



Evaluación de las diferencias entre datos de precipitación del IDEAM y sensores satelitales en la cuenca media del río Lebrija, Colombia.

**Camilo Enrique Mojica Gutiérrez
Fredy Vladimir Jones Navas**

Universidad Santo Tomás
Bogotá D.C., Colombia
Año 2025

Evaluación de las diferencias entre datos de precipitación del IDEAM y sensores satelitales en la cuenca media del río Lebrija, Colombia.

**Camilo Enrique Mojica Gutiérrez y
Fredy Vladimir Jones Navas**

Trabajo de grado presentada(o) como requisito parcial para optar al título de:
Master en Gestión de cuencas Hidrográficas.

Director (a):

Elizabeth Flórez Carvajal P.hD

Línea de Investigación:

Gestión del riesgo y adaptación al cambio climático

Grupo de Investigación:

Centro de altos estudios en cuencas hidrográficas – CAECHI

Universidad Santo Tomás

Bogotá D.C., Colombia

Año 2025

(Dedicatoria o lema)

Camilo Enrique Mojica Gutiérrez

Agradezco a Dios, por permitirme finalizar esta tesis, a mi familia que me ha acompañado en mis momentos difíciles y en especial a mis hijos que son mi motor y mi impulso, para ellos a quienes con el simple ejemplo de este documento quiero transmitirles el valor del sacrificio que trae bondad en las cosas cotidianas.

Fredy Vladimir Jones Navas

A mi familia, por su paciencia, apoyo incondicional y palabras de aliento en los momentos de más cansancio. A mis padres, por enseñarme el valor del esfuerzo y la constancia. A quienes, con su compañía silenciosa, hicieron posible que este camino no se sintiera tan largo. Esta tesis es también de ustedes.

Agradecimientos

Con este documento queremos agradecerle a la profesora Elizabeth por su apoyo este tiempo, a la universidad por la excelencia académica en el programa, a nuestras familias y allegados que de manera directa han estado presentes en este proceso y a Dios porque nos ha permitido realizar esta investigación, de manera consciente y oportuna para mostrar el potencial técnico que tenemos en nuestro país en beneficio de ofrecer una mejor desarrollo eficiente, climático y ambiental.

Han sido largas jornadas dedicadas a realizar esta investigación, con el resultado queremos compartir la experiencia de implementar nuevas técnicas en el desarrollo de estudios hidrológicos, climatológicos y ambientales, deseamos que sea mucha utilidad para el desarrollo de nuevas investigaciones y que como resultado los lectores puedan determinar si esas herramientas son válidas para implementar en sus estudios cotidianos.

Resumen

Los estudios hidro-climatológicos en diferentes cuencas necesitan bases de datos de variables climáticas. El IDEAM es la entidad principal que las monitorea en Colombia, con datos oficiales, libres y gratuitos. También existen otras entidades locales e internacionales que ofrecen información proveniente de sensores satelitales de teledetección.

Los datos de estaciones climatológicas presentan, en algunos casos, incertidumbres por factores externos y, además, su distribución es limitada en ciertas zonas. Frente a esta problemática, se desarrolló la presente investigación en la cuenca media del río Lebrija, utilizando tres fuentes de información para comparar las muestras.

Se recolectaron datos de precipitación de 25 estaciones del IDEAM, de las cuales se usaron 6, y se compararon con datos de sensores satelitales tomados de los portales The POWER Data Access Viewer y Giovanni de EARTH DATA. La información fue revisada, depurada y completada. Además, se aplicaron análisis estadísticos de tendencias, variabilidad, correlaciones, validación cruzada, pruebas de significación y análisis de probabilidades.

Los resultados muestran que, en zonas sin estaciones, los sensores permiten contar con datos continuos; sin embargo, al compararlos con los del IDEAM, las correlaciones son bajas en la mayoría de los casos. Esto sugiere que, aunque los productos satelitales son útiles en áreas sin cobertura, la información de estaciones sigue mostrando mayor coherencia y mejor ajuste estadístico, por lo que debe considerarse como fuente principal cuando esté disponible.

Palabras claves: Precipitación, hidrología, cuenca, estaciones, sensores satelitales.

Abstract

Hydro-climatological studies in different basins require databases of climatic variables. IDEAM is the main entity monitoring these variables in Colombia, providing official, free, and accessible data. There are also other local and international entities that offer information from remote sensing satellite sensors.

Climatological station data sometimes presents uncertainties due to external factors, and its distribution is limited in certain areas. To address this issue, the present research was developed in the middle basin of the Lebrija River, using three sources of information to compare the samples.

Precipitation data from 25 IDEAM stations was collected, of which 6 were used, and compared with data from satellite sensors obtained from The POWER Data Access Viewer and Giovanni portals of EARTH DATA. The information was reviewed, refined, and completed. What is more, statistical analyses of trends, variability, correlations, cross-validation, significance tests, and probability analysis were applied.

The results show that, in areas without stations, sensors provide continuous data; however, when compared to IDEAM data, the correlations are low in most cases. This suggests that although satellite products are useful in areas without coverage, station data still shows greater coherence and better statistical fit, making it the primary source when available.

Keywords: Precipitation, hydrology, basin, stations, satellite sensors.

Tabla de contenido

Introducción	5
1 Planteamiento del problema	11
1.1 Pregunta problema de investigación	14
1.2 Justificación	15
2 Objetivos.....	17
2.1 Objetivo general.....	17
2.2 Objetivos específicos.....	17
3 Marco referencial.	18
3.1 Localización de la cuenca media hidrográfica del río Lebrija Medio.	26
3.2 Factores económicos de la cuenca media hidrográfica del río Lebrija.	28
3.3 Usos de suelo de la Cuenca Media hidrográfica del Río Lebrija.	29
3.4 Factores climatológicos de la cuenca media hidrográfica del río Lebrija.	30
3.5 Datos climatológicos de precipitación.....	32
3.6 Fuentes de información de datos variable de precipitación para la cuenca del río Lebrija medio	34
4 Marco teórico y conceptual.	40
4.1 ¿Qué es el ciclo hidrológico?.....	40
4.2 Cuenca Hidrográfica	43
4.3 Variables Climatológicas	43
4.3.1 Precipitación	44
4.3.2 Instrumentos de medición	44
4.4 Estudios Hidrológicos	48
4.5 Bases teóricas Hidrológicas	49
5 Estado del arte	51
5.1 Antecedentes.....	51
5.2 Estudios locales.....	53
5.3 Estudios nacionales.....	54
5.4 Estudios internacionales.....	56
6 Aspectos metodológicos	58
6.1 Paradigma de investigación.....	58
6.2 Enfoque de la investigación.....	59
6.3 Temporalidad.....	60

6.4	Análisis de datos.....	60
6.5	Metodología propuesta.	62
6.6	Selección de información de las estaciones climatológicas que se encuentren localizadas en el área de la cuenca media del Río Lebrija.	63
6.6.1	Identificación de estaciones climatológicas.....	64
6.6.2	Definición y llenado de los datos faltantes	66
6.7	Contrastar la información disponible de las tres fuentes seleccionadas	70
6.7.1	Análisis de estadística descriptiva de datos del IDEAM completos y de la NASA. 71	
6.7.2	Análisis de correlación entre las fuentes de información	73
6.7.3	Modelo de predicción y estudio de tendencias.....	76
7	Resultados y discusión	79
7.1	Resultado de Identificación de estaciones climatológicas.	79
7.2	Datos de sensores satelitales implementados.	83
7.3	Comparación de series temporales y validación estadística	83
7.3.1	Evaluación del Sensor GIOVANNI frente a Estaciones de Referencia (datos diarios) 83	
7.3.2	Comparación entre PowerData y datos observados del IDEAM (datos diarios) 87	
7.3.3	Análisis mensual de precipitación: IDEAM vs Giovanni	89
7.3.4	Análisis mensual de precipitación: IDEAM vs PowerData.....	91
7.4	Recomendaciones según los resultados obtenidos	93
8	Conclusiones	95
9	Referencias bibliográficas	98
10	Anexos	101
	Anexo 1: Análisis comparativo datos a nivel diario estación Caobo.	101
	Anexo 2: Análisis comparativo datos a nivel diario estación Escuela Agrícola Cachirá.	105
	Anexo 3: Análisis comparativo datos a nivel diario estación Porvenir.....	109
	Anexo 4: Análisis comparativo datos a nivel diario estación Provincia	113
	Anexo 5: Análisis comparativo datos a nivel diario estación San Rafael.	117
	Anexo 6: Análisis comparativo datos a nivel diario estación La Vega.....	121
	Anexo 7: Análisis comparativo datos a nivel Mensual estación Caobo.....	125

Anexo 8: Análisis comparativo datos a nivel Mensual estación Escuela Agrícola Cachirá.	127
Anexo 9: Análisis comparativo datos a nivel Mensual estación Porvenir.	130
Anexo 10: Análisis comparativo datos a nivel Mensual estación Provincia.	132
Anexo 11: Análisis comparativo datos a nivel Mensual estación San Rafael.	135
Anexo 12: Análisis comparativo datos a nivel Mensual estación La Vega.....	138
Anexo 13: Resultado análisis SARIMAX	141

Lista de figuras

Figura 1 Localización Cuenca medía río Lebrija y sus límites con los municipios que atraviesa.	27
Figura 2 Localización Cuenca medía río Lebrija con red hidrográfica.	28
Figura 3 Valores totales mensuales de precipitación.	32
Figura 4 Ilustración del ciclo Hidrológico.	42
Figura 5 Ilustración del pluviómetro y registro de agua en estaciones del IDEAM.	46
Figura 6 Ilustración del pluviógrafo y extracción del tambor del pluviógrafo.	47
Figura 7 Ilustración del pluviógrafo y extracción del tambor del pluviógrafo.	48
Figura 8 Localización de las estaciones climatológicas de Cuenca medía río Lebrija.	80
Figura 9 Localización de las estaciones climatológicas de Cuenca medía río Lebrija.	82

