Version 2.0. The form is titled "ERDDAP > griddap > Data Access Form" and displays the following information:

- Dataset Title:** CHIRPS Version 2.0, Precipitation, Global, 0.05°, Daily, 1981-present
- Information:** UCMB Climate Hazards Group (Dataset ID: chrs200GlobalDailyP05), Summary, Links, FIDOC (ISO 19115), Metadata, Background #, Files, Make a graph.
- Dimensions:**
 - time (UTC):** Start: 1980-01-01T00:00:00Z, Stop: 2009-12-31T00:00:00Z, Size: 15,706, Spacing: 1 day (event)
 - latitude (degrees_north):** Start: -5.700, Stop: 8.140, Size: 2000, Spacing: 0.05 (nearest)
 - longitude (degrees_east):** Start: -74.525, Stop: -72.477, Size: 7200, Spacing: 0.05 (nearest)
- Grid Variables:** chrs2p (Precipitation, mm/day) [Checked]
- File type:** nc (Download a NetCDF-3 binary file with COARDS/GFACD0 metadata). Just generate the URL: <https://coastwatch.pfeg.noaa.gov/erddap/griddap/chrs200GlobalDailyP05>
- Submit:** (Please be patient. It may take a while to get the data.)
- The Dataset Attribute Structure (.das) for this Dataset:**

```

attributes {
  time {
    string _SourceAttributeType "Time";
    float# actual_range 3.471532e+5, 1.103940e+6;
    string axis "T";
    string calendar "gregorian";
    string comment "This is the time at the start of the day."
  }
}

```

Nota. El gráfico representa la interfaz del portal web de descarga de datos CHIRPS. Tomado de (ERDDAP, 2023).

3.2.1.2 Descarga y procesamiento de datos ERA5-Land. Para acceder a esta base de datos fue necesaria 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”. El proceso de descarga se realizó mediante la función “download_ERA”, de la librería “KrigR” de R [24], a través de un código en R; en este se 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 E presenta el script empleado para la descarga de las variables.

3.2.2 Implementación del método Kriging para la reducción de escala espacial.

El Kriging es un método geoestadístico que permite estimar algunos valores no registrados de una variable dentro de un área determinada usando información existente dentro de esta. Para la estimación se pueden usar funciones de apoyo tal como medias, funciones de covarianza y variogramas. Existen tres distintas metodologías de Kriging lineal: Kriging Simple, que se basa en el uso de una media constante conocida y semivariogramas sin ajustes; Kriging ordinario, que se ajusta a las funciones de los semivariogramas; y el Kriging universal, que consiste en el ajuste de los datos a un modelo de regresión global mediante el uso de variogramas (Blanco, 2022; Chilès & Delfiner, 2012).

Para el desarrollo del proceso de escalamiento de las variables climatológicas del proyecto se parte del uso del sistema de ecuaciones del Kriging universal.

$$\sum_{\beta} \lambda_{\beta} \gamma_{\alpha\beta} + \sum_l \mu_l f_{\alpha}^l = \gamma_{\alpha 0} \quad \alpha = 1, 2, \dots, N$$

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

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)

Los variogramas se pueden calcular como “la mitad de la esperanza matemática del cuadrado de la diferencia de dos valores de una 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).

Una de las principales aplicaciones del Kriging Universal es el escalamiento espacial de variables meteorológicas mediante archivos covariantes, usualmente de elevación, convirtiéndolas en una fuente de información climatológica de alta resolución para diversos estudios. Por ejemplo, se ha utilizado para realizar estudios de salud pública en el departamento del Cauca, Colombia mediante el escalamiento de variables climatológicas como precipitación y temperatura (Blanco, 2022); construir mapas de temperatura mínima para la evaluación de la posibilidad teórica de distribución de *Bulinus globosus* en China (Wang 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 (M. Wu et al., 2022); analizar la distribución espacial de los niveles freáticos en una región de Turquía (Gundogdu & Guney, 2007); y 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).

La implementación del método geoestadístico Kriging Universal para 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), la cual requiere el uso de datos de entrenamiento

de las variables a escalar y datos de archivos covariantes disponibles tanto en la resolución de los datos satelitales como en la resolución a la que se quiere escalar los datos. Información a partir de la cual la herramienta ajusta variogramas a los datos de entrenamiento y establece funciones de covarianza con los datos de las covariantes con resolución original para así generar funciones que describen la correlación espacial entre los datos de entrenamiento y los covariantes (Kusch & Davy, 2022a); posteriormente, se realiza la predicción de los valores de la variable requerida en las nuevas ubicaciones, a partir de los datos de las covariantes en la resolución objetivo (Kusch & Davy, 2022b).

Para realizar el escalamiento de las bases de datos CHIRPSv2 para la variable de precipitación diaria y ERA5-Land para temperatura máxima y mínima, se utilizó el modelo digital de elevación (MDE) con la resolución de los datos originales y resolución de 0.01° (~1km) del Servicio Geológico de los Estados Unidos (USGS) como archivo covariante para la reducción de escala, ya que este fue validado por (Kusch & Davy, 2022b) y (Blanco, 2022). El código está disponible en el Apéndice E.

La extracción de las series de tiempo para los archivos netCDF escalados para cada una de las variables estudiadas en las coordenadas correspondientes a las estaciones IDEAM seleccionadas para la comparación se realizó mediante el lenguaje de programación NCAR Command Language (NCL) para la manipulación de datos climatológicos que corre sobre Ubuntu. El código está disponible en el Apéndice E.

Figura 5

Diagrama de flujo de reducción de escala



Nota. El gráfico representa un diagrama de flujo sobre el proceso de reducción de escala de manera más clara y evidente.

3.2.3 Retos enfrentados

El escalamiento de los datos implicó retos técnicos y computacionales que tuvieron que ser superados. En primer lugar, el grado de requisitos computacionales para realizar el análisis y procesamiento de datos expuesto en la metodología fue significativo, ya que para esto fue necesario el uso de diversos programas, tales como: Excel, QGIS, R, Xming, Climate Data Operator y NCAR Command Language (NCL) de Ubuntu. Para el uso de algunos de ellos se requirió tiempo de aprendizaje durante la ejecución del proyecto para así desarrollar los códigos y procesar datos de manera eficiente.

También es relevante nombrar los tiempos requeridos para el procesamiento de los datos escalados. Por ejemplo, para la descarga de los datos ERA5-LAND, en un principio se optó por descargar la información de los 20 años de manera continua, pero durante el proceso la interfaz Rstudio cerraba sesión inesperadamente, el computador se apagaba de manera repentina o el disco duro se saturaba con la memoria llena. En una ocasión donde supuestamente la descarga llegó a 100%, el archivo no se generó lo que obligó a buscar otras estrategias para la descarga. Al final se decidió descargar los datos para cada año independientemente y realizando respaldos en disco externo. En el caso de temperatura máxima y mínima el proceso tomó alrededor de más de 24 horas para cada una de las variables.

Posteriormente, se pudo notar que al intentar realizar el escalamiento para archivos de 20 años completos (tal como se habían descargado los datos) resultaba en un proceso muy demorado y poco recomendable, ya que este duró alrededor de 5 días escalando el archivo correspondiente, para finalmente fallar por problemas computacionales del equipo. Finalmente, fue necesario realizar el escalamiento de los datos en paquetes anuales resultando en un proceso que tomó entre 6 y 8 horas por cada año trabajado.

Por último, durante el escalamiento de datos satelitales de precipitación CHIRPSv2 con la librería KrigR de R, se pudo identificar que el código no estaba preparado para escalar datos para días en los que no se registró precipitación en ninguno de los píxeles demarcados dentro de la zona de estudio. Esta condición produjo un error no identificado inicialmente que no permitió realizar el proceso de escalamiento de los archivos correspondientes a algunos de los años estudiados de la forma adecuada. Dichos días tuvieron que ser procesados aparte mediante QGIS debido a que solo se debía escalar el tamaño de píxeles con el mismo valor.

Debido al covariante usado (modelo digital de elevación), el código generó resultados inconsistentes en los casos en que no existían datos de precipitación en zonas con altos gradientes de elevación. Por lo cual los archivos escalados tuvieron que ser sometidos a procesos de modificación para poder limitar los rangos de los valores negativos a cero, mediante el uso de las funciones de Climate Data Operator (CDO) en Ubuntu.

Por último, al extraer las series de tiempo de los datos escalados de precipitación para la comparación punto-píxel, se encontraron valores extraños de orden exponencial exageradamente grandes, como por ejemplo $4.71E^{+37}$; valores que no coincidían con los valores mínimos y máximos identificados dentro de los metadatos de los archivos satelitales CHIRPSv2 escalados. Además, estos valores problemáticos, siempre se presentaban en la misma fecha. Al corroborar con el visor de datos Xming, se logró determinar que esos valores eran 0, por lo cual se concluyó que este error se encontraba ligado a errores de aproximación de las cifras decimales de la precipitación escalada al extraer las series de tiempo. El apéndice F da mayor aclaración sobre los errores causados por precipitaciones nulas.

3.3 Fase III – Determinación del ajuste de los datos escalados para las estaciones seleccionadas mediante el cálculo de indicadores estadísticos.

Para la determinación del ajuste de los datos escalados se realizó la comparación punto-píxel de las series de tiempo obtenidas mediante el proceso de escalamiento Kriging en el píxel correspondiente a las coordenadas de cada una de las estaciones IDEAM seleccionadas para el cálculo de los indicadores estadísticos. Este análisis no tuvo en cuenta los datos imputados pues esto generaría más incertidumbre en el análisis.

Para evaluar el ajuste se usaron seis indicadores estadísticos: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Modified Kling-Gupta Efficiency (KGE), Bias (β), Variabilidad (γ) y Coeficiente de correlación de Pearson (r); estos han sido implementados en estudios similares (Blanco, 2022; Valencia et al., 2023). Se implementó un código de R que permite calcular el error entre las series de tiempo diarias de los datos observados y escalados.

A continuación, se detallan cada una de las métricas evaluadas. Las ecuaciones respectivas a cada una estas, junto con su explicación, se encuentran en el Apéndice G.

Root Mean Square Error (RMSE): Es el indicador de validación cruzada más común y es más sensible a los valores atípicos. Los valores de RMSE cercanos a cero indican un rendimiento más óptimo (Njoku et al., 2023).

$$RMSE = \sqrt{\frac{1}{n} * \sum_{i=1}^n (S_i - O_i)^2}$$

Mean Absolute Error (MAE): Se define como el promedio de las diferencias absolutas entre las variables medidas y estimadas. Los valores de MAE cercanos a cero indican mejor ajuste (Njoku et al., 2023).

$$MAE = \frac{1}{n} * \sum_{i=1}^n |S_i - O_i|$$

Modified Kling-Gupta Efficiency (KGE): Indicador estadístico ampliamente usado para evaluar el desempeño de modelos hidrológicos en términos de sus tres componentes: Coeficiente de Correlación de Pearson (r), Bias (β) y Variabilidad (γ). Los valores del KGE oscilan en el rango de $-\infty$ hasta 1 (mayor precisión del modelo) (Valencia et al., 2023).