Lista de tablas

Tabla 1 Información descargada de las estaciones existentes en la cuenca.....	81
Tabla 2 Información Seleccionada para el estudio de las estaciones existentes en la cuenca .	83
Tabla 3 Indicadores estadísticos de comparación entre datos diarios del sensor Giovanni y estaciones IDEAM.	87
Tabla 4 Indicadores estadísticos de comparación entre datos diarios de PowerData y estaciones IDEAM.	88
Tabla 5 Distribuciones probabilísticas comparados datos IDEAM vs datos Giovanni diario estación Caobo.	102
Tabla 6 Distribuciones probabilísticas comparados datos IDEAM vs datos Power Data diario estación Caobo.	104
Tabla 7 Distribuciones probabilísticas comparados datos IDEAM vs datos Giovanni diario estación Escuela Agrícola Cachirá.	106
Tabla 8 Distribuciones probabilísticas comparados datos IDEAM vs datos Power Data diario estación Escuela Agrícola Cachirá.	108
Tabla 9 Distribuciones probabilísticas comparados datos IDEAM vs datos Giovanni diario estación Porvenir..	110
Tabla 10 Distribuciones probabilísticas comparados datos IDEAM vs datos Power Data diario estación Porvenir.	112
Tabla 11 Distribuciones probabilísticas comparados datos IDEAM vs datos Giovanni diario estación Provincia.....	114
Tabla 12 Distribuciones probabilísticas comparados datos IDEAM vs datos Power Data diario estación Provincia.....	116
Tabla 13 Distribuciones probabilísticas comparados datos IDEAM vs datos Giovanni diario estación San Rafael.	118
Tabla 14 Distribuciones probabilísticas comparados datos IDEAM vs datos Power Data diario estación San Rafael.	120
Tabla 15 Distribuciones probabilísticas comparados datos IDEAM vs datos Giovanni diario estación La Vega.	122
Tabla 16 Distribuciones probabilísticas comparados datos IDEAM vs datos Power Data diario estación La Vega.	124
Tabla 17 Valores Medios Mensuales de precipitación IDEAM estación Caobo..	125
Tabla 18 Valores Medios Mensuales de precipitación Giovanni estación Caobo.	125
Tabla 19 Valores Medios Mensuales de precipitación Power Data estación Caobo.....	126
Tabla 20 Valores Medios Mensuales de precipitación IDEAM estación Escuela Agrícola Cachirá.....	127
Tabla 21 Valores Medios Mensuales de precipitación Giovanni estación Escuela Agrícola Cachirá.....	128
Tabla 22 Valores Medios Mensuales de precipitación Power Data estación Escuela Agrícola Cachirá.....	128
Tabla 23 Valores Medios Mensuales de precipitación IDEAM estación Porvenir.	130
Tabla 24 Valores Medios Mensuales de precipitación Giovanni estación Porvenir.	130
Tabla 25 Valores Medios Mensuales de precipitación Power Data estación Porvenir.....	131

Tabla 26	Valores Medios Mensuales de precipitación IDEAM estación Provincia.....	132
Tabla 27	Valores Medios Mensuales de precipitación Giovanni estación Provincia.....	133
Tabla 28	Valores Medios Mensuales de precipitación Power Data estación Provincia.	133
Tabla 29	Valores Medios Mensuales de precipitación IDEAM estación San Rafael.....	135
Tabla 30	Valores Medios Mensuales de precipitación Giovanni estación San Rafael.	136
Tabla 31	Valores Medios Mensuales de precipitación Power Data estación San Rafael.	136
Tabla 32	Valores Medios Mensuales de precipitación IDEAM estación La Vega	138
Tabla 33	Valores Medios Mensuales de precipitación Giovanni estación Provincia.	138
Tabla 34	Valores Medios Mensuales de precipitación Power Data estación Provincia.	139

Lista de Anexos.

Anexo A	Análisis comparativo datos IDEAM vs datos Giovanni diario estación Caobo.	101
Anexo B	Análisis comparativo de dispersión datos IDEAM vs datos Giovanni diario estación Caobo.	101
Anexo C	Análisis comparativo de distribución de precipitación datos IDEAM vs datos Giovanni diario estación Caobo.	102
Anexo D	Análisis comparativo de distribución de precipitación datos IDEAM vs datos Giovanni diario estación Caobo.	102
Anexo E	Análisis comparativo datos IDEAM vs datos Power Data diario estación Caobo....	103
Anexo F	Análisis comparativo de dispersión datos IDEAM vs datos Power Data i diario estación Caobo.	103
Anexo G	Análisis comparativo de distribución de precipitación datos IDEAM vs datos Power Data diario estación Caobo.....	104
Anexo H	Análisis comparativo de distribución de precipitación datos IDEAM vs datos Power Data diario estación Caobo.....	104
Anexo I	Análisis comparativo datos IDEAM vs datos Giovanni diario estación Escuela Agrícola Cachirá.....	105
Anexo J	Análisis comparativo de dispersión datos IDEAM vs datos Giovanni diario estación Escuela Agrícola Cachirá.....	105
Anexo K	Análisis comparativo de distribución de precipitación datos IDEAM vs datos Giovanni diario estación Escuela Agrícola Cachirá.	106
Anexo L	Análisis comparativo de distribución de precipitación datos IDEAM vs datos Giovanni diario estación Escuela Agrícola Cachirá.	106
Anexo M	Análisis comparativo datos IDEAM vs datos Power Data diario estación Escuela Agrícola Cachirá.	107
Anexo N	Análisis comparativo de dispersión datos IDEAM vs datos Power Data diario estación Escuela Agrícola Cachirá.	107
Anexo O	Análisis comparativo de distribución de precipitación datos IDEAM vs datos Power Data diario estación Escuela Agrícola Cachirá.	108
Anexo P	Análisis comparativo de distribución de precipitación datos IDEAM vs datos Power Data diario estación Escuela Agrícola Cachirá.	108
Anexo Q	Análisis comparativo datos IDEAM vs datos Giovanni diario estación Porvenir....	109
Anexo R	Análisis comparativo de dispersión datos IDEAM vs datos Giovanni diario estación Porvenir.....	109
Anexo S	Análisis comparativo de distribución de precipitación datos IDEAM vs datos Giovanni diario estación Porvenir.	110
Anexo T	Análisis comparativo de distribución de precipitación datos IDEAM vs datos Giovanni diario estación Porvenir..	110
Anexo U	Análisis comparativo datos IDEAM vs datos Power Data diario estación Porvenir.	111
Anexo V	Análisis comparativo de dispersión datos IDEAM vs datos Power Data diario estación Porvenir.	111

Anexo W Análisis comparativo de distribución de precipitación datos IDEAM vs datos Power Data diario estación Porvenir.....	112
Anexo X Análisis comparativo de distribución de precipitación datos IDEAM vs datos Power Data diario estación Porvenir.....	112
Anexo Y Análisis comparativo datos IDEAM vs datos Giovanni diario estación Provincia....	113
Anexo Z Análisis comparativo de dispersión datos IDEAM vs datos Giovanni diario estación Provincia.....	113
Anexo AA Análisis comparativo de distribución de precipitación datos IDEAM vs datos Giovanni diario estación Provincia.....	114
Anexo BB Análisis comparativo de distribución de precipitación datos IDEAM vs datos Giovanni diario estación Provincia.....	114
Anexo CC Análisis comparativo datos IDEAM vs datos Power Data diario estación Provincia.....	115
Anexo DD Análisis comparativo de dispersión datos IDEAM vs datos Power Data diario estación Provincia.....	115
Anexo EE Análisis comparativo de distribución de precipitación datos IDEAM vs datos Power Data diario estación Provincia.....	116
Anexo FF Análisis comparativo de distribución de precipitación datos IDEAM vs datos Power Data diario estación Provincia.....	116
Anexo GG Análisis comparativo datos IDEAM vs datos Giovanni diario estación San Rafael.....	117
Anexo HH Análisis comparativo de dispersión datos IDEAM vs datos Giovanni diario estación San Rafael.....	117
Anexo II Análisis comparativo de distribución de precipitación datos IDEAM vs datos Giovanni diario estación San Rafael.....	118
Anexo JJ Análisis comparativo de distribución de precipitación datos IDEAM vs datos Giovanni diario estación San Rafael.....	118
Anexo KK Análisis comparativo datos IDEAM vs datos Power Data diario estación San Rafael.....	119
Anexo LL Análisis comparativo de dispersión datos IDEAM vs datos Power Data diario estación San Rafael.....	119
Anexo MM Análisis comparativo de distribución de precipitación datos IDEAM vs datos Power Data diario estación San Rafael.....	120
Anexo NN Análisis comparativo de distribución de precipitación datos IDEAM vs datos Power Data diario estación San Rafael.....	120
Anexo OO Análisis comparativo datos IDEAM vs datos Giovanni diario estación La Vega .	121
Anexo PP Análisis comparativo de dispersión datos IDEAM vs datos Giovanni diario estación La Vega.....	121
Anexo QQ Análisis comparativo de distribución de precipitación datos IDEAM vs datos Giovanni diario estación La Vega.....	122
Anexo RR Análisis comparativo de distribución de precipitación datos IDEAM vs datos Giovanni diario estación La Vega.....	122
Anexo SS Análisis comparativo datos IDEAM vs datos Power Data diario estación La Vega.....	123