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

Coefficiente de Correlación de Pearson (r): Indicador estadístico de correlación lineal entre los valores observados y los valores estimados. Los valores de r oscilan en el rango de -1 a 1, donde los negativos indican correlación inversa entre las variables, si $r=0$ significa que no existe correlación alguna entre las variables y los valores más cercanos a uno indican correlación lineal positiva (Valencia et al., 2023).

$$r = \frac{\sum_{i=1}^n (O_i - \bar{O})(S_i - \bar{S})}{\sqrt{\sum_{i=1}^n (O_i - \bar{O})^2} \sqrt{\sum_{i=1}^n (S_i - \bar{S})^2}}$$

Bias (β): Mide la tendencia promedio que presentan los modelos para estimar los valores requeridos frente a los valores observados. Los valores de β pueden variar de $-\infty$ a $+\infty$, donde los valores menores a uno indican subestimación y los mayores a uno indican sobreestimación de las variables del modelo (Valencia et al., 2023).

$$\beta = \frac{\mu_s}{\mu_0}$$

Variabilidad (γ): Evalúa la relación de la dispersión de los valores simulados frente a los observados, indicando si estos tienen una dispersión similar a los datos observacionales cuando los valores son cercanos a 1. (Valencia et al., 2023).

$$\gamma = \frac{CV_s}{CV_0}$$

4. Resultados

Con este trabajo se generó una base de datos climatológica (precipitación y temperatura) para el departamento de Santander para la ventana de tiempo 1990-2009 con una alta resolución espacial (~1000 m) que permite su uso para proyectos de carácter local y regional (Apéndices J, K

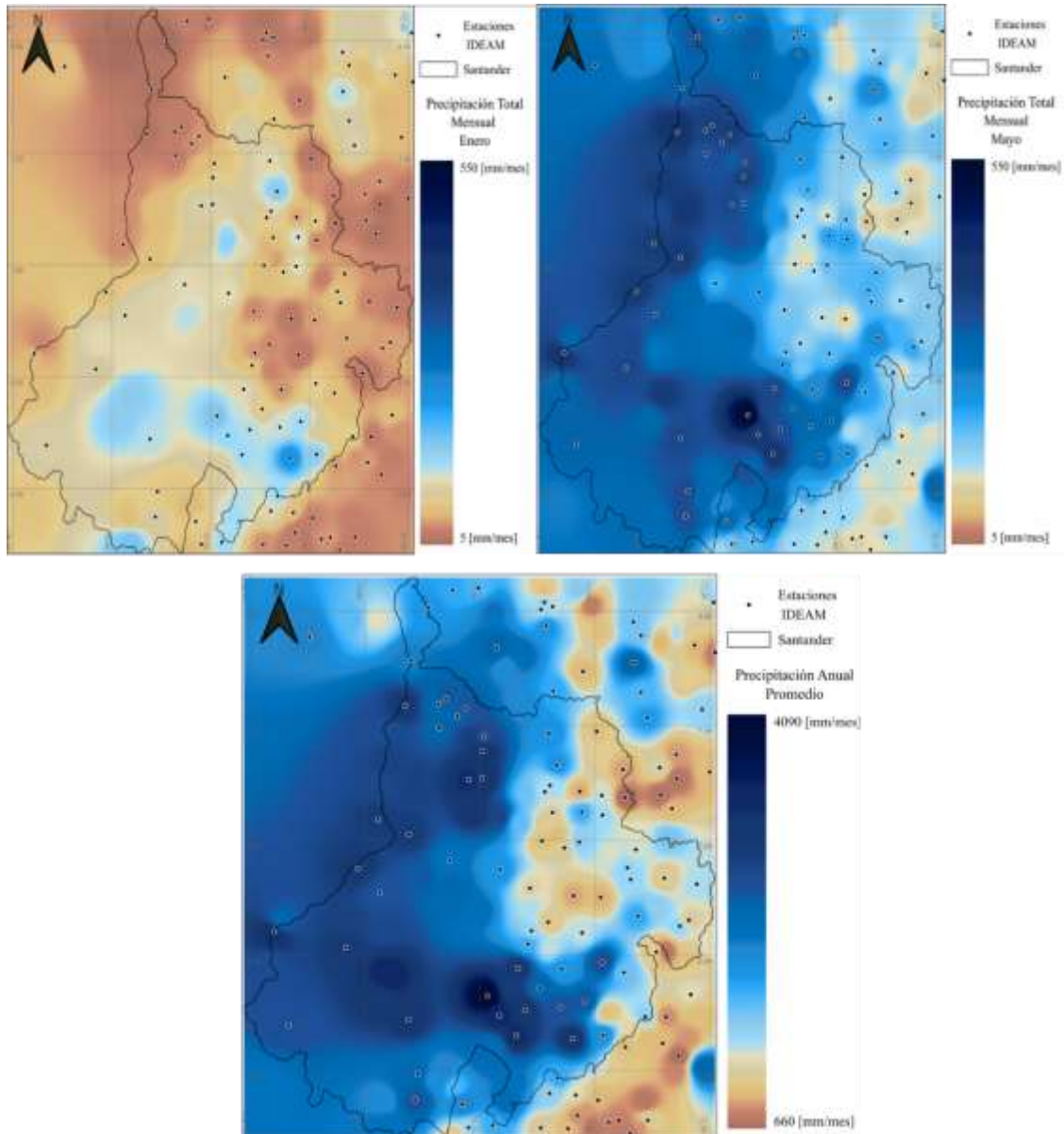
y L). Los datos IDEAM procesados y consolidados para fines de verificación del ejercicio de escalamiento también son parte de los productos generados (Apéndice B). Adicionalmente, los códigos implementados para los diferentes procesos metodológicos se ofrecen al lector de este trabajo de tal manera que sirvan de herramienta para ampliar la disponibilidad de datos a nuevas ventanas de tiempo o para implementar un ejercicio similar en otras zonas de estudio (Apéndices C, D, y E). Se realiza un análisis del ajuste entre los datos observados y escalados y se hace una evaluación del impacto de estos datos escalados a partir de la comparación entre las superficies interpoladas de precipitación y temperatura, mensuales y anuales, para la zona de estudio.

4.1 Superficies climatológicas de precipitación total (mensual y anual) y temperatura media (mensual y anual) a partir de series de tiempo diarias de estaciones del IDEAM

A partir de los datos IDEAM procesados según se describe en la sección 3.1.1. de la metodología se generaron las superficies climatológicas de precipitación, y temperatura mínima y máxima para el departamento de Santander (Figuras 6-8). Para el caso de la precipitación (Figura 6), la distribución existente de estaciones muestra adecuadamente, desde una perspectiva regional, los patrones geográficos indicando mayores precipitaciones al oeste y sur del departamento y claras diferencias de precipitación de carácter estacional (enero vs. mayo). La poca densidad de estaciones de temperatura no permite identificar un gradiente claro a pesar de que existe una variedad de temperaturas en el departamento definidas por su topografía y propiedades biofísicas. La existencia de la estación Berlín permite que se dé un indicio de las menores temperaturas características de la zona de páramo. Sin embargo, se requeriría la existencia de más estaciones en esta región para poder hacer una delimitación clara de su extensión (Figura 8).

Figura 6

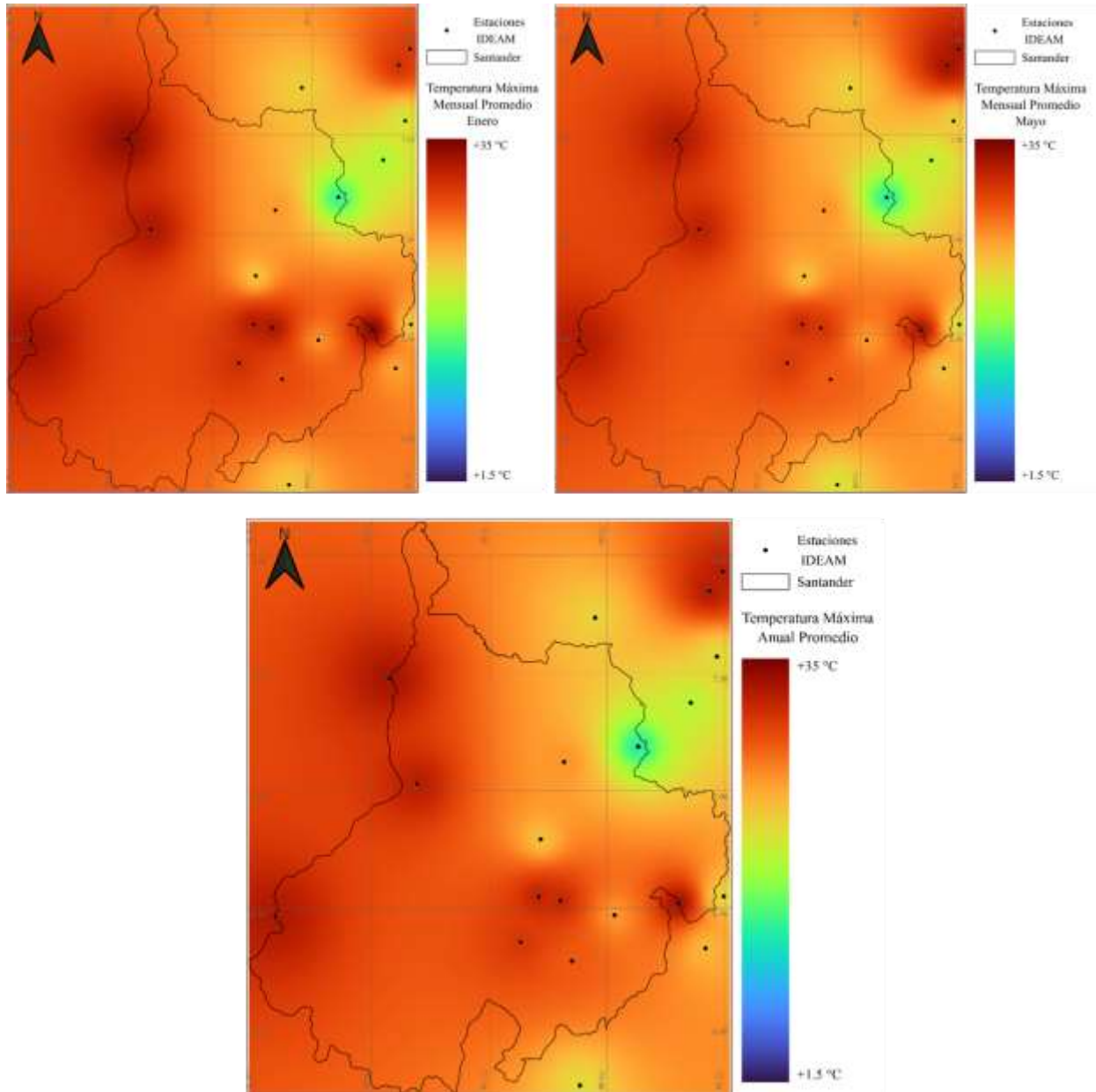
Promedios de precipitación mensual y anual para Santander con datos IDEAM



Nota. El gráfico representa los promedios de la precipitación total mensual de los meses enero y mayo (arriba), y de la precipitación total anual (abajo) en Santander con la información procesada de estaciones del IDEAM.

Figura 7

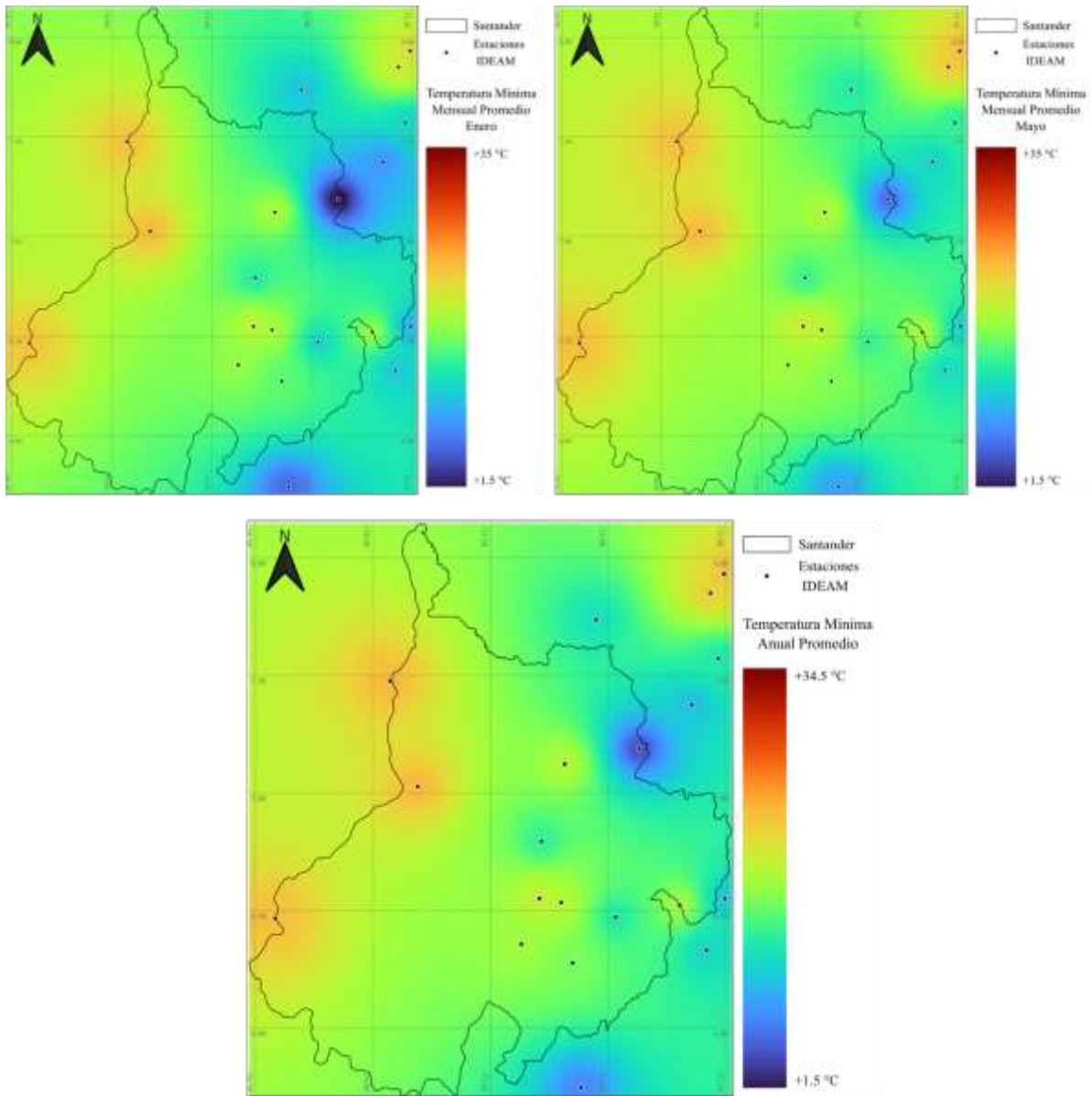
Temperatura máxima mensual y anual para Santander con datos IDEAM



Nota. El gráfico representa la temperatura máxima anual y mensual promedio de los meses de enero y mayo en Santander con la información procesada del IDEAM.

Figura 8

Temperatura mínima mensual y anual para Santander con datos IDEAM



Nota. El gráfico representa la temperatura mínima anual y mensual promedio de los meses de enero y mayo en Santander con la información procesada del IDEAM.

4.2 Superficies climatológicas de precipitación total (mensual y anual) y temperatura media (mensual y anual) a partir de bases de datos escaladas

Mediante los datos obtenidos de CHIRPSv2 para precipitación y ERA5-Land para temperatura, como se describe en la sección 3.2 de la metodología, se generaron las superficies climatológicas escaladas de alta resolución espacial de precipitación y temperatura mínima y máxima para el departamento de Santander a escala diaria (Figuras 9-11), mensual y anual (Figuras 12-14).

4.2.1 Bases de datos crudos de las variables meteorológicas

4.2.1.1 Información CHIRPS. Se logró la descarga de los datos de precipitación total diaria con una absoluta continuidad espacial y temporal para la ventana del proyecto y una resolución espacial de 0.05° (5 km). En la figura 9 (izquierda), se muestra un ejemplo de la información presente en CHIRPS en su resolución original. En el Apéndice H se encuentran los archivos “netCDF” descargados de la base de datos CHIRPS.

4.2.1.2 Información ERA5-Land. Se logró la descarga de los datos de temperatura máxima y mínima con absoluta continuidad espacial y temporal para la ventana del proyecto y una resolución espacial de 0.1° (10 km). En las figuras 10 y 11 (izquierda), se muestra un ejemplo de la información presente en ERA5-land. En el apéndice I se encuentra el código de descarga y los archivos “netCDF” descargados de la base de datos CHIRPS.

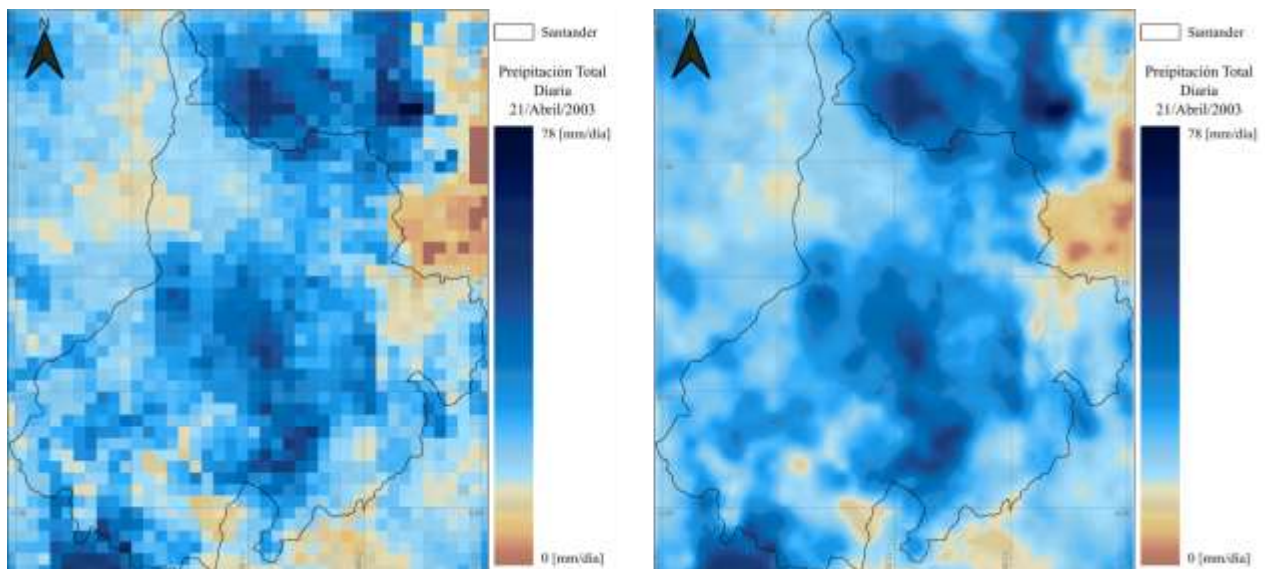
4.2.2 Bases de datos escalados de las variables meteorológicas.

El proceso de escalamiento generó datos de precipitación y temperatura máxima y mínima en una resolución de 0.01° (~1000 m) para la zona de estudio en la ventana de tiempo seleccionada.

En las Figuras 9-11, se muestra un comparativo entre los datos crudos (izquierda) y los datos escalados (derecha) para una fecha específica. Para el caso de la precipitación total diaria, el escalamiento define y suaviza los gradientes de precipitación, respetando las tendencias generales de los datos crudos. Para el caso de la temperatura, el escalamiento logra mejorar significativamente la distribución de valores de las dos variables. Se hace evidente el papel que juega el covariante usado (DEM) en el proceso de escalamiento. El Apéndice J contiene los datos escalados de precipitación, y los apéndices K y L contienen los datos escalados de temperatura máxima y mínima, respectivamente.

Figura 9

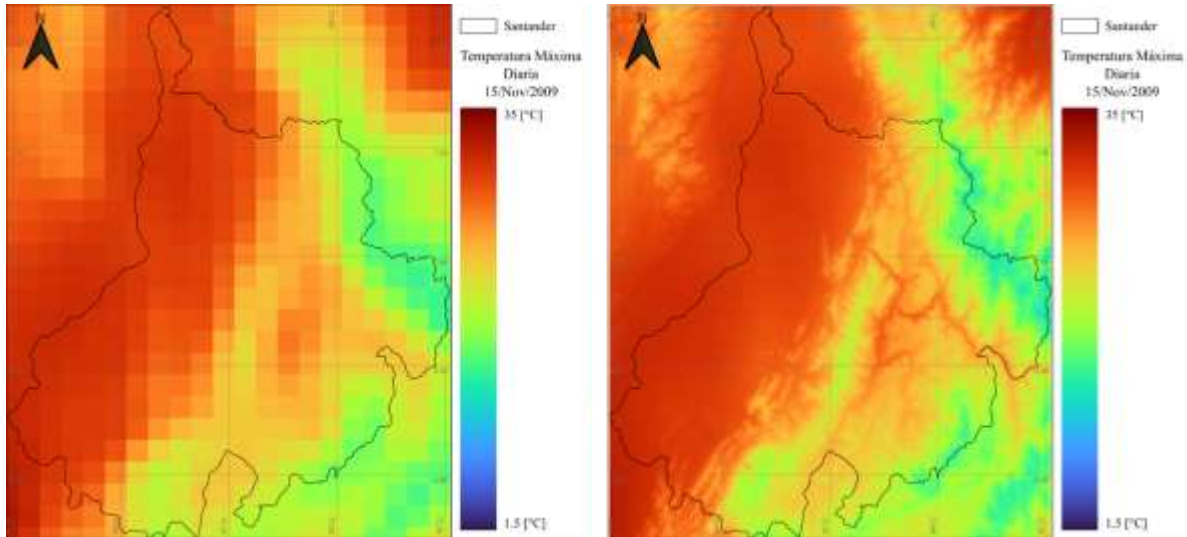
Precipitación total diaria CHIRPS



Nota. El gráfico representa un ejemplo de datos de precipitación total diaria – 15/11/2009, 21/04/2003, resolución espacial 0.05° (5 km).

Figura 10

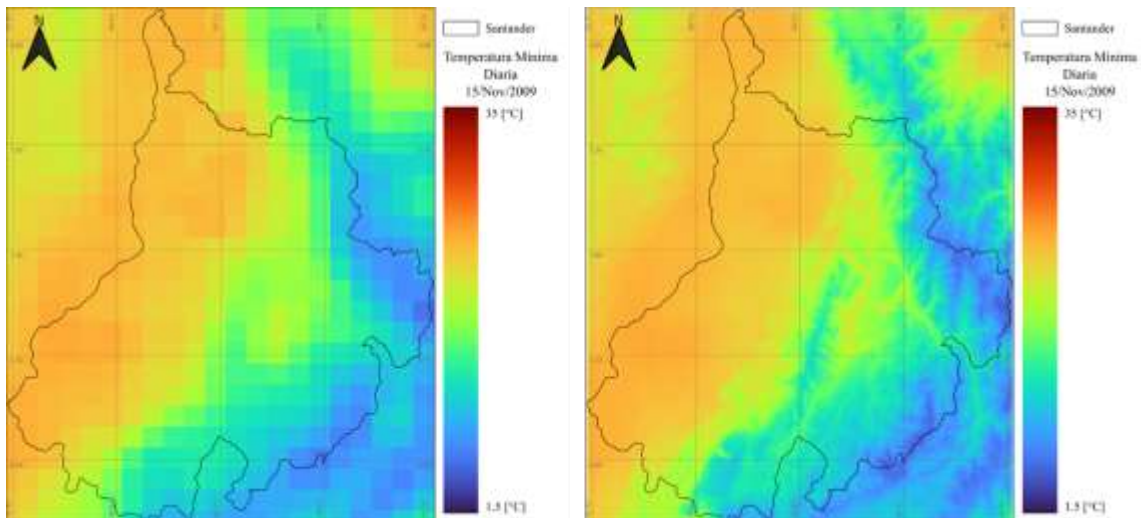
Temperatura máxima diaria ERA5-Land



Nota. El gráfico representa un ejemplo de datos de temperatura máxima– 15/11/2009, resolución espacial 0.1° (10 km).

Figura 11

Temperatura mínima diaria ERA5-Land



Nota. El gráfico representa un ejemplo de datos de temperatura máxima– 15/11/2009, resolución espacial 0.1° (10 km).