Anexo TT Análisis comparativo de dispersión datos IDEAM vs datos Power Data diario estación La Vega.	123
Anexo UU Análisis comparativo de distribución de precipitación datos IDEAM vs datos Power Data diario estación La Vega.	124
Anexo VV Análisis comparativo de distribución de precipitación datos IDEAM vs datos Power Data diario estación La Vega.	124
Anexo WW Análisis comparativo datos IDEAM vs datos Giovanni Mensual estación Caobo.	126
Anexo XX Análisis comparativo datos IDEAM vs datos Power Data Mensual estación Caobo.	127
Anexo YY Análisis comparativo datos IDEAM vs datos Giovanni diario estación Escuela Agrícola Cachirá.	129
Anexo ZZ Análisis comparativo datos IDEAM vs datos Power Data diario estación Escuela Agrícola Cachirá.	129
Anexo AAA Análisis comparativo datos IDEAM vs datos Giovanni diario estación Porvenir.	131
Anexo BBB Análisis comparativo datos IDEAM vs datos Power Data diario estación Porvenir.	132
Anexo CCC Análisis comparativo datos IDEAM vs datos Giovanni diario estación Provincia.	134
Anexo DDD Análisis comparativo datos IDEAM vs datos Power Data diario estación Provincia.	135
Anexo EEE Análisis comparativo datos IDEAM vs datos Giovanni diario estación San Rafael.	137
Anexo FFF Análisis comparativo datos IDEAM vs datos Power Data diario estación San Rafael.	137
Anexo GGG Análisis comparativo datos IDEAM vs datos Giovanni diario estación Provincia.	140
Anexo HHH Análisis comparativo datos IDEAM vs datos Power Data diario estación Provincia.	140
Anexo III Análisis de modelo estadístico de SAIMAX Giovanni vs IDEAM estación Caobo..	141
Anexo JJJ Análisis de modelo estadístico de SAIMAX PoweData vs IDEAM estación Caobo..	141
Anexo KKK Análisis de modelo estadístico de SAIMAX Giovanni vs IDEAM estación Escuela Agrícola Cachirá..	142
Anexo LLL Análisis de modelo estadístico de SAIMAX PoweData vs IDEAM estación Escuela Agrícola Cachirá..	142
Anexo MMM Análisis de modelo estadístico de SAIMAX Giovanni vs IDEAM estación Porvenir.	143
Anexo NNN Análisis de modelo estadístico de SAIMAX PoweData vs IDEAM estación Porvenir.	143
Anexo OOO Análisis de modelo estadístico de SAIMAX Giovanni vs IDEAM estación Provincia..	144

Anexo PPP Análisis de modelo estadístico de SAIMAX PoweData vs IDEAM estación Provincia..	144
Anexo QQQ Análisis de modelo estadístico de SAIMAX Giovanni vs IDEAM estación San Rafael.....	145
Anexo RRR Análisis de modelo estadístico de SAIMAX PoweData vs IDEAM estación San Rafael.....	145
Anexo SSS Análisis de modelo estadístico de SAIMAX Giovanni vs IDEAM estación Vega.	146
Anexo TTT Análisis de modelo estadístico de SAIMAX PoweData vs IDEAM estación Vega.	146

Introducción

El territorio colombiano cuenta con mecanismos de medición y control de eventos meteorológicos en diferentes zonas de su territorio, dicha información es utilizada para el manejo y alimentación de bases de datos que son usadas para la elaboración de diferentes estudios. Estos mecanismos se extienden en el área de influencia de diferentes cuencas hidrográficas, tal es el caso de la cuenca media del río Lebrija, la cual tiene instrumentadas estaciones climatológicas que recopilan datos sobre variables meteorológicas como precipitación, temperatura del aire, humedad relativa, brillo solar, nubosidad, radiación solar, evaporación y velocidad del viento, entre otras, con el objetivo de proporcionar datos utilizados en estudios hidro-climatológicos.

Las estaciones climatológicas en el territorio colombiano son administradas por diferentes entidades gubernamentales públicas y privadas, en particular, el Instituto de Hidrología, Meteorología y Estudios Ambientales – IDEAM, es la principal fuente de información con datos disponibles para el desarrollo de estudios climatológicos y ambientales, dentro de su catálogo disponen de estaciones agrometeorológica, climatológica ordinaria, climatológica principal, limnigráfica, limnimétrica, mareográfica, meteorológica especial, pluviográfica, pluviométrica, radio sonda, sinóptica principal y sinóptica secundaria.

La medición de parámetros meteorológicos se hace de manera manual o automática, en una periodicidad de tiempo diario con un rango de horario específico y los datos se pueden descargar del portal de datos abiertos DHIME (Datos Hidrológicos y Meteorológicos) del IDEAM dada la disponibilidad de años y variables que cuentan con registros.

Los datos meteorológicos son de vital importancia para el territorio colombiano, por cuanto permiten entender, la dinámica climatológica que se produce en la región por su localización en el continente Sur Americano. De acuerdo con Poveda (2004) las escalas

temporales de la variabilidad hidro-climatología en el territorio colombiano, son Escala Inter-Decadal, la cual está dominada por el cambio climático y cambio ambiental, Escala Interanual El Niño/Oscilación del Sur (ENSO), tiene una variabilidad estacional cálida con aumento de las temperaturas superficiales del mar, escala Anual dominada por los fenómenos físicos con una dinámica trans-ecuatorial, escala Intra-anual donde los fenómenos que mayor influencia ejercen en Colombia son oscilaciones de 30-60 días y escala Diurna resultante del ciclo de insolación superficial, el cual es un mecanismo muy importante para entender el génesis de la lluvia tropical.

El IDEAM durante varias décadas ha administrado el manejo de las bases de datos de parámetros climatológicos de las variables mencionadas, su acceso a la información en años anteriores era restringida y se debía solicitar de manera oficial a la entidad mediante algunos mecanismos establecidos para tal fin, en la actualidad existe el Sistema de Información para la Gestión de Datos Hidrológicos y Meteorológicos DHIME de dicha entidad, el cual es un portal de base de datos abierto en el que se puede acceder de manera gratuita y virtual, además solicitar la información geoespacial para diferentes estaciones climatológicas en periodos multianual, diario o mensual.

Para Colombia el DHIME se convierte en la principal fuente de información de diferentes variables climatológicas en zonas establecidas en su territorio, las cuales se implementan en el desarrollo de estudios hidrológicos, hidráulicos, climatológicos y ambientales, sin embargo, existen diferentes entidades como Corporaciones Autónomas Regionales e instituciones educativas o privadas que de manera oficial, también toman muestras y el acceso a dicha información en algunos casos es restringida.

Por lo anterior, se considera que los datos del IDEAM son una fuente de información primaria que asegura una confiabilidad en las muestras obtenidas, sin embargo, existen

muchos problemas asociados a la disponibilidad de estaciones en algunas zonas donde no se han instrumentado estaciones climatológicas y los datos al ser tomados de manera manual contienen una incertidumbre que se produce por errores humanos involuntarios o falta de información levantada.

En la actualidad existen a nivel global portales web de bases de datos abiertos de parámetros climatológicos, los cuales son usados para descargar información de diferentes variables, éstos usan sensores satelitales que son monitoreados diariamente y se pueden localizar espacialmente en muchas zonas del planeta tierra. Esto sin duda se convierte en una solución para la búsqueda de información meteorológica en zonas que no cuentan con estaciones climatológicas y permiten identificar las condiciones climáticas a nivel regional para áreas determinadas.

Con el desarrollo tecnológico y la disponibilidad actual de diferentes fuentes de información para recolección de datos climatológicos, se desea realizar la caracterización de portales de datos abiertos de sensores satelitales que se puedan implementar en la presente investigación y analizar cuál es el comportamiento, su correlación entre los datos obtenidos y compararlos con la información existente de diferentes estaciones climatológicas del IDEAM, determinando si es pertinente usar estas bases de datos en cuencas similares a la cuenca de estudio.

En esta investigación se seleccionó una cuenca de interés dentro de la cual previamente se hizo la verificación de la existencia de gran cantidad de estaciones climatológicas, que contaran con datos de la variable de precipitación, una vez realizada la verificación se determinó que la cuenca medía del río Lebrija, cumplía con los parámetros para realizar un estudio pertinente.

Para lograr desarrollar el análisis hidrológico de la cuenca en estudio, se ha realizado la búsqueda de portales web de bases de datos abiertos de parámetros climatológicos derivados de modelos de sensores satelitales que tengan disponibilidad de información en el territorio colombiano y que se hayan usado en estudios de caso para cuencas hidrográficas similares en la región latinoamericana, cabe mencionar que en Colombia la mayor parte de la información disponible ha sido levantada de manera oficial mediante estaciones climatológicas instauradas en diferentes zonas a lo largo de su territorio.

En la región latinoamericana y específicamente en el territorio Colombiano existen diferentes sensores satelitales que se han instrumentado para monitorear el medio ambiente, cambio climático, la agricultura y la seguridad, los cuales han sido lanzados gracias a diferentes convenios y colaboración internacional, actualmente se tienen los satélites FACSAT 1 (Fábrica de Satélites Colombianos) lanzado en el año 2018 y FACSAT 2 lanzado en el año 2022, los cuales fueron desarrollados por Agencia Espacial Colombiana (CANSAT) en colaboración de la NASA (National Aeronautics and Space Administration) y la Agencia Espacial Europea (ESA), según Rodríguez et al., (2020) en su artículo manifiesta la importancia de las observaciones de los sensores lanzados por la CANSAT en el FACSAT 1 y especifica el análisis del segmento espacial como parte del proceso de estructuración de un proyecto aeroespacial que se desarrolla en Colombia, por otra parte, Cárdenas et al., (2024) indica que tras los éxitos de los lanzamientos de los FACSAT 1 y FACSAT 2 se desarrolla un proceso para una misión satelital FACSAT 3 que busca posicionar un tercer satélite en la órbita terrestre baja en el territorio Colombiano..

Los datos de las misiones Colombianas son operados por la Fuerza Aérea Colombiana, sin embargo, su divulgación no se hace de manera abierta en portales nacionales diferentes al IDEAM, con lo anterior se hizo la verificación de que portales de otros países cuentan con

información de datos de sensores satelitales abierta y de fácil acceso, que permite extraer datos para estudios de caso en cuencas localizadas en el territorio colombiano, basado en lo anterior se logró determinar que el portal web The POWER Data Access Viewer y Giovanni de la EARTH DATA ambos de la NASA del departamento de Prediction of Worldwide Energy Resources, cuentan con información disponible para el análisis indicado.

The POWER Data Access Viewer permite extraer datos de parámetros climatológicos, de manera diaria, mensual y anual, de modelos satelitales, ideales para estudios de impacto climatológico, energía solar, agrícola e ingeniería etc., por otra parte Giovanni Geospatial Interactive Online Visualization and Analysis Infrastructure, incluye algunos parámetros adicionales a los mencionados que no se pueden consultar en el anterior portal, tiene interacción visual para análisis y visualización, mejor disponibilidad de información y su principal función es su implementación en estudios de investigación ambientales y climáticos.

Para esta investigación se usó la variable de precipitación de 25 estaciones climatológicas las cuales se analizaron de manera independiente, verificando que tengan un mínimo de 20 años de información y que los años cuenten con la suficiente disponibilidad de datos para poder realizar un análisis estadístico, seguido de esto se seleccionaron 6 estaciones a las cuales se les realizó el análisis de la tendencia de las muestras, definición y llenado de datos faltantes mediante interpolación lineal, comparación de la correlación para los datos de las estaciones del IDEAM y de los sensores de la NASA, variabilidad entre datos atípicos y se analizaron estadísticamente las muestras mediante diferentes distribuciones probabilísticas tales como Pearson, Spearman y Kendall, comparando los resultados de tal manera que se establezca cual es la correlación que se desarrolla entre las muestras y definir cuales datos presentan una mejor calidad para la cuenca hidrográfica del río Lebrija medio.

El principal objetivo fue realizar la validación y comparación de las muestras obtenidas entre los datos de las estaciones operadas por el IDEAM y los sensores de la NASA, una vez realizado dichas acciones se pretendió obtener los resultados y determinar si es recomendado el uso de estas fuentes alternativas, en la identificación de bases de datos de condiciones climatológicas para cuencas con características similares.

Algunos autores han realizado el análisis estadístico de variables climatológicas tomando como insumo los datos obtenidos de sensores satelitales, tal es el caso de Conde (2018) el cual dentro de su investigación para la cuenca del río Júcar en la provincia del Huamanga en Perú, utilizó 5 estaciones climatológicas y datos de los sensores de Tropical Rainfall Measuring Mission (TRMM) y Global Precipitation Measurement (GPM), evaluando las estimaciones de precipitación cuantitativamente, calificándolas como satisfactorio el uso de las muestras en su estudio, pero concluye que necesitan correcciones entre las mismas.

Este documento está compuesto por 11 capítulos dentro de los cuales se destacan como principales, el Marco Referencial donde se hace la descripción de la información relevante a nivel geográfico y los aspectos más importantes que se encuentran en la cuenca hidrográfica del río Lebrija, el Marco Teórico y Conceptual en el que se hace la descripción de los impactos en la problemática evidenciada, así como la importancia en el desarrollo de análisis de este tipo de investigaciones, basado en estudios de caso de cuencas similares. También cuenta con la metodología donde se hace una descripción detallada de los análisis realizados de cada elemento identificado así como las fórmulas estadísticas usadas y los resultados y discusión en la que se describen los valores finales de los análisis que permiten identificar si se considera viable usar dichas fuentes de información en otros estudios similares.

1 Planteamiento del problema

Colombia es un país localizado en la zona noroccidental de América del Sur, su ubicación es determinante para entender la diversidad de climas con que cuenta y su geografía que resulta ser tan variada, su topografía es predominantemente ondulada que se ve altamente influenciada por sus tres cordilleras oriental, central y occidental, esto le permite tener múltiples condiciones para la vida de diferentes especies de flora y fauna que habitan los diferentes ecosistemas que se encuentran en su territorio.

Tal como lo indica Mejía et al. (1999) en Colombia las condiciones climáticas se ven altamente influenciadas por la dinámica de los vientos alisios y la oscilación meridional de la Zona de Convergencia Intertropical ZCIT, así como la vecindad con los océanos Pacífico y Atlántico, los cuales penetran humedad hacia el interior del país; por otra parte, su conformación fisiográfica es atravesada en parte de su extensión del noreste a sureste por la cordillera de los andes, además gran parte de la extensión del área de la cuenca hidrográfica del río Amazonas se localizan en el territorio colombiano, la cual presenta las mayores precipitaciones y la variabilidad de procesos hidrológicos influenciados por la vegetación y tipo de suelo, generan la dinámica climatológica en el país.

La localización en la ZCIT le brinda la bondad de contar con variabilidad de ecosistemas y climas, la línea ecuatorial lo atraviesa en costado sur, esto permite que durante todo el año se cuente con la misma iluminación solar, por su localización no tiene estaciones, además los días y noches cuentan con la misma duración, se distingue por tener dos periodos climáticos un invierno lluvioso y un verano seco, con tendencia climatológica bimodal el cual se distingue por dos periodos al año cada uno.

Colombia es un país con una gran capacidad de retención hídrica dado por sus precipitaciones, su ubicación en la ZCIT y por la influencia de los vientos alisios quienes a su

vez marcan los periodos de lluvia y de sequía. Según Mejía et al. (1999) otros fenómenos climatológicos que influyen en la variación de las condiciones climatológicas y que condiciona la tendencia de las épocas de invierno y sequia son el fenómeno de El Oscilación Sur-El Niño (ENSO por sus siglas en inglés) asociado a sequias y temperaturas más altas por el calentamiento de las aguas del Pacífico ecuatorial y su fase opuesta denominada La Niña con tendencia hacia fuertes precipitaciones excesivas que generar fuertes inundaciones producto del enfriamiento de las aguas del pacífico ecuatorial. La topografía incide en la variabilidad de diferentes temperaturas que se generan según la elevación, en Colombia gran parte de su extensión es de alta montaña, es atravesada por tres cordilleras, lo que permite encontrarse con ecosistemas en zonas de pie de monte, nevados, páramos, humedales, zona llana, desiertos, ciénagas, sabanas, bosque alto andino y andino.

Según Eslava (1993) el clima es una expresión del comportamiento de la atmósfera, expresado como conjunto de valores para diversos fenómenos meteorológicos en determinado lugar y tiempo, es establecido en términos de frecuencia, además de las características en masa de aire. Entender las condiciones climáticas en el contexto de diferentes estudios, permite establecer soluciones ante diversas problemáticas que se desarrollan en contextos, económicos, sociales y ambientales, por eso es tan importante su análisis en el marco de proyectos a nivel mundial.

Las observaciones meteorológicas según Vega (2022), son mediciones instrumentales y evaluaciones visuales cuantitativas de los procesos y fenómenos atmosféricos, las cuales permiten establecer el clima en una zona específica, en términos de lluvia, vientos y humedad. Para Colombia el IDEAM tiene registros en bases de datos históricos, de mediciones climáticas que realiza de manera diaria, mensual y anual mediante la localización de estaciones meteorológicas y climatológicas, así mismo, existen Corporaciones Autónomas Regionales,

empresas de servicios públicos e instituciones privadas o públicas, que monitorean las condiciones climáticas de manera independiente en el territorio nacional colombiano.

Para Colombia es muy importante poder monitorear sus condiciones climáticas en el desarrollo económico de actividades agrícolas, industriales, mineras, ganaderas, comerciales entre otras y así, establecer las cadenas de consumo de recurso hídrico para la población que habita su territorio, por otra parte le permite mantener identificadas las zonas afectadas por los fenómenos de El Niño y La Niña , así como las posibles zonas amenazadas en el marco de la gestión del riesgo.

Actualmente existe información en bases de datos de variables climáticas establecidas en la localización de estaciones en el territorio colombiano, sin embargo, hay zonas donde no hay cobertura ya sea por la dificultad de acceso o bajo desarrollo económico de la región. Por otra parte, existe una incertidumbre en la veracidad de algunos datos, considerando que el levantamiento de muestras se hace de manera convencional en el caso de las estaciones manuales, tal levantamiento algunas veces se realiza por personal que no está capacitado para esta actividad o se presentan errores por circunstancias ajenas al tomador de la muestra. Lo anterior incide de manera directa en la obtención de datos climatológicos para casos específicos donde la información obtenida puede tener datos erróneos, ocasionando que se generen análisis que no corresponden a las condiciones reales en la cuenca hidrográfica de estudio.

Las entidades que operan y gestionan las estaciones climatológicas a nivel nacional han avanzado en la metodología para la toma de muestras y procesamiento de datos adaptándose a los avances tecnológicos, tanto en la distribución de información como en la alimentación de bases de datos, en algunos casos ya está disponible la información en portales digitales abiertos de consulta gratuita, además, se han implantado estaciones en zonas rurales donde

no existían, sin embargo, la problemática que se ha identificado se concentra en la disponibilidad de información en regiones con bajo desarrollo económico, zonas en las que la construcción de estaciones se dificulta y en estos sectores no se cuenta con información disponible de datos climatológicos.

Sumado a lo anterior, muchas veces no se tiene buena información dentro de las estaciones seleccionadas, ya sea porque que en algunos años no se registraron datos o no se hizo el levantamiento de manera adecuada, esto dificulta el procesamiento de modelos climatológicos inclusive en zonas donde se tienen operadas más de una estación, la problemática se ha tratado utilizando diferentes metodologías de corrección o llenado de datos faltantes, analizadas mediante diferentes distribuciones probabilísticas y análisis estadísticos.

Al identificar la problemática existente en el área de diferentes cuencas hidrográficas en el territorio colombiano, asociada a la falta de información en sectores específicos o baja calidad de las muestras, se busca determinar cuál es la incidencia entre los datos tomados de sensores satelitales de los dos portales de la NASA y compararlos con la información del IDEAM, para lo cual en esta investigación se ha seleccionado (6) estaciones climatológicas en la cuenca medía del río Lebrija ,localizada entre los departamentos de Cesar, Santander y Norte de Santander, verificado que cuentan con información a nivel multianual disponible con datos de al menos 20 años operadas por el IDEAM, además extraer datos climatológicos de The POWER Data Access Viewer y Giovanni de la EARTH DATA de la NASA para verificar la diferencia que puede existir entre los datos obtenidos de la variable de Precipitación dentro de la zona de la cuenca de estudio.