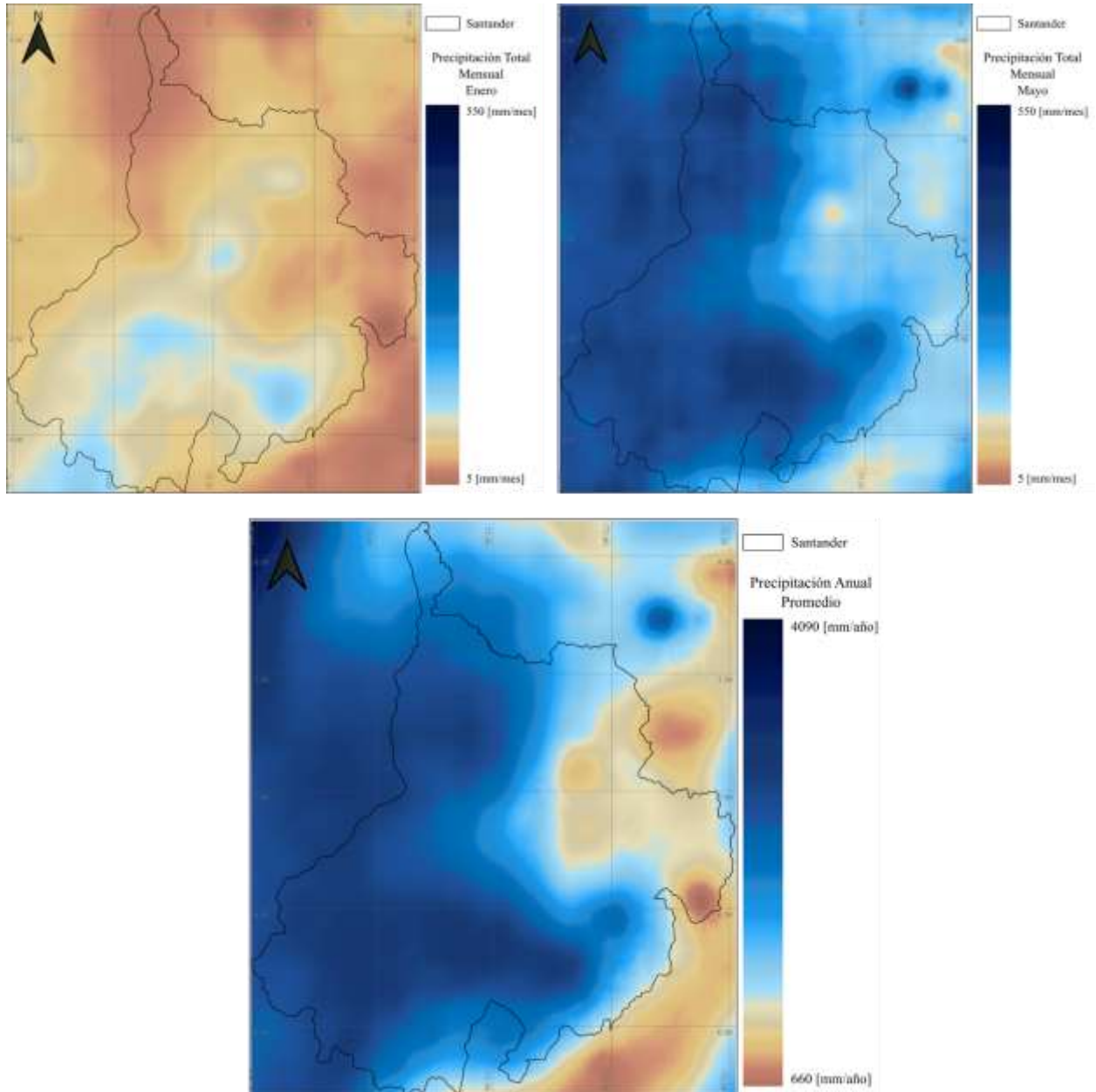
4.2.3 Superficies climatológicas generadas con datos satelitales escalados

Los datos escalados permitieron generar superficies mensuales y anuales que ofrecen una apreciación más clara de los patrones espaciales y temporales de precipitación y temperatura en el departamento de Santander. Para el caso de la precipitación, el promedio el promedio anual (Figura 12) confirma el gradiente establecido en el mapa climatológico del IDEAM Precipitación media total anual – Promedio multianual 1981-2010), pero con la ventaja de una mayor resolución espacial. El comparativo entre un mes seco (enero) y uno lluvioso (mayo), informa sobre los rangos de precipitación promedio que ocurren en dichas condiciones y la distribución geográfica detallada de los patrones de lluvia en el departamento (Figuras 12). Se concluye entonces que, en promedio, las mayores precipitaciones se ubican al oeste y sur del departamento, pero que hay claras diferencias estacionales de precipitación, como ya se ha establecido para la región andina (Valencia et al., 2023). El Apéndice M presenta las superficies de precipitación total mensual promedio para los 12 meses del año.

Los patrones de temperatura promedio anual son consistentes con los reportados por el IDEAM (Temperatura máxima media promedio multianual y temperatura mínima media promedio multianual). El comparativo entre las superficies presentadas en las figuras 13 y 14 evidencian la poca robustez que pueden ofrecer los datos *in-situ* cuando no existe una densidad adecuada de estaciones. Las superficies de los promedios mensuales de temperatura máxima y mínima muestran poca variabilidad entre temporadas, confirmando que es más relevante el gradiente topográfico en la distribución de la temperatura que los cambios estacionales en esta zona ecuatorial. El Apéndice N presenta las superficies de temperatura promedio máxima y mínima para los 12 meses del año.

Figura 12

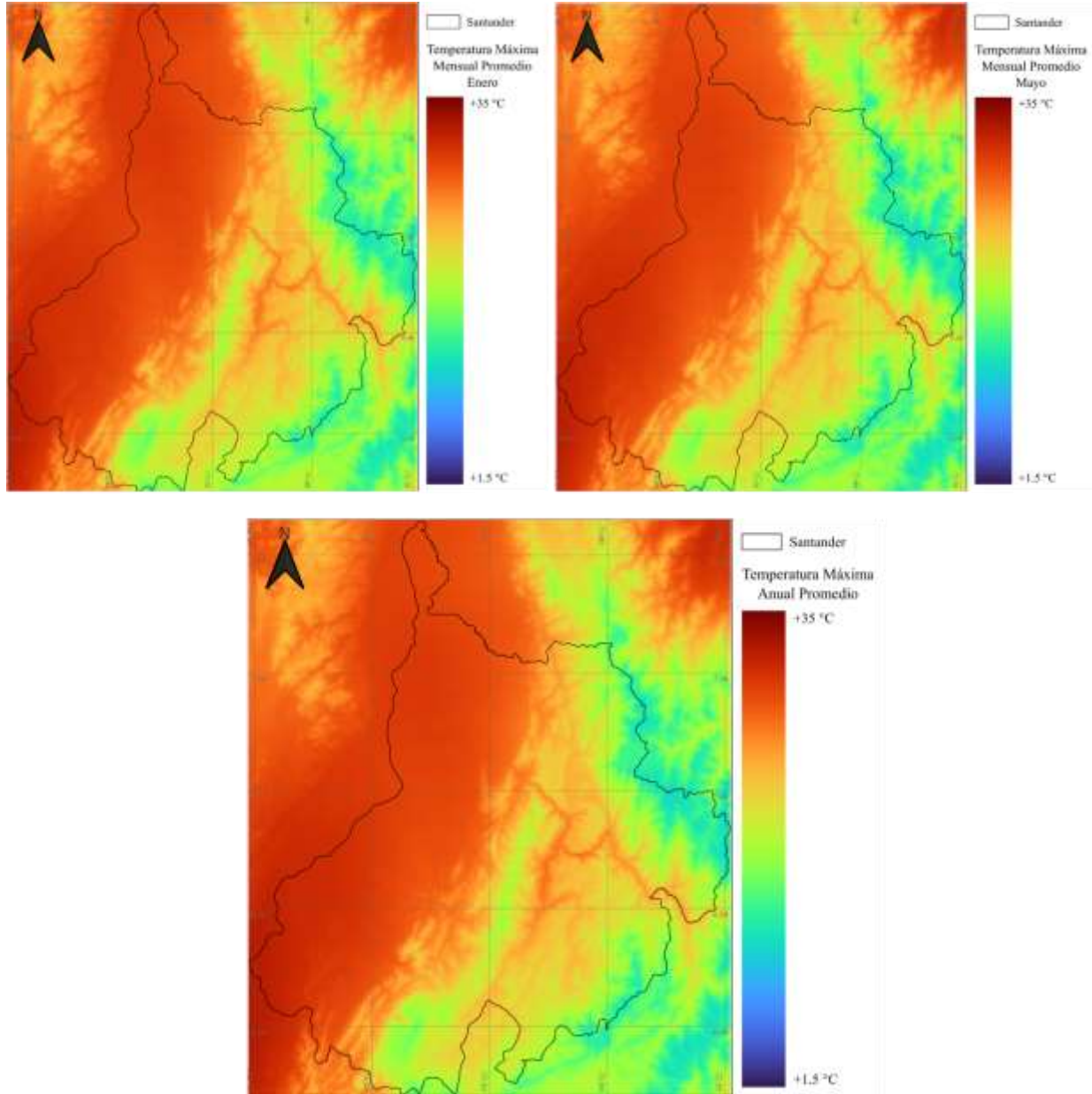
Precipitación mensual y anual para Santander con datos CHIRPSv2



Nota. El gráfico representa la precipitación anual y mensual total de los meses de enero y mayo en Santander con la información escalada de la base de datos CHIRPS.

Figura 13

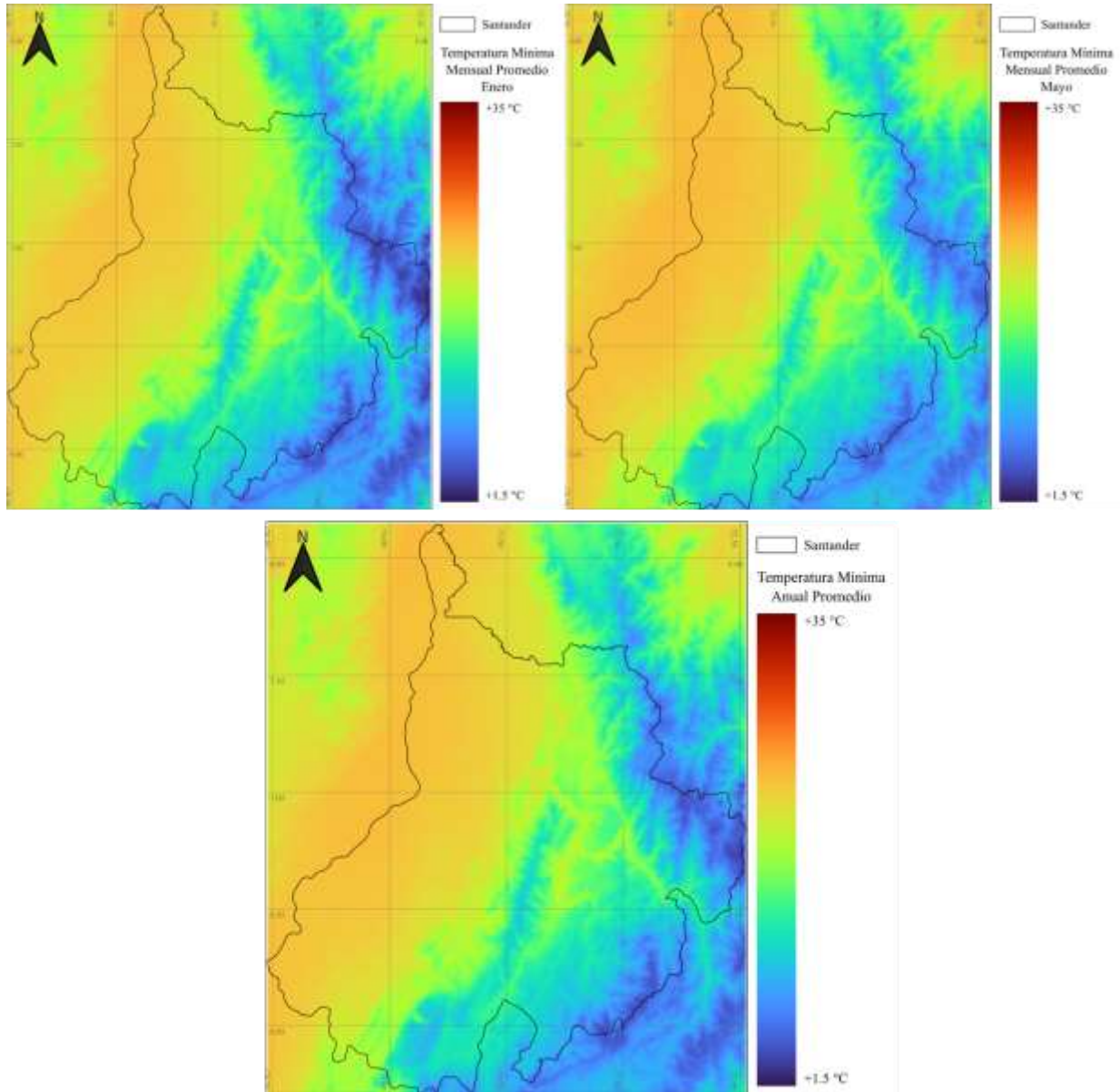
Temperatura máxima mensual y anual para Santander con datos ERA5-Land



Nota. El gráfico representa la temperatura máxima anual y mensual promedio de los meses de enero y mayo en Santander con la información escalada de la base de datos ERA5-Land.