1.1 Pregunta problema de investigación

¿Qué diferencias espaciales y temporales existen entre los datos de precipitación del IDEAM y los obtenidos de sensores satelitales en la cuenca media del río Lebrija, y en qué

medida pueden estos últimos complementar o sustituir la información de estaciones convencionales en zonas con baja densidad de datos?

1.2 Justificación

En la mayor parte del desarrollo de estudios hidrológicos, climatológicos y ambientales en el territorio Colombiano, se han usado datos de estaciones climatológicas operadas por diferentes entidades, al analizar dicha información se ha logrado determinar que existe una incertidumbre en los datos suministrados por diferentes factores, por otra parte la falta de estaciones o su localización tan lejana entre sí, incide en la baja existencia de información para el desarrollo de diferentes estudios.

Con el desarrollo tecnológico a nivel global, han surgido diversas herramientas que permiten consultar datos y generar predicciones a eventos meteorológicos en diferentes zonas geográficas, los cuales se usan en algunos países inclusive reemplazando las estaciones convencionales, para Colombia la implementación de estas nuevas herramientas no se ha implementado con mucha frecuencia y aún se tienen muchas dudas con base a la veracidad o disponibilidad de información disponible.

En el marco de este estudio de investigación, se busca determinar la incidencia de la disponibilidad de datos, utilizando como fuente de información las bases de datos climatológicos provenientes de sensores satelitales. Se comparará la correlación de estos datos con los obtenidos de las estaciones del IDEAM, con el objetivo de evaluar su viabilidad para ser implementados en zonas con baja calidad de datos o donde no existan estaciones instaladas en determinadas cuencas hidrográficas.

Es de vital importancia para los departamentos en el territorio colombiano donde se localizan sus cuencas hidrográficas, contar con información precisa y confiable de las diferentes

variables climatológicas, pues permiten entender el comportamiento del sistema ecológico en determinadas zonas, realizar políticas públicas, analizar los aspectos sociales, definir parámetros o criterios de usos de suelos, actividades económicas, fijar los volúmenes de consumo del recurso hídrico a nivel nacional, así como la definición de sistemas de alerta temprana para la prevención de gestión de riesgo en zonas de amenazas relacionadas con la probabilidad de ocurrencia de eventos climatológicos, a su vez, permite realizar la predicción de los diferentes escenarios que se pueden presentar con la eventualidad de los fenómenos de El Niño y La Niña.

Conforme lo anterior se requiere buscar fuentes de información primaria, oficial y precisa donde, se logre identificar la existencia de parámetros climatológicos, para poderlos implementar en el desarrollo de estudios de caracterización de las variables en diferentes cuencas hidrográficas localizadas en el territorio colombiano, de esta manera determinar si el uso de las fuentes de información local con estaciones presenta mayor confiabilidad o se puede determinar como una buena práctica el uso de datos de portales internacionales que tengan datos de sensores satelitales y compararlos entre sí.

2 Objetivos

2.1 Objetivo general

Comparar la información climatológica de la variable de precipitación, proveniente de estaciones convencionales del IDEAM con datos satelitales obtenidos de los portales POWER Data Access Viewer y Giovanni de la NASA, para la cuenca media del río Lebrija en Colombia, con el fin de identificar diferencias en magnitud, comportamiento temporal y representatividad, y así evaluar su aplicabilidad en estudios hidrológicos de cuencas similares.

2.2 Objetivos específicos

- Seleccionar la información de las estaciones climatológicas que se encuentren localizadas en el área de la cuenca medía del Río Lebrija, basado en las tres fuentes de bases de datos establecidas, los periodos de las series de tiempo a nivel diario, multianual y la disponibilidad de muestras obtenidas.
- Contrastar la información disponible de las tres fuentes seleccionadas de las que se obtuvieron los datos de precipitación, mediante procesos estadísticos de validación de las muestras, estableciendo correlación y tendencia de los datos.
- Establecer recomendaciones según los resultados obtenidos, definiendo si es viable usar datos tomados de sensores satelitales diferentes al IDEAM en estudios climatológicos en cuencas hidrográficas similares.

3 Marco referencial.

A nivel mundial para la elaboración de proyectos agrícolas, ambientales, biológicos, climatológicos e hidrológicos relacionados con el manejo, disposición y aprovechamiento de los recursos hídricos, se debe realizar la evaluación de diferentes factores que determinan las condiciones climáticas de variables de medición en las zonas con diversidad de ecosistemas. Diferentes autores han desarrollado investigaciones en cuencas hidrográficas a las cuales se han realizado análisis climatológicos de la variable de precipitación para determinar el comportamiento de los caudales, producto de las precipitaciones en el área de la mismas, tal es el caso de Pineda y De Piñerez (2020), quien desarrolló una investigación para comparar el balance hídrico de la cuenca del río Checua en el departamento de Cundinamarca Colombia usando datos de 7 estaciones climatológicas de la Corporación Autónoma de Cundinamarca CAR y sensores de teledetección remotos tomados del portal de Giovanni de la NASA para de la misión Tropical Rainfall Measuring Mission (TRMM) a partir de análisis de sensibilidad.

Por otra parte, Ortega (2018), desarrolló una investigación en la cual evaluó la influencia de la precipitación en el área de la cuenca hidrográfica del río Guayllabamba en Ecuador (CARG), tomando como fuente de información climatológica, 27 estaciones del Instituto Nacional de Meteorología e Hidrología (INAMHI) y datos de sensores remotos, de la Misión “Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis (TMPA)” y la misión de Global Precipitation Measurement” (GPM) ambos de la NASA, para la misma localización geográfica de las tres fuentes de información.

Con dicha investigación el Ortega (2018) realizó la toma de muestras a nivel diario multianual de variables de precipitación, evapotranspiración, temperatura, heliofanía humedad relativa. en los periodos de 1998 a 2015, completó los datos faltantes de las estaciones climatológicas cuyos días faltantes eran superiores a 5 días consecutivos, mediante la

metodología de llenado denominada “*media de las muestras entre los días y meses faltantes*” y validó la correlación existente entre las muestras obtenidas de las tres fuentes de información satelital y ordinaria.

Con los datos obtenidos de las tres fuentes de información del (INAMHI), y (TMPA) se realizaron procesos estadísticos dado que los resultados de los datos crudos presentaron índices de baja calidad, con lo cual se hizo un análisis para medir el desempeño de forma cuantitativa y mejorar las muestras; este se basó en la metodología de coeficiente de correlación (Pearson), error medio cuadrático (RMSE), error absoluto medio (MAE), error estándar (ME), sesgo porcentual (BIAS) y coeficiente de Eficiencia de Nash-Sutcliffe (NSE).

Como resultado del proceso estadístico, se logró mejorar la correlación entre las muestras presentando un desempeño superior y ajuste de los datos atípicos, dichos datos corregidos se usaron para un modelo hidrológico específico que se realizó a la cuenca de estudio (CARG).

Seguido de este análisis, los datos corregidos se tomaron para la elaboración de un modelo de lluvia escurrentía, desarrollado mediante el software HEC-HMS con el cual logró determinar la dinámica de las microcuencas y cuál es su comportamiento de las lluvias según las características geomorfológicas de la misma, para determinados periodos de lluvia y diferentes periodos de retorno, así mismo se convirtió en un estudio que permitió valorar la continuidad de la respuesta hidrológica de la cuenca en los años donde se incorporaron los datos de los productos satelitales (TMPA).

Ortega (2018) concluye que la información de las muestras obtenidas de los sensores satelitales (TMPA) comparadas con la información de las estaciones climatológicas convencionales (INAMHI), presentan una variación entre la correlación de las muestras, con lo

cual debieron ser sometidas a diferentes procesos estadísticos de identificación de valores atípicos y relleno de datos faltantes mediante los valores más probables a nivel diario.

Los datos satelitales fueron tratados de acuerdo a las metodologías de corrección de sesgo, todo lo anterior para mejorar la calidad de los datos, finalmente el autor determinó que para la cuenca del río Guayllabamba, los datos de información satelital (TMPA) mediante las correcciones evaluadas a las series históricas, produjo una respuesta hidrológica, en todos sus escenarios, que presenta un mejor rendimiento comparado con la red base de estaciones convencionales (INAMHI.), con lo cual recomienda implementar dicha bases de datos satelitales debido a su retroalimentación y calibración de los sensores en comparación con las estaciones ordinarias.

Finalmente Ortega (2018) concluye que,

“el desempeño de los escenarios de los productos satelitales, en particular el producto TMPA, es efectivo, coherente, y pertinente, para densificar la red de estaciones terrestres, como un aporte adicional en las zonas de baja cobertura o heterogeneidad. Lo que demuestra, a criterio del autor de este estudio, que la posibilidad de incorporar esta información, es apropiada, debido a que en ningún caso la densificación de la red de la CARG, ha dejado de ser representativa de la cuenca, estadísticamente. **(pp. 231 – 232)**

Por otra parte Conde (2018) abordó una investigación para la cuenca de la provincia de Huamanga en Perú, la cual hizo el análisis de datos meteorológicos de la variable de precipitación para 5 estaciones climatológicas del Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI) y los sensores de TRMM y GPM de la (NASA) para los periodos de 2014 a 2016 a nivel diario multianual, el objetivo de la investigación fue obtener y analizar la calidad de los datos, procesar las estimaciones de precipitación tomadas de

información satelital y realizar la comparación de los datos climatológicos tomados de las tres fuentes de información.

El autor estableció el número de días faltantes y verifico qué su porcentaje para llenado, no fuera superior al 20%; en caso tal sugiere según bibliografía utilizada, se debe descartar la estación, además, realizó histogramas determinando la variación de la precipitación mensual para las 5 estaciones, observando que tienen un comportamiento similar entre ellas.

Por otra parte Conde (2018), realizó diferentes análisis estadísticos de homogeneidad para determinar la correlación de las muestras, tales como análisis de doble masa, el cual es usado para analizar la consistencia entre los datos pluviométricos y llenado de los datos faltantes, en el caso de la investigación lo realizó mediante el software HEC-4, por otra parte analizó los resultados obtenidos de los errores estadísticos mediante la metodología coeficiente de Nash – Sutcliffe (NSE), sesgo Porcentual (PBIAS) y ratio de la relación entre RMSE y desviación estándar de las observaciones (RSR).

Para el coeficiente Nash – Sutcliffe (NSE), los datos de los sensores (TRMM) y (GPM) presentan un ajuste aceptable comparado con los registros de las estaciones convencionales, por otra parte, para la relación entre RMSE y Desviación Estándar de las muestras (RSR) tomadas para los sensores satelitales (TRMM) y (GPM) presentaron un error de estimación aceptable a diferencia de los datos de estaciones, con lo anterior se hicieron correcciones a las muestras mediante regresiones simples.

Como resultado de su investigación Conde (2018) concluye que **“Se recomienda continuar con el análisis de descarga, procesamiento y comparación de las imágenes satelitales de estimación de precipitación. Asimismo, sugiere aplicar estos datos en estudios hidrológicos de oferta hídrica para la provincia de Huamanga.” (p. 112)**, según lo anterior el autor no presenta una descripción detallada en sus resultados o conclusiones de si

considera oportuno implementar datos satelitales para cuencas similares en el territorio peruano, sin embargo en dicha descripción deja entender que recomienda usar datos pero no para que tipo de escenarios y de que fuente de información.

En ese contexto, las investigaciones realizadas a las cuencas del río Guayllabamba en Ecuador y la provincia de Huamanga en Perú, resaltan la necesidad de encontrar fuentes de información climatológica diferente a la de estaciones convencionales en algunas regiones de América latina, ambos autores plantean el problema que corresponde a la calidad y disponibilidad de datos tomados de estaciones climatológicas de entidades oficiales.

Para países como Colombia, Ecuador y Perú, al igual que otros de la región latinoamericana, existe una incertidumbre en las muestras que se encuentran para las diferentes variables climatológicas, esto ha llevado a que se busquen otros repositorios de información oficial de entidades que monitorean variables y condiciones climáticas de la región de Sur América, en ese contexto se han desarrollado diferentes investigaciones como las descritas anteriormente, las cuales encuentran en los portales web de sensores satelitales un lugar ideal la elaboración de estudios hidrológicos en cuencas localizadas en los países de la región.

En el contexto Colombiano los datos climatológicos son tomados de información primaria principalmente de la información del (DHIME) del IDEAM, en el cual se pueden descargar muestras de diferentes variables, tal es el caso de Forero y Rodríguez (2016) quien desarrolló una investigación para evaluar la influencia de los métodos convencionales para la estimación de datos faltantes de precipitación para la cuenca El Cune en Villeta, Cundinamarca, el cual uso información de 5 estaciones convencionales.

Dicho análisis contempló la consistencia de los datos por el método de la curva de doble masa, el cual definió una correlación aceptable de las muestras, estos datos pasaron por un

proceso estadístico de reconstrucción de series de precipitación, el cual utilizó el método de ponderación del inverso de la distancia (IDW), también el método de regresión lineal y la combinación lineal ponderada (CLP), con lo cual se pudo determinar que los datos en su condición inicial debían ser sometidos a un proceso de complementación de las muestras y ajuste para garantizar una adecuado estudio.

Para realizar investigaciones a diferentes cuencas hidrográficas en el territorio colombiano, como la Cuenca Media del Río Lebrija, es importante entender la variabilidad de climas que se producen en Colombia, debido a su localización geográfica en el continente Sur americano y planeta tierra, además de la importancia que tiene para la estabilidad energética del ecosistema global terrestre.

Según Eslava (1993) Colombia es un país tropical Montañoso que posee una variedad de climas y contraste de valores climatológicos, también denominadas variables, estas permiten definir el comportamiento de factores climatológicos en el tiempo y la consecuencia de eventos que ocurrieron en una época determinada.

Para realizar una proyección de escenarios que sucedieron y que pueden ocurrir de diversas variables se debe consultar la información meteorológica de entidades oficiales que monitorean las condiciones climáticas en temporalidades o estacionalidades establecidas en zonas del planeta, estos análisis se hacen en el desarrollo de estudios hidro-climatológicos.

La hidro-climatología es una ciencia encargada de analizar la dinámica del clima que existe en la atmósfera con relación a cada una de las variables independientes que forman parte del ciclo hidrológico en una escala de tiempo, en ella convergen diferentes aspectos que se relacionan entre sí, generando el balance energético y variabilidad climática que se produce en el planeta tierra, para la diversidad de ambientes que existen en cada uno de los ecosistemas a nivel global. Según Poveda (2004), existen tres fenómenos involucrados en la

variabilidad climática, las variaciones causadas por forzamientos externos periódicos las cuales están asociadas a la rotación de la tierra en su eje, variaciones debidas a la interacción no lineal, las cuales se producen por oscilaciones climatológicas autorreguladas y variaciones asociadas con fluctuaciones aleatorias en los factores físicos o químicos, las cuales pueden ser externas como (volcanes o meteoritos).

Para el caso colombiano los estudios hidro-climatológicos se realizan mediante la obtención de datos primarios tomados por entidades oficiales, tales como el IDEAM, Corporaciones Autónomas Regionales, instituciones educativas, entre otros, con lo cual se logra establecer la variedad de climas que se producen entre los diversos ecosistemas y cuál es la incidencia en sus condiciones climatológicas a lo largo de su territorio, la localización geográfica del país le permite obtener una gran oferta hídrica, con un alto potencial en zonas con lluvias frecuentes y grandes cauces que desembocan en los mares que la rodean.

Colombia está localizado en el hemisferio norte de Sur América en el trópico cerca de la línea ecuatorial, limita al Norte (N) con el mar de las Antillas o mar Caribe, al Este (E) su frontera limita con los países de Venezuela y Brasil, al Oeste (W) con el océano Pacífico, al Noreste con el país de Panamá y al Sur con los países de Perú y Ecuador, su topografía va de zonas onduladas o montañosas atravesada por tres cordilleras (occidental, central y oriental), cuenta con zonas valles o llanuras predominante planas hasta llegar a sus playas en los mares de los océanos pacífico y caribe. Bernal et al. (2016) manifiesta que Colombia es el único país de Suramérica que tiene costas en los océanos Atlántico y Pacífico, la línea ecuatorial atraviesa el sur del país, con lo cual el 87% de su extensión continental forma parte de hemisferio Norte.

Por otra parte, Colombia está localizada en la Zona de Convergencia Inter Tropical (ZCIT) la cual según Ortega (2018), es un cinturón de baja presión que rodea el globo terrestre de la región ecuatorial, a su vez Poveda (2004) sugiere que las condiciones de precipitación

con una dinámica bimodal de dos periodos de lluvia y dos de sequía en la región son efecto de la localización del territorio en la (ZCIT).

Su localización en la zona ecuatorial (ZCIT) permite la existencia de diversidad de climas y ecosistemas, lo cual genera una gran bondad hídrica por las diversas condiciones climáticas que se producen en sus entornos establecidos por sus cuencas hidrográficas.

En Colombia existen cinco macrocuencas hidrográficas principales que alimentan los volúmenes de caudal de los principales afluentes que cruzan el territorio, se trata de la cuenca del Caribe, Magdalena Cauca, Pacífico, Orinoco y Amazonas, Según Bernal et al. (2016), el territorio continental se extiende desde Punta Gallinas en la Guajira hasta la desembocadura de la quebrada San Antonio en el río Amazonas en Leticia, a su vez dicho autor indica que las tierras bajas de Colombia comprenden la Amazonia, los Llanos Orientales, planicie del Pacífico, llanura del Caribe, valles del Magdalena y el Catatumbo, las corrientes de los ríos corren hacia el oriente, la cuenca del río Amazonas es la más extensa de Colombia con un área de alrededor de 483.000 km², por otra parte existen diferentes afluentes que forman parte de las vertientes de las cuencas hidrográficas del río Magdalena – Cauca, río Atrato, cuencas de la Sierra Nevada de Santa Marta y la Guajira, cuencas ríos Sinú, Patria, San Juan, Mira, Baudó, Guaviare, Meta, Vichada, Tomo, Arauca, Orinoco, Guaviare, Inírida, Caquetá, Putumayo, Vaupés, Sardinata y Zulía entre otros.

Entre esas macrocuencas colombianas se encuentra la cuenca del hidrográfica del río Magdalena, la cual según Bernal et al. (2016), nace en el departamento del Huila en el páramo las Papas, atraviesa de Sur a Norte pasando por las cordilleras central y oriental de los andes hasta desembocar en el mar Caribe en la ciudad de Barranquilla en Bocas de Ceniza, tiene una longitud de 1540 km y un área de cuenca de alrededor de 250.00 km², esta cuenca es alimentada por diversas sub cuencas que descargan los volúmenes de caudal producidos por

las vertientes, tal es el caso de la cuenca hidrográfica del río Lebrija localizado entre los departamentos de Cesar, Santander y Norte de Santander, el cual nace en Piedecuesta y desemboca en el municipio de Puerto Wilches, la cuenca hidrográfica del río Lebrija está dividida en tres micro cuencas, Lebrija Medio, Alto y Bajo. La presente investigación se desarrolla en el área establecida en la Cuenca Media del Río Lebrija.