Figura 14

Temperatura mínima mensual y anual para Santander con datos ERA5-Land

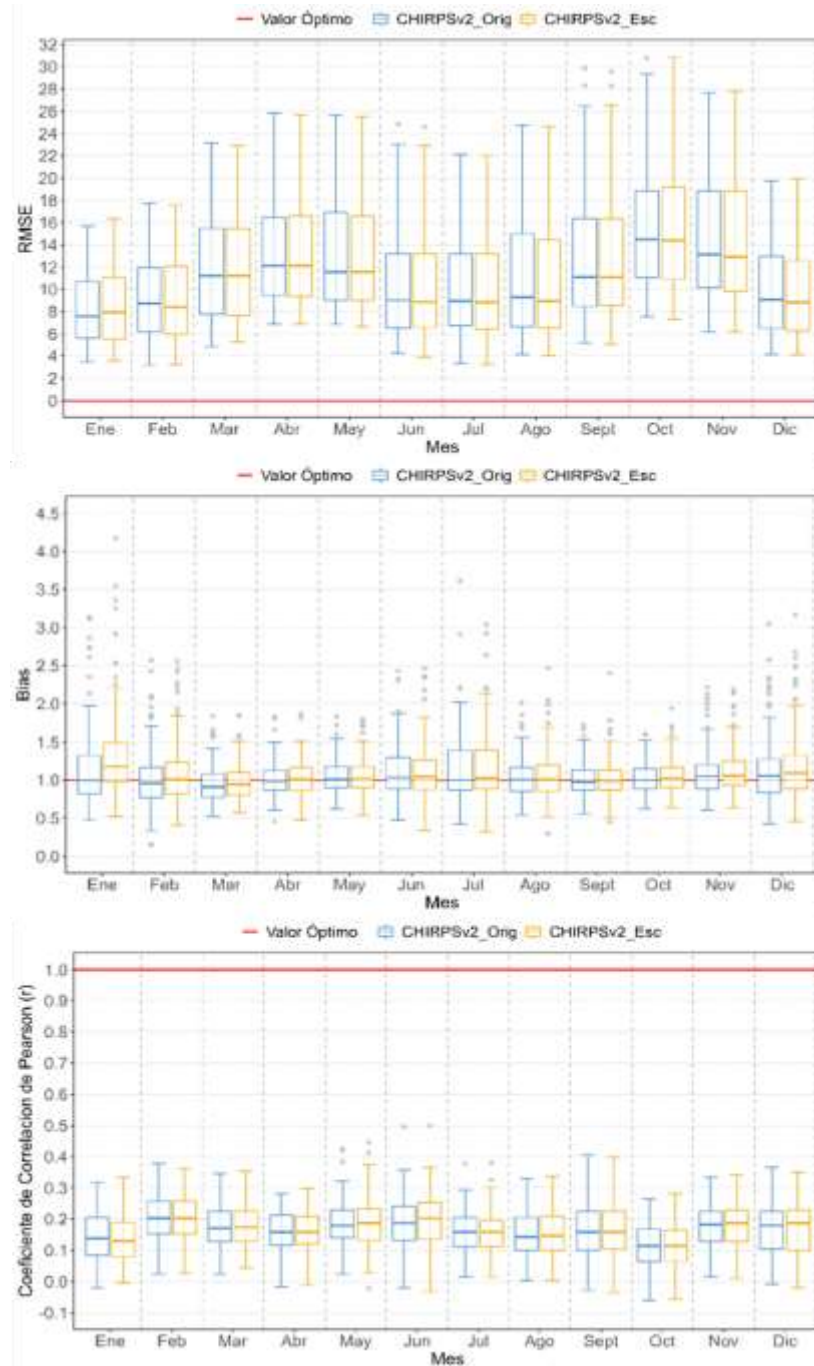


Nota. El gráfico representa la temperatura mínima anual y mensual promedio de los meses de enero y mayo en Santander con la información escalada de la base de datos ERA5-Land

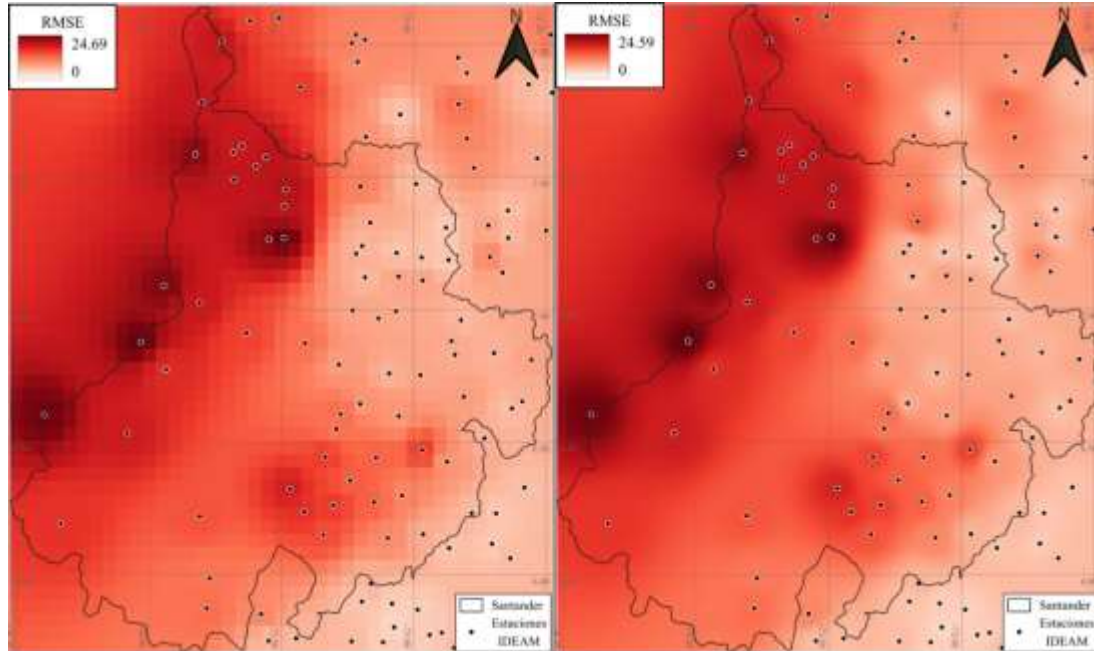
4.3 Grado de incertidumbre

4.3.1 Precipitación

En la Figura 15 se presentan los resultados más significativos de las métricas de desempeño estimadas para la variable de precipitación con los datos CHIRPS originales y escalados. De la gráfica de RMSE se observa que en los meses de marzo-mayo y septiembre-noviembre -los meses más lluviosos del año- se presentan mayores diferencias entre los datos. Por otro lado, la gráfica del coeficiente de correlación de Pearson (r) evidencia que los datos a lo largo del año poseen una correlación lineal positiva con valores iguales o inferiores a 0.5. Adicionalmente, el sesgo (Bias) expone un comportamiento de los datos bastante cercano a 1; sin embargo, se debe mencionar que en los meses de enero y julio se encontraron valores considerablemente atípicos. Para ninguno de las métricas se evidencia un incremento del error como resultado del escalamiento. Especialmente, los errores más grandes se localizan al oeste del departamento, donde se presenta mayor precipitación (Figura 16). La información completa con las 6 métricas de desempeño (Bias, RSME, KGE, Pearson, variabilidad y MAE) se presentan en el Apéndice O.

Figura 15*Métricas estimadas precipitación CHIRPS*

Nota. El gráfico representa las métricas estadísticas estimadas que se obtuvieron para precipitación en la base de datos CHIRPS, la información completa está en el apéndice O.

Figura 16*Distribución RMSE agosto – precipitación total diaria*

Nota. El gráfico representa la distribución del RMSE en Santander del mes de agosto, con datos sin escalar y escalados de la variable precipitación total diaria en CHIRPS.

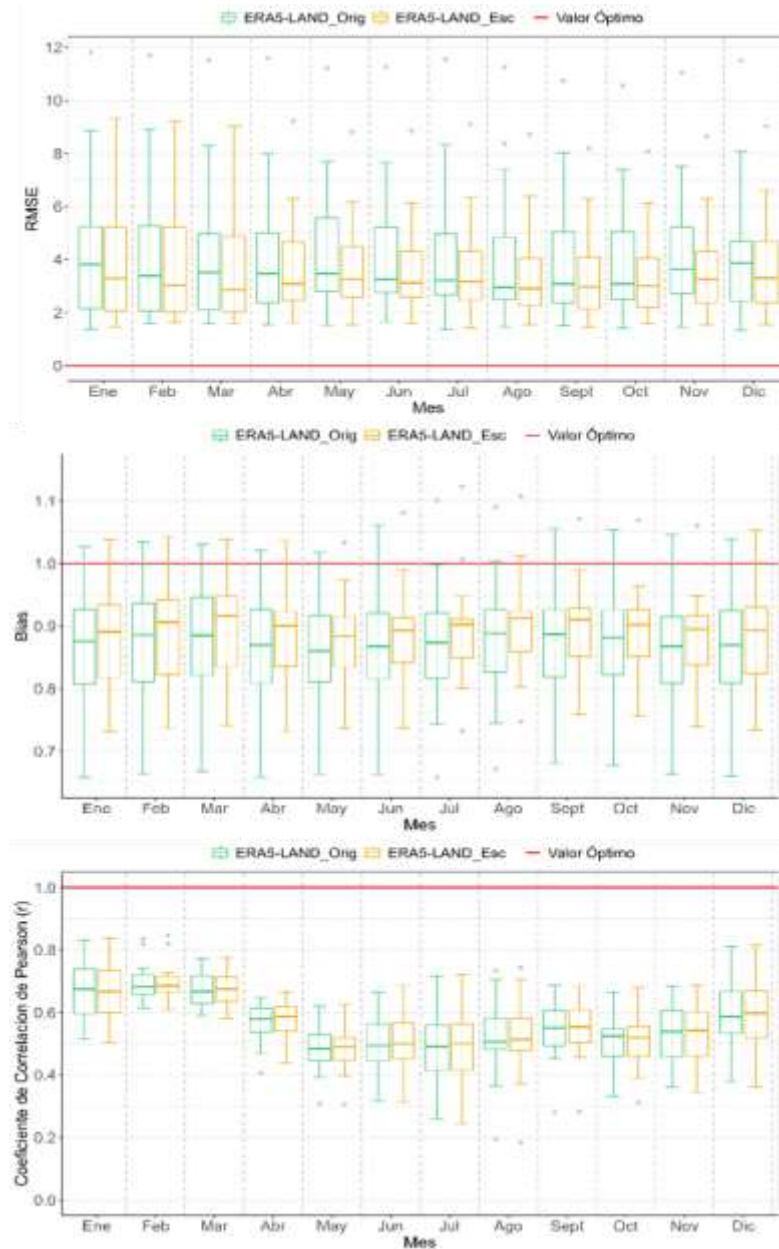
4.3.2 Temperatura máxima

En la Figura 17 se presentan los resultados más significativos de las métricas de desempeño estimadas para la variable de temperatura máxima con los datos ERA5-Land originales y escalados. Como se esperaba, las métricas de error son mucho mejores para esta variable y tanto para el RMSE como para el sesgo se aprecia una reducción del rango de error con los datos escalados. No se evidencia una variabilidad estacional del error RMSE, pero al comparar los datos originales con los escalados se encontró una mejora en 9 de los 12 meses (75%). Por otro lado, la gráfica del coeficiente de correlación de Pearson (r) evidencia que los datos a lo largo del año poseen una correlación lineal positiva con valores entre 0.8 y 0.2; es destacable que en los primeros

3 meses del año la correlación está entre 0.85 y 0.5. La gráfica de sesgo expone una tendencia a la subestimación de los datos satelitales con respecto de los observados, sin embargo, se debe mencionar que los datos escalados presentan una ligera mejora en el rango de error. Especialmente, los errores más grandes (RMSE) se localizan al este del Departamento, aunque vale la pena aclarar que esta distribución está limitada por la poca disponibilidad de estaciones de temperatura (Figura 18). La información completa con las 6 métricas de desempeño (Bias, RSME, KGE, Pearson, variabilidad y MAE) para temperatura máxima se encuentran en el Apéndice O.

Figura 17

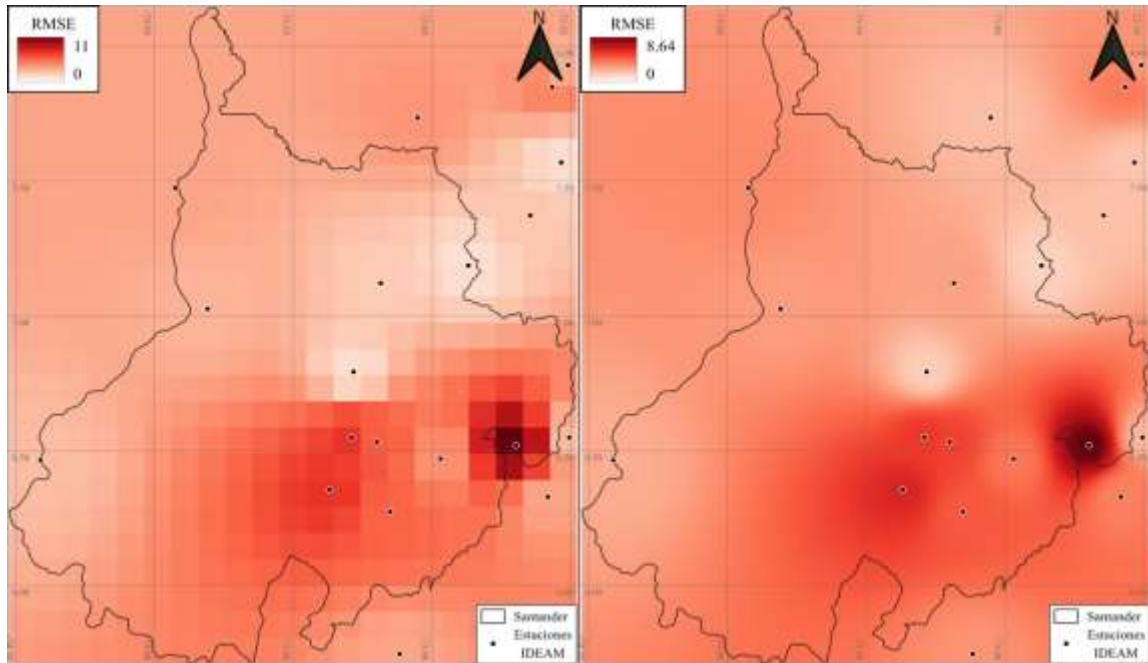
Métricas estimadas temperatura máxima ERA5-Land



Nota. El gráfico representa las métricas estadísticas estimadas que se obtuvieron para temperatura máxima en la base de datos ERA5-Land, la información completa está en el apéndice O.

Figura 18

Distribución RMSE noviembre – temperatura máxima

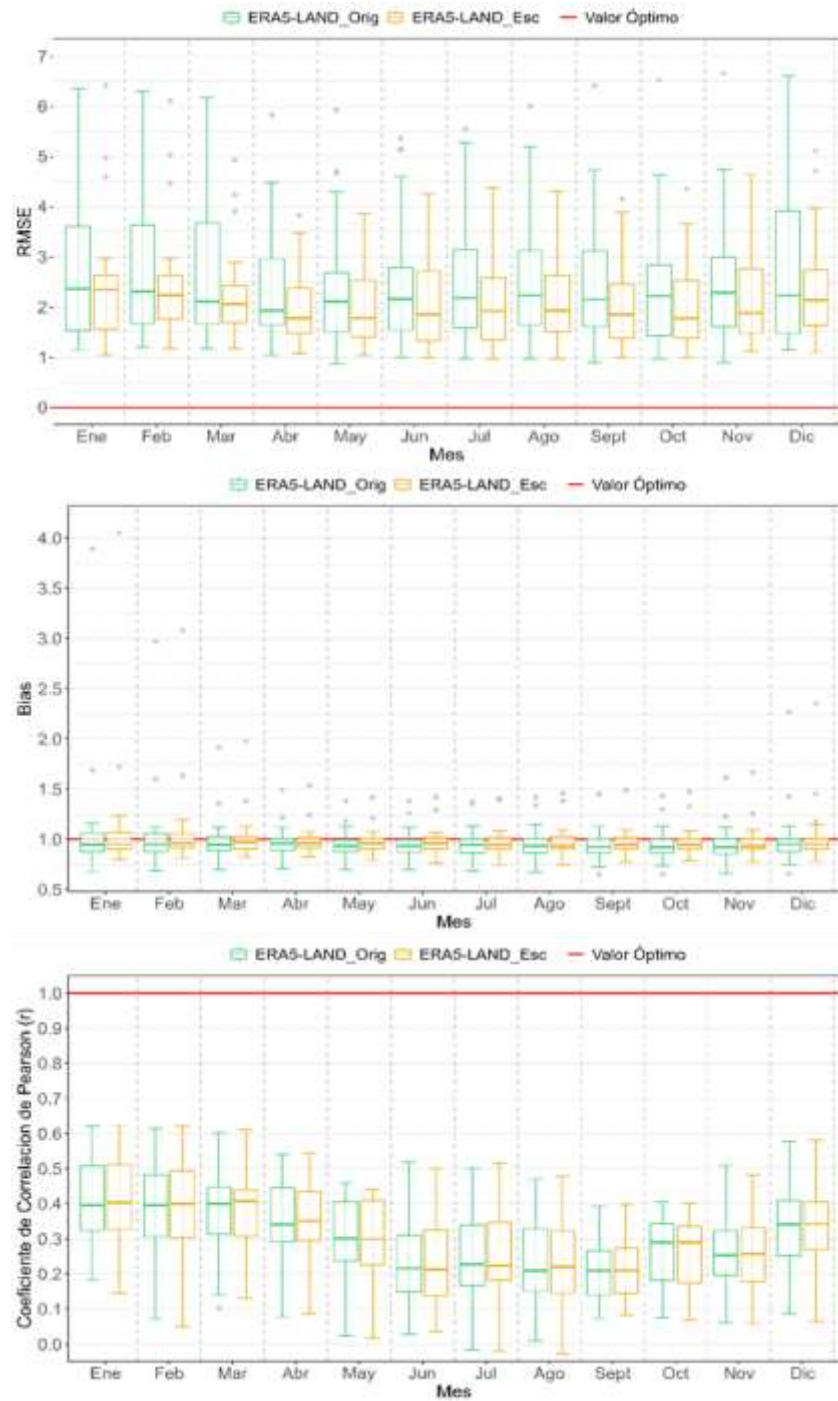


Nota. El gráfico representa la distribución del RMSE en Santander del mes de noviembre, con datos sin escalar y escalados de la variable temperatura máxima en ERA5-Land.

4.3.3 Temperatura mínima

En la Figura 19 se presentan los resultados más significativos de las métricas de desempeño estimadas para la variable de temperatura mínima con los datos ERA5-Land originales y escalados. En general, con respecto a las tendencias de error para la temperatura máxima, se evidencia una mejora de las métricas de error para el caso de RMSE y sesgo, y un incremento del error estimado por el coeficiente de correlación de Pearson. No se aprecia una tendencia de error RMSE estacional, pero sí una mejora del error al comparar el desempeño de los datos escalados con

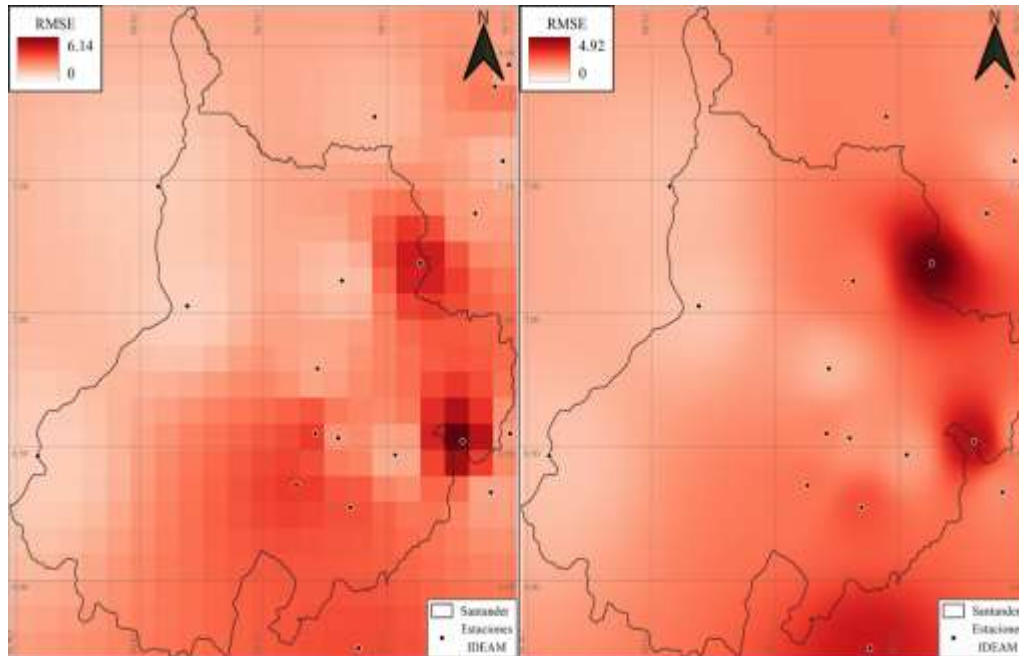
respecto a los datos crudos. El rango de error para esta métrica mejoró para todos los meses del año. La gráfica del coeficiente de correlación de Pearson (r) evidencia que los datos a lo largo del año poseen una correlación lineal positiva con valores entre 0.65 y 0, pero los rangos de error no mejoran al comparar los ajustes para datos crudos y escalados. El sesgo expone un comportamiento de los datos bastante cercano a 1 para los comparativos de datos crudos y escalados, pero se presentan valores altamente atípicos en el período de diciembre a marzo. Para algunos meses se aprecia la mejora del error para el caso de los datos escalados. La distribución espacial del error (RMSE) (Figura 20) es similar a la de temperatura máxima. La información completa con las 6 métricas de desempeño (Bias, RSME, KGE, Pearson, variabilidad y MAE) se encuentran en el apéndice O.

Figura 19*Métricas estimadas temperatura mínima ERA5-Land*

Nota. El gráfico representa las métricas estadísticas estimadas que se obtuvieron para temperatura mínima en la base de datos ERA5-Land, la información completa está en el apéndice O.

Figura 20

Distribución RMSE marzo – temperatura mínima



Nota. El gráfico representa la distribución del RMSE en Santander del mes de marzo, con datos sin escalar y escalados de la variable temperatura mínima en ERA5-Land.

5. Conclusiones

Con este trabajo se generó una base de datos climatológica (precipitación y temperatura máxima y mínima) para el departamento de Santander para la ventana de tiempo 1990-2009 con una alta resolución espacial (~1000 m) que permite su uso para proyectos de carácter local y regional. Los datos IDEAM procesados y consolidados para fines de verificación del ejercicio de escalamiento también son parte de los productos generados. Adicionalmente, los códigos implementados para los diferentes procesos metodológicos se ofrecen al lector de este trabajo de

tal manera que sirvan de herramienta para ampliar la disponibilidad de datos a nuevas ventanas de tiempo o para implementar un ejercicio similar en otras zonas de estudio.