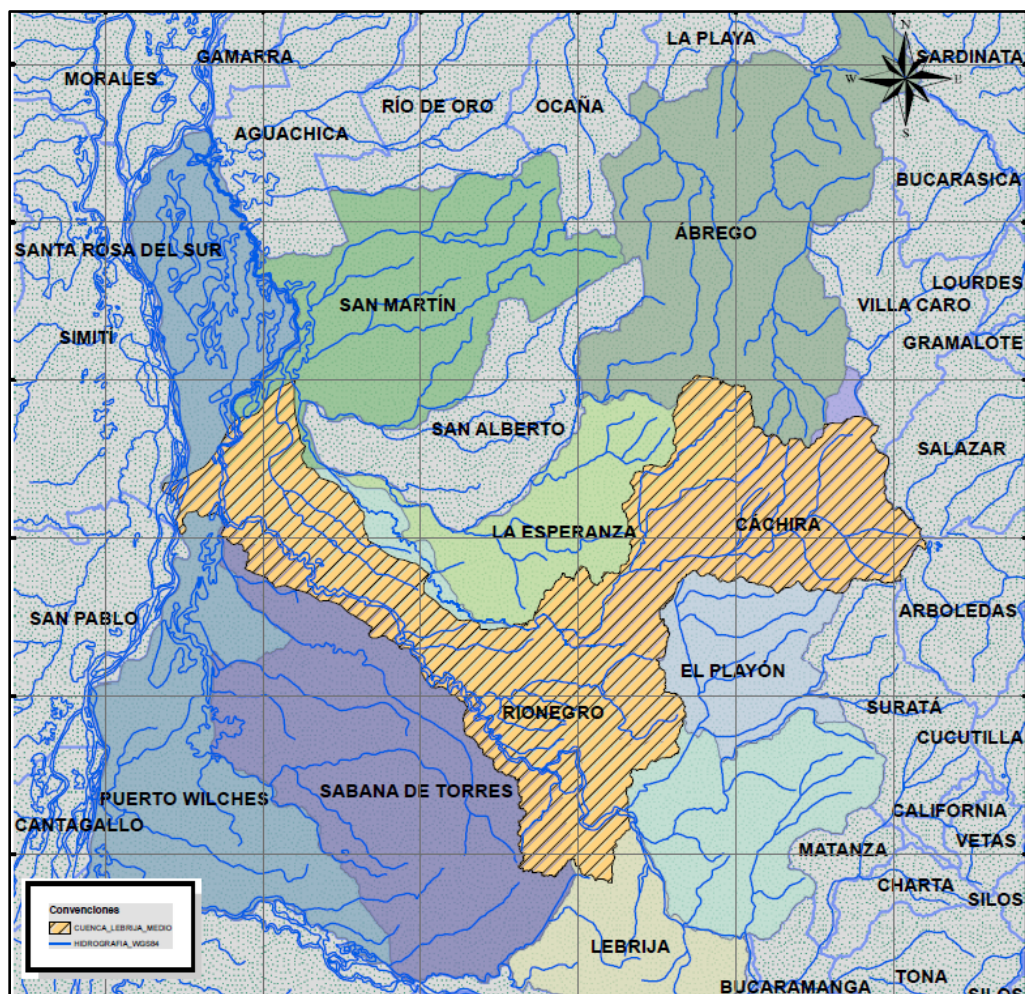
3.1 Localización de la cuenca media hidrográfica del río Lebrija Medio.

Según manifiesta el Plan de Ordenación y Manejo de Cuencas Hidrográficas (POMCA) del río Lebrija Medio de la Corporación Autónoma Regional para la Defensa de la Meseta de Bucaramanga - CDMB (2019) la cuenca se encuentra localizada en Sur América, Colombia entre los departamentos de Cesar, Santander y Norte de Santander, pasa por los municipios de Puerto Wilches, San Martín, Sabana de Torres, Rionegro, La Esperanza, Lebrija, El Playón, Cachirá y Abrego en el territorio colombiano.

Tal como lo indica la Figura 1 la localización de la cuenca media (indicada en amarillo), cuenta con una gran cantidad de afluentes que alimentan los flujos que descargan en el río Lebrija, según Aguirre (2007) la geometría y el destino de escurrimiento la consolidan como una Cuenca de Vertiente, por otra parte la CDMB, (2019) en el (POMCA), sugiere que **“alcanza un área aproximada de 192.707 hectáreas, y una longitud del cauce principal de 168.67 km” (p. 32)**, con lo cual Aguirre (2007) sugiere que corresponde a una cuenca “intermedia – grande” dado que está en un rango de (500 a 2500 km²).

Figura 1

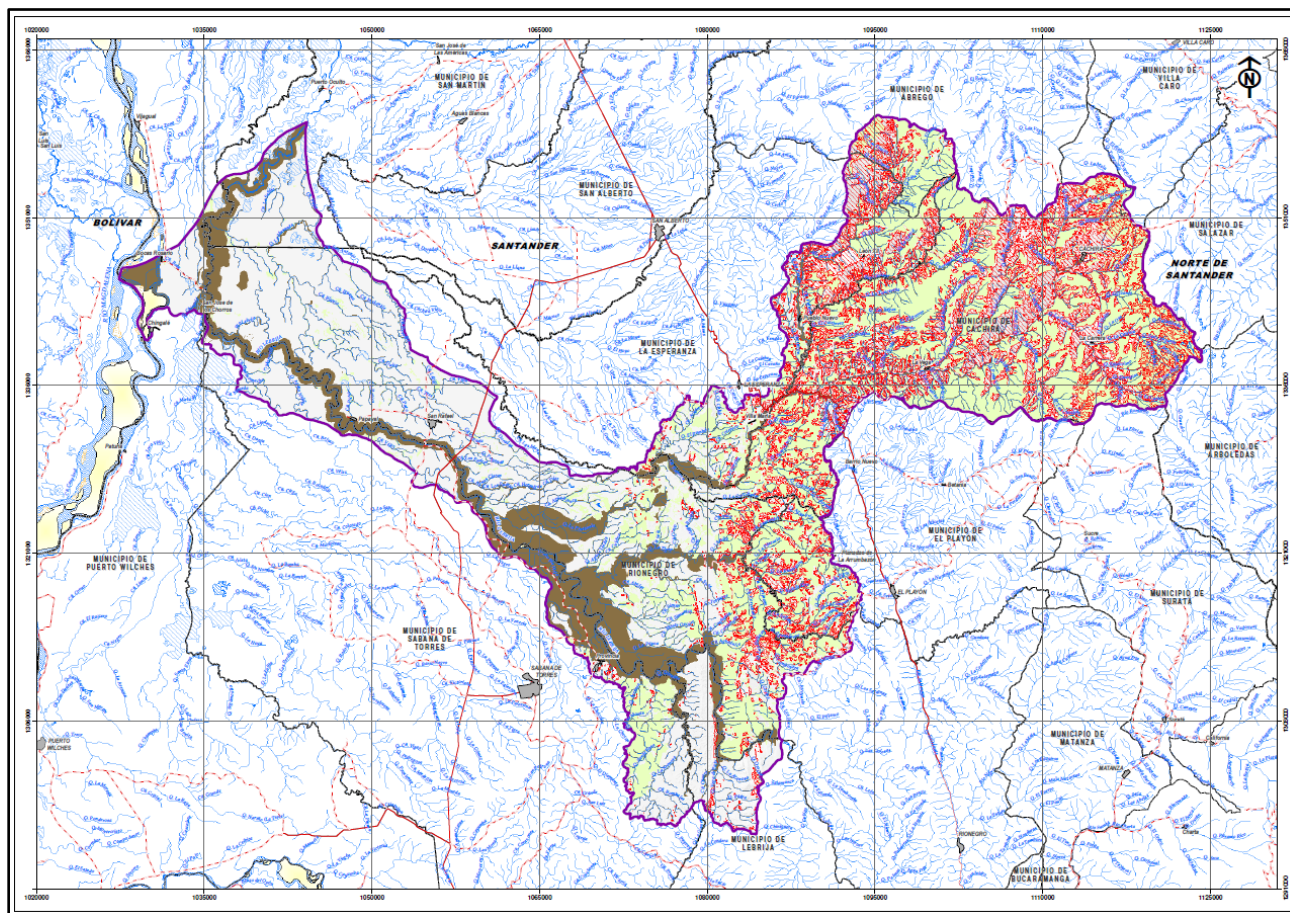
Localización Cuenca medía río Lebrija y sus límites con los municipios que atraviesa.



Nota. Fuente elaboración propia.

Figura 2

Localización Cuenca media río Lebrija con red hidrográfica.



Nota. Tomado de “Consultoría para el ajuste (actualización) del plan de ordenación y manejo de la cuenca hidrográfica del Río Lebrija medio (código 2319-03) de conformidad con los términos establecidos por el Fondo de adaptación dentro del convenio interadministrativo no. 021 de 2014” , Mapa del escenario tendencial cuenca río Lebrija medio , por Corporación Autónoma Regional para la Defensa de la Meseta de Bucaramanga -CDMB, 2019.

3.2 Factores económicos de la cuenca media hidrográfica del río Lebrija.

Tal cómo indica CDMB (2019), la cuenca media del río Lebrija por su localización estratégica entre los departamentos del Cesar, Norte de Santander y Santander, así como sus condiciones climáticas a diferentes altitudes de sus ecosistemas, le permiten un desarrollo el cual establece un uso de suelo para actividades agrícolas, ganaderas, pesca, comerciales, turísticos, mineros, industriales, transporte viales y férreos, por otra parte, el autor también

establece que las coberturas del suelo se caracterizan por tener condiciones de pastos limpios y vegetación secundaria alta.

Por otra parte las modificaciones antrópicas del uso del suelo para diversos usos, principalmente la ganadería excesiva y la construcción, ha modificado las condiciones climáticas del área de la cuenca hidrográfica.

3.3 Usos de suelo de la Cuenca Media hidrográfica del Río Lebrija.

Según establece CDMB (2019) en el (POMCA) del río Lebrija Medio las condiciones geográficas del terreno muestran una variabilidad de ecosistemas, caracterizados por las diferencias en la topografía, predominantemente ondulado o montañoso lo cual establece las diferencias en las condiciones climáticas por las diferentes altitudes y las características del suelo. Las cobertura del suelo o tierra en la Cuenca, se clasificaron en Tejido urbano discontinuo producto del desarrollo urbanístico en la cabecera de municipios que la conforman, la red vial y líneas férreas de conexiones entre los diferentes centros poblados, además se caracteriza por la explotación de hidrocarburos, minería, actividades agrícolas tales como arroz, palma de aceite, pastos limpios, arbolados, enmalezados, mosaico de pastos y cultivos de diferentes insumos para consumo.