El análisis del ajuste de los datos brinda herramientas para justificar el uso de los datos en proyectos aplicados. En general, las métricas de error usadas indican que los datos satelitales y de reanálisis tienen un buen ajuste, aunque hay diferencias en el desempeño para las tres variables estudiadas. Para el caso de la precipitación total diaria, los errores fueron mayores que para las variables de temperatura y se evidencia una variabilidad temporal donde los errores fueron superiores durante las temporadas lluviosas del año (marzo-mayo y septiembre-noviembre). Así mismo, al hacer la comparación punto-píxel entre datos escalados y observados vs. datos crudos y observados no se evidenció un incremento en las métricas de error; esto se considera ventajoso en tanto que los datos escalados ofrecen una mayor resolución espacial y podrían, por lo tanto, ser usados preferencialmente a los datos CHIRPSv2 crudos en zonas donde no exista información de precipitación.

Para el caso de la variable temperatura, las métricas de error fueron mucho mejores que para la precipitación, y el desempeño fue mejor para temperatura mínima que para temperatura máxima. Adicionalmente, en la mayoría de los casos se evidenció una reducción de los rangos de error en la comparación punto-píxel de los datos observados y escalados con respecto a la comparación entre datos observados y crudos. Se reconoce aquí el papel que juega el covariante modelo digital de elevación (MDE) en el escalamiento adecuado de la temperatura, teniendo en cuenta las características topográficas del área de estudio.

Las superficies de precipitación y temperatura (anuales y mensuales) generadas con datos satelitales y de reanálisis se comparan positivamente con las disponibles a través del IDEAM en términos de distribución de valores y rangos. Estas superficies generadas ofrecen la ventaja de una

mayor resolución espacial que permite su aplicación para estudios locales y regionales. Estos resultados contrastan con el caso de las superficies generadas usando datos observacionales de estaciones IDEAM, especialmente para el caso de las variables de temperatura.

6. Recomendaciones

Aunque el ejercicio de escalamiento de los datos de precipitación CHIRPSv2 no mejoró las métricas de error, se recomienda usar esta base de datos dado que ofrece una mayor resolución espacial y cuenta con absoluta continuidad temporal. Los datos y superficies escaladas evidencian claramente los patrones climáticos con tendencia bimodal presentes en la región andina y los patrones geográficos que indican mayores precipitaciones al oeste y sur del departamento. Los resultados del proceso de escalamiento se podrían mejorar para esta variable con la incorporación de un covariante que complemente el MDE. Se sugiere que, a futuro, se haga un estudio que identifique y produzca, en las escalas requeridas, este nuevo insumo.

En el caso de los datos observacionales (IDEAM), el proceso de selección de estaciones con datos de alta calidad en términos de disponibilidad continua de datos y extensión temporal resultó en el uso de 113 estaciones de precipitación y 20 de temperatura mínima y máxima, lo cual representa aproximadamente el 26% y 31% respectivamente de las estaciones descargadas inicialmente. Es necesario realizar un esfuerzo interinstitucional para robustecer la red observacional de datos hidrometeorológicos del Departamento y del País. Estos datos a su vez servirán para robustecer la calidad de las bases de datos satelitales y de reanálisis.

Referencias Bibliográficas

- Achá, N. A., Saavedra, O. C., Ureña, J. E., Achá, N. A., Saavedra, O. C., & Ureña, J. E. (2022). MODELACIÓN HIDROLÓGICA EN LA CUENCA DEL RÍO ROCHA INCORPORANDO LINEAMIENTOS DE CAUDAL ECOLÓGICO. *Investigación & Desarrollo*, 22(1), 49–62. <https://doi.org/10.23881/idupbo.022.1-5i>
- Anagua, F., & Villazon, M. (2018, septiembre 7). *VALIDACIÓN DEL SATÉLITE METEOROLÓGICO CHIRPS EN LA CUENCA RURRENABAQUE-BOLIVIA*.
- Antonio, L.-R., William. (2018). *Clima, hidrología y meteorología.: Para ciencias ambientales e ingeniería*. Universidad Piloto de Colombia.
- Apaiico, G. P. (2023). *GeomarPerales/ideamR* [R]. <https://github.com/GeomarPerales/ideamR> (Obra original publicada en 2021)
- Araya-Osses, D., Casanueva, A., Román-Figueroa, C., Uribe, J. M., & Paneque, M. (2020). Climate change projections of temperature and precipitation in Chile based on statistical downscaling. *Climate Dynamics*, 54(9), 4309–4330. <https://doi.org/10.1007/s00382-020-05231-4>
- ASALE, R.-, & RAE. (2024, enero 2). *Temperatura | Diccionario de la lengua española*. «Diccionario de la lengua española» - Edición del Tricentenario. <https://dle.rae.es/temperatura>
- Beck, H. E., Dijk, A. I. J. M. van, Larraondo, P. R., McVicar, T. R., Pan, M., Dutra, E., & Miralles, D. G. (2022). MSWX: Global 3-Hourly 0.1° Bias-Corrected Meteorological Data Including Near-Real-Time Updates and Forecast Ensembles. *Bulletin of the American Meteorological Society*, 103(3), E710–E732. <https://doi.org/10.1175/BAMS-D-21-0145.1>

- Bernal, L. A., Bernal, G., Cardona, O. D., Castaño-Castaño, A. N., Castillo-González, Y. C., Coca-Domínguez, O., Cuervo, Y. C., Mila, F. D., Díaz-Rentería, J. A., Echeverry, J., Escovar, M. A., Fernández, R., Carrillo, R. F., Rojas, A. F., Fuentes, G., Gutiérrez, W. O. G., García, A., García-Millán, N., M, C. A. H., ... Estacio, N. Z. (2021). *Investigaciones en gestión del riesgo de desastres para Colombia: Avances, perspectivas y casos de estudio*. Comisión Nacional Asesora de Investigación de Gestión del Riesgo de Desastres.
- Berndt, C., & Haberlandt, U. (2018). Spatial interpolation of climate variables in Northern Germany—Influence of temporal resolution and network density. *Journal of Hydrology: Regional Studies*, 15, 184–202. <https://doi.org/10.1016/j.ejrh.2018.02.002>
- Blanco, K. (2022). *Reducción de Escala Espacial de Datos Diarios de Precipitación y Temperatura para Estudios de Salud Pública en el Departamento del Cauca, Colombia* [Tesis documental de investigación]. UNIVERSIDAD INDUSTRIAL DE SANTANDER.
- Burbano, L. L., & Zarama, M. V. (1996). LA IMPORTANCIA DE LA HIDROLOGIA EN EL MANEJO DE CUENCAS HIDROGRAFICAS. *Revista de Ciencias Agrícolas*, 14(1 y 2), Article 1 y 2. <https://revistas.udenar.edu.co/index.php/rfacia/article/view/1163>
- Carrera-Villacrés, D. V., Guevara-García, P. V., Tamayo-Bacacela, L. C., Balarezo-Aguilar, A. L., Narváez-Rivera, C. A., & Morocho-López, D. R. (2016). Relleno de series anuales de datos meteorológicos mediante métodos estadísticos en la zona costera e interandina del Ecuador, y cálculo de la precipitación media. *Idesia (Arica)*, 34(3), 81–90. <https://doi.org/10.4067/S0718-34292016000300010>
- Chen, F., Gao, Y., Wang, Y., & Li, X. (2020). 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, 101324. <https://doi.org/10.1016/j.uclim.2022.101324>
- Cheng, L., & AghaKouchak, A. (2014). Nonstationary Precipitation Intensity-Duration-Frequency Curves for Infrastructure Design in a Changing Climate. *Scientific Reports*, 4(1), Article 1. <https://doi.org/10.1038/srep07093>
- Chilès, J.-P., & Delfiner, P. (2012). *Geostatistics: Modeling Spatial Uncertainty*. John Wiley & Sons, Incorporated. <http://ebookcentral.proquest.com/lib/bibliouis-ebooks/detail.action?docID=693179>
- Copernicus Climate Change Service. (2019a). *ERA5-Land hourly data from 2001 to present* [dataset]. ECMWF. <https://doi.org/10.24381/CDS.E2161BAC>
- Copernicus Climate Change Service. (2019b). *ERA5-Land hourly data from 2001 to present* [dataset]. ECMWF. <https://doi.org/10.24381/CDS.E2161BAC>
- dhime. (2024). *Consulta y Descarga de Datos Hidrometeorológicos*. <http://dhime.ideam.gov.co/atencionciudadano/>
- Dibike, Y. B., Gachon, P., St-Hilaire, A., Ouarda, T. B. M. J., & Nguyen, V. T.-V. (2008). Uncertainty analysis of statistically downscaled temperature and precipitation regimes in Northern Canada. *Theoretical and Applied Climatology*, 91(1), 149–170. <https://doi.org/10.1007/s00704-007-0299-z>
- Duan, Z., Tuo, Y., Liu, J., Gao, H., Song, X., Zhang, Z., Yang, L., & Mekonnen, D. F. (2019). Hydrological evaluation of open-access precipitation and air temperature datasets using SWAT in a poorly gauged basin in Ethiopia. *Journal of Hydrology*, 569, 612–626. <https://doi.org/10.1016/j.jhydrol.2018.12.026>

- Eguiguren-Velepucha, P. A., Chamba, J. A. M., Aguirre Mendoza, N. A., Ojeda-Luna, T. L., Samaniego-Rojas, N. S., Furniss, M. J., Howe, C., & Aguirre Mendoza, Z. H. (2016). Tropical ecosystems vulnerability to climate change in southern Ecuador. *Tropical Conservation Science*, 9(4), 1940082916668007. <https://doi.org/10.1177/1940082916668007>
- ERDDAP. (2023). *ERDDAP - CHIRPS Version 2.0, Precipitation, Global, 0.05°, Daily, 1981-present—Data Access Form*. <https://coastwatch.pfeg.noaa.gov/erddap/griddap/chirps20GlobalDailyP05.html>
- Fan, Y., & van den Dool, H. (2008a). A global monthly land surface air temperature analysis for 1948–present. *Journal of Geophysical Research: Atmospheres*, 113(D1). <https://doi.org/10.1029/2007JD008470>
- Fan, Y., & van den Dool, H. (2008b). A global monthly land surface air temperature analysis for 1948–present. *Journal of Geophysical Research: Atmospheres*, 113(D1). <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(1), Article 1. <https://doi.org/10.1038/sdata.2015.66>
- García Abad, J. (2012). *Calibración Local de Predicciones Numéricas de Viento con Técnicas Estadísticas no Lineales (Downscaling Estadístico)*.
- Gubler, S., Hunziker, S., Begert, M., Croci-Maspoli, M., Konzelmann, T., Brönnimann, S., Schwierz, C., Oria, C., & Rosas, G. (2017). The influence of station density on climate data

- homogenization. *International Journal of Climatology*, 37(13), 4670–4683.
<https://doi.org/10.1002/joc.5114>
- Gundogdu, K. S., & Guney, I. (2007). Spatial Analysis of Groundwater Levels Using Universal Kriging. *Journal of Earth System Science*, 116, 49–55. <https://doi.org/10.1007/s12040-007-0006-6>
- Hadi, S. J., & Tombul, M. (2018). Comparison of Spatial Interpolation Methods of Precipitation and Temperature Using Multiple Integration Periods. *Journal of the Indian Society of Remote Sensing*, 46(7), 1187–1199. <https://doi.org/10.1007/s12524-018-0783-1>
- Hortua, D. S. M., Calderón, M. D. H., & Espinosa, L. D. B. (2022). Viabilidad de uso de un sistema integrado de producción agropecuaria de tipo VAC como estrategia de mitigación de inundaciones para la cuenca alta del río Caquetá. *INVENTUM*, 17(32), Article 32. <https://doi.org/10.26620/uniminuto.inventum.17.32.2022.68-77>
- Hundecha, Y., Sunyer, M. A., Lawrence, D., Madsen, H., Willems, P., Bürger, G., Kriaučiūnienė, J., Loukas, A., Martinkova, M., Osuch, M., Vasiliades, L., von Christierson, B., Vormoor, K., & Yücel, I. (2016). Inter-comparison of statistical downscaling methods for projection of extreme flow indices across Europe. *Journal of Hydrology*, 541, 1273–1286. <https://doi.org/10.1016/j.jhydrol.2016.08.033>
- IDEAM. (2024, enero 15). *AGUAS SUPERFICIALES*. <http://www.ideam.gov.co/web/agua/aguas-superficiales>
- Jarkas, A. M., & Bitar, C. G. (2012). Factors Affecting Construction Labor Productivity in Kuwait. *Journal of Construction Engineering and Management*, 138(7), 811–820. [https://doi.org/10.1061/\(ASCE\)CO.1943-7862.0000501](https://doi.org/10.1061/(ASCE)CO.1943-7862.0000501)
- Jarraud, M., & Bokova, I. (2012). *International Glossary of Hydrology*.

- Jiménez-Ramos, G., Echaveguren, T., Vargas-Baecheler, J., & Chamorro, A. (2023). Traffic interruption risk induced by cut-slope failure: The rainfall effect. *Transportation Geotechnics*, *41*. Scopus. <https://doi.org/10.1016/j.trgeo.2023.100993>
- Khashei, M., & Bijari, M. (2010). An artificial neural network (p, d, q) model for timeseries forecasting. *Expert Syst. Appl.*, *37*, 479–489. <https://doi.org/10.1016/j.eswa.2009.05.044>
- Kusch, E., & Davy, R. (2022a). 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>
- Kusch, E., & Davy, R. (2022b). 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>
- Lala, L. (2020). *UNA REVISIÓN DE LOS MÉTODOS DE DOWNSCALING PARA LAS PROYECCIONES DE CAMBIO CLIMÁTICO RESILIENCIA AFRICANA Y LATINOAMERICANA AL CAMBIO CLIMÁTICO (ARCC) SEPTIEMBRE 2014*.
- Lanzante, J. R., Dixon, K. W., Nath, M. J., Whitlock, C. E., & Adams-Smith, D. (2018). Some Pitfalls in Statistical Downscaling of Future Climate. *Bulletin of the American Meteorological Society*, *99*(4), 791–803. <https://doi.org/10.1175/BAMS-D-17-0046.1>
- Li, Y., Dou, Q., Lu, Y., Xiang, H., Yu, X., & Liu, S. (2020). Effects of ambient temperature and precipitation on the risk of dengue fever: A systematic review and updated meta-analysis. *Environmental Research*, *191*, 110043. <https://doi.org/10.1016/j.envres.2020.110043>
- Liu, K., Wang, Q., Wang, M., & Koks, E. E. (2023). Global transportation infrastructure exposure to the change of precipitation in a warmer world. *Nature Communications*, *14*(1). Scopus. <https://doi.org/10.1038/s41467-023-38203-3>

- 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/articles/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, Parts A/B/C*, 127, 103184. <https://doi.org/10.1016/j.pce.2022.103184>
- Lu, Y., Fang, J., Pan, Y., Wang, S., Zhou, P., Yang, Y., Shao, M., & Tang, J. (2021). Evaluation of 12-Years Chinese Regional Reanalysis (1998–2009): Comparison of Dynamical Downscaling Methods With/Without Local Data. *Journal of Geophysical Research: Atmospheres*, 126(11). Scopus. <https://doi.org/10.1029/2020JD034259>
- Marin Mamani, G., Cahui Galarza, A., Gallegos Ramos, J. R., Bolívar Espinoza, N., & Enríquez Mamani, V. (2021). MODELAMIENTO KRIGING EN LA TASACIÓN DE PRECIOS PREDIALES URBANOS. *Investigación & Desarrollo*, 21(1), 109–118.
- Masello, L., Castignani, G., Sheehan, B., Guillen, M., & Murphy, F. (2023). Using contextual data to predict risky driving events: A novel methodology from explainable artificial intelligence. *Accident Analysis and Prevention*, 184. Scopus. <https://doi.org/10.1016/j.aap.2023.106997>
- Meier, I. C., & Leuschner, C. (2008). Leaf Size and Leaf Area Index in *Fagus sylvatica* Forests: Competing Effects of Precipitation, Temperature, and Nitrogen Availability. *Ecosystems*, 11(5), 655–669. <https://doi.org/10.1007/s10021-008-9135-2>

- Menebo, M. M. (2020). Temperature and precipitation associate with Covid-19 new daily cases: A correlation study between weather and Covid-19 pandemic in Oslo, Norway. *Science of The Total Environment*, 737, 139659. <https://doi.org/10.1016/j.scitotenv.2020.139659>
- Miller, N. L., Bashford, K. E., & Strem, E. (2003). Potential Impacts of Climate Change on California Hydrology. *JAWRA Journal of the American Water Resources Association*, 39(4), 771–784. <https://doi.org/10.1111/j.1752-1688.2003.tb04404.x>
- MSWX-GloH2O, M. (2021, febrero 18). *MSWX - GloH2O*. <https://www.gloh2o.org/mswx/>
- 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, 100092. <https://doi.org/10.1016/j.resenv.2022.100092>
- NOAA. (2021). *Gridded Climate: NOAA Physical Sciences Laboratory*. <https://www.psl.noaa.gov/data/gridded/>
- 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), 101739. <https://doi.org/10.1016/j.asej.2022.101739>
- 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->

- Ozturk, D., & Kilic, F. (2016). Geostatistical Approach for Spatial Interpolation of Meteorological Data. *Anais Da Academia Brasileira de Ciências*, 88, 2121–2136. <https://doi.org/10.1590/0001-3765201620150103>
- Park, S., Navratil, S., Gregory, A., Bauer, A., Srinath, I., Szonyi, B., Nightingale, K., Anciso, J., Jun, M., Han, D., Lawhon, S., Ivanek, R., & Griffiths, M. W. (2015, marzo 12). Multifactorial Effects of Ambient Temperature, Precipitation, Farm Management, and Environmental Factors Determine the Level of Generic Escherichia coli Contamination on Preharvested Spinach. *Applied and Environmental Microbiology*, 81(7), 2635–2650. <https://doi.org/10.1128/AEM.03793-14>
- Posada-Marín, J. A., Rendón, A. M., Salazar, J. F., Mejía, J. F., & Villegas, J. C. (2019). WRF downscaling improves ERA-Interim representation of precipitation around a tropical Andean valley during El Niño: Implications for GCM-scale simulation of precipitation over complex terrain. *Climate Dynamics*, 52(5), 3609–3629. <https://doi.org/10.1007/s00382-018-4403-0>
- Punia, A. (2021). Role of temperature, wind, and precipitation in heavy metal contamination at copper mines: A review. *Environmental Science and Pollution Research*, 28(4), 4056–4072. <https://doi.org/10.1007/s11356-020-11580-8>
- Río San José, J. del. (2010). *Tratamiento de datos espaciales en hidrología*. Bubok Publishing.
- Serrano-Notivoli, R., & Tejedor, E. (2021). From rain to data: A review of the creation of monthly and daily station-based gridded precipitation datasets. *WIREs Water*, 8(6), e1555. <https://doi.org/10.1002/wat2.1555>

- Sigalla, O. Z., Valimba, P., Selemani, J. R., Kashaigili, J. J., & Tumbo, M. (2023). Analysis of spatial and temporal trend of hydro-climatic parameters in the Kilombero River Catchment, Tanzania. *Scientific Reports*, 13(1), 7864. <https://doi.org/10.1038/s41598-023-35105-8>
- Siles Soria, G. (2022). REANÁLISIS CLIMATOLÓGICO ERA5: UNA REVISIÓN SOBRE SU USO EN EL CÁLCULO DE ATENUACIÓN ATMOSFÉRICA EN SISTEMAS DE COMUNICACIONES SATELITALES. *Investigación & Desarrollo*, 22(1), 145–159. <https://doi.org/10.23881/idupbo.022.1-12i>
- Souvignet, M., & Heinrich, J. (2011). Statistical downscaling in the arid central Andes: Uncertainty analysis of multi-model simulated temperature and precipitation. *Theoretical and Applied Climatology*, 106(1), 229–244. <https://doi.org/10.1007/s00704-011-0430-z>
- Sun, Y., Wendi, D., Kim, D. E., & Liang, S.-Y. (2019). Deriving intensity–duration–frequency (IDF) curves using downscaled in situ rainfall assimilated with remote sensing data. *Geoscience Letters*, 6(1), 17. <https://doi.org/10.1186/s40562-019-0147-x>
- Sunyer, M. A., Hundedcha, Y., Lawrence, D., Madsen, H., Willems, P., Martinkova, M., Vormoor, K., Bürger, G., Hanel, M., Kriaučiūnienė, J., Loukas, A., Osuch, M., & Yücel, I. (2015). Inter-comparison of statistical downscaling methods for projection of extreme precipitation in Europe. *Hydrology and Earth System Sciences*, 19(4), 1827–1847. <https://doi.org/10.5194/hess-19-1827-2015>
- Tamerius, J. D., Zhou, X., Mantilla, R., & Greenfield-Huitt, T. (2016). Precipitation Effects on Motor Vehicle Crashes Vary by Space, Time, and Environmental Conditions. *Weather, Climate, and Society*, 8(4), 399–407. <https://doi.org/10.1175/WCAS-D-16-0009.1>
- TessaDem, & OpenStreetMap. (2023). *Mapa topográfico Santander, altitud, relieve*. Mapas topográficos. <https://es-co.topographic-map.com/map->

flcf3/Santander/?center=6.75462%2C-72.79266&zoom=9&popup=7.02185%2C-72.77756

- Tuel, A., & El Moçayd, N. (2023). Evaluating extreme precipitation in gridded datasets with a novel station database in Morocco. *Stochastic Environmental Research and Risk Assessment*. Scopus. <https://doi.org/10.1007/s00477-023-02437-w>
- 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>
- Valencia, S., Marín, D. E., Gómez, D., Hoyos, N., Salazar, J. F., & Villegas, J. C. (2023). Spatio-temporal assessment of Gridded precipitation products across topographic and climatic gradients in Colombia. *Atmospheric Research*, 285, 106643. <https://doi.org/10.1016/j.atmosres.2023.106643>
- Vano, J. A., Das, T., & Lettenmaier, D. P. (2012). Hydrologic Sensitivities of Colorado River Runoff to Changes in Precipitation and Temperature. *Journal of Hydrometeorology*, 13(3), 932–949. <https://doi.org/10.1175/JHM-D-11-069.1>
- Vargas Godoy, M. R., & Markonis, Y. (2023). pRecipe: A global precipitation climatology toolbox and database. *Environmental Modelling & Software*, 165, 105711. <https://doi.org/10.1016/j.envsoft.2023.105711>
- Vianna, L. F. N., de Souza, R. V., Schramm, M. A., & Alves, T. P. (2023). Using climate reanalysis and remote sensing-derived data to create the basis for predicting the occurrence of algal blooms, harmful algal blooms and toxic events in Santa Catarina, Brazil. *Science of the Total Environment*, 880. Scopus. <https://doi.org/10.1016/j.scitotenv.2023.163086>
- 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), 52.
<https://doi.org/10.1186/s40249-022-00980-2>
- Werner, A. T., & Cannon, A. J. (2016). Hydrologic extremes – an intercomparison of multiple gridded statistical downscaling methods. *Hydrology and Earth System Sciences*, 20(4), 1483–1508. <https://doi.org/10.5194/hess-20-1483-2016>
- Wu, H., Chen, B., Snelgrove, K., & Lye, L. (2016). Quantification of Uncertainty Propagation Effects during Statistical Downscaling of Precipitation and Temperature to Hydrological Modeling. *Journal of Environmental Informatics*, 2, 139–148.
<https://doi.org/10.3808/jei.201600347>
- 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), 04022009.
[https://doi.org/10.1061/\(ASCE\)CR.1943-5495.0000286](https://doi.org/10.1061/(ASCE)CR.1943-5495.0000286)
- Wu, Z., Dijkstra, P., Koch, G. W., Peñuelas, J., & Hungate, B. A. (2011). Responses of terrestrial ecosystems to temperature and precipitation change: A meta-analysis of experimental manipulation. *Global Change Biology*, 17(2), 927–942. <https://doi.org/10.1111/j.1365-2486.2010.02302.x>
- Yan, N., & Baas, A. C. W. (2015). Parabolic dunes and their transformations under environmental and climatic changes: Towards a conceptual framework for understanding and prediction. *Global and Planetary Change*, 124, 123–148.
<https://doi.org/10.1016/j.gloplacha.2014.11.010>