Por otra parte la cobertura de la tierra y clasificación del suelo según el (POMCA) del río Lebrija medio CDMB (2019), corresponde a

“Bosque denso alto de tierra firme, Bosque denso bajo de tierra firme, Bosques fragmentados con pastos y cultivos, Bosque de galería y ripario, Plantación forestal de coníferas, Herbazales densos de tierra firme, Arbustal denso, Arbustal abierto, Vegetación secundaria alta, Vegetación secundaria baja, Zonas pantanosas, Ríos, Lagunas, lagos y ciénagas naturales.” (p. 66)

3.4 Factores climatológicos de la cuenca media hidrográfica del río Lebrija.

La cuenca hidrográfica media del río Lebrija se caracteriza por la diversidad de climas tropicales gracias a su localización estratégica en la cordillera oriental en el territorio colombiano, tal como indica CDMB (2019)

Debido a la localización geográfica de la cuenca del río Lebrija Medio, cuyo centroide se encuentra aproximadamente a los 7°36' de latitud norte (N) y 73° 60' de longitud oeste (W), sobre la vertiente occidental de la cordillera Oriental en la zona Andina colombiana, lo cual le confiere una variación de la precipitación debido al efecto el Efecto Foehn originado por la masas de aire húmedo que viene de Venezuela y Brasil. **(p. 58)**

Dicho esto se entiende que sus condiciones climáticas le generan una gran oferta hídrica caracterizada por las precipitaciones que recargan los niveles de agua que se transitan por diferentes cueces o corrientes hasta alimentar el río Lebrija, las temperaturas son medias y varían según la elevación de la topografía del terreno a lo largo de la cuenca.

Las condiciones de las diferentes variables climatológicas en la actualidad son tomadas y monitoreadas por entidades gubernamentales e instituciones educativas, centros de investigaciones, entre otros, los cuales permiten obtener y descargar bases de datos de eventos que ocurrieron con periodicidad o estacionalidad de tiempo a nivel multianual, mensual diario y horario de muestras en zonas específicas del planeta tierra.

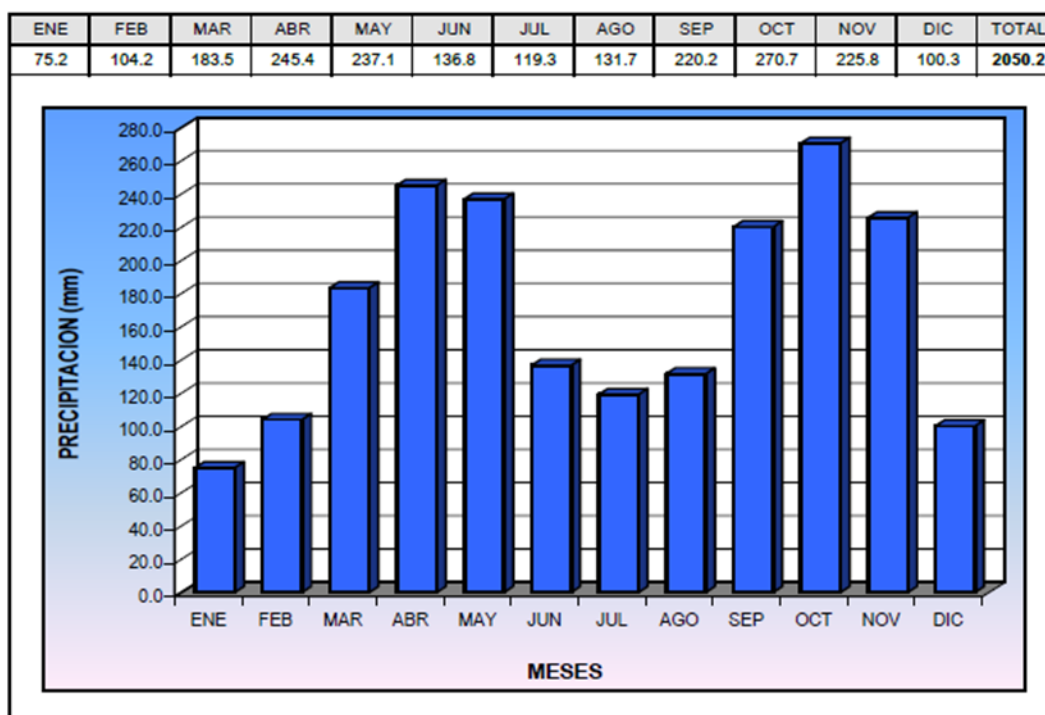
En la cuenca hidrográfica media del río Lebrija, el IDEAM tiene instrumentadas diferentes estaciones de las cuales se puede consultar de manera gratuita a través del portal de datos abiertos DHIME las variables de precipitación, temperatura del aire, humedad relativa, brillo solar, nubosidad, radiación solar, evaporación y velocidad del viento, entre otras y realizar

la descarga para el desarrollo de estudios climatológicos, ambientales, hidrológicos, hidráulicos entre otros.

Según, CDMB (2019), en la cuenca media del río Lebrija existen 101 estaciones climatológicas del IDEAM para las diferentes variables climáticas, a lo largo de los municipios de El Playón, Girón, Lebrija, Puerto Wilches, Rionegro, Sabana Torres, San Martín y Suratá, su tendencia está marcada por la dinámica de la (ZCIT), correspondiente a una franja de bajas presiones donde llegan corrientes de aire cálido y húmedo provenientes de los cinturones de alta presión, ubicados en la zona subtropical de los hemisferios Sur y Norte, lo cual le permite tener dos periodos de estacionalidad de lluvias o tendencia Bimodal a lo largo del año.

Dichos periodos de lluvia, tal como se indica en la Figura 3 inician con la primera temporada de invierno en el mes de marzo alcanzando su pico más alto entre abril y mayo y baja en junio hasta agosto, cuando se generan temperaturas cálidas o de verano. Adicionalmente, un segundo periodo de lluvia el cual genera un incremento en las precipitaciones en septiembre alcanzando su pico más alto en octubre y noviembre, por ende los meses de diciembre a febrero están marcados por una temporada de verano.

Figura 3
Valores totales mensuales de precipitación.



Nota. Tomado de “Consultoría para el ajuste (actualización) del plan de ordenación y manejo de la cuenca hidrográfica del Río Lebrija medio (código 2319-03) de conformidad con los términos establecidos por el Fondo de adaptación dentro del convenio interadministrativo no. 021 de 2014” , (p. 285), por Corporación Autónoma Regional para la Defensa de la Meseta de Bucaramanga -CDMB, 2019.

3.5 Datos climatológicos de precipitación

Históricamente a nivel mundial se ha desarrollado el monitoreo de variables climáticas de diversas maneras, la más común es la toma de datos convencionales que se basa en el muestreo diario, de forma manual o automática, desde estaciones climatológicas que se localizan estratégicamente en los territorios y alimentan bases de datos para diferentes entidades. Además, existen sensores satelitales que monitorean también variables desde el espacio tomando muestras de ondas para recolectar datos de la superficie terrestre, estos sensores miden diversas longitudes de onda (radios, microondas, infrarroja, luz visible etc.) que son reflejados en la tierra en las extensiones de diversas zonas del planeta, dicha información según el país donde se realizan las muestras puede ser confidencial o abierta.

Ambas metodologías son muy útiles para el desarrollo de modelos climatológicos, la forma convencional hasta la fecha es la principal fuente de información que se utiliza para el desarrollo de estudios hidro-climatológicos, pues las muestras de los datos están avaladas por entidades gubernamentales y se monitorean los eventos de manera directa con información que se puede medir de un suceso que existió en determinado tiempo.

Por otra parte los portales web que contienen imágenes y bases de datos sensores satelitales de variables climatológicas, durante los últimos años se han convertido en una fuente de información muy completa, la cual ha generado una metodología muy bien sustentada, basado en la inclusión de herramientas ofimáticas modernas, quien a su vez también presenta muy buenos resultados y se puede obtener muestras de zonas donde no se cuenta con estaciones instrumentadas, además los datos contienen información completos a nivel diario.

En el desarrollo de proyectos hidrológicos, climatológicos y ambientales se busca encontrar fuentes de información que ayuden a reducir la incertidumbre en los datos obtenidos para diferentes variables, tomando como referencia los datos de las estaciones climatológicas que en ocasiones tienen baja precisión en las muestras, esto se ha logrado mediante el desarrollo tecnológico de herramientas ofimáticas, que buscan permitir una mejor disponibilidad de información y correlación estadística para los diferentes análisis.

En Colombia la extracción de bases de datos de sensores satelitales no se ha implementado con mucha frecuencia, la mayor parte de los estudios hidro-climatológicos se desarrollan basados en la información de datos abiertos para estaciones de entidades gubernamentales, especialmente del IDEAM. Gran parte de los análisis de información que se ha desarrollado con datos climatológicos tomados de sensores satelitales se ha implementado en proyectos de investigación, tal es el caso de Sánchez (2024) quien desarrollo una

investigación para la comparación de los datos de los programas satelitales de CHIPS, ERA 5 y el IDEAM a una cuenca localizada entre los departamentos de Boyacá y Casanare, así mismo Pineda y De Piñerez (2020), también en su proyecto de grado generó una comparación del balance hidrológico de la cuenca hidrográfica del río Checua en el departamento de Cundinamarca tomando información de estaciones meteorológicas de la CAR y sensores satelitales remotos de teledetección TRMM.

Por otra parte, en algunos países de la región es más común incluir datos extraídos de sensores satelitales, en estudios hidro-climatológicos como lo son Argentina, Ecuador, Brasil, Perú, Bolivia, Chile y México, además se han desarrollado diversas investigaciones que presentan los análisis tomando como fuente de información datos de sensores satelitales, en ese contexto Vilchis et al.(2015), realizó la evaluación de datos de precipitación a nivel diario de datos de estaciones meteorológicas automáticas (EMA) y los sistemas satelitales de PERSIANN y CMORPH, tomando información de los años de 2003 a 2009 en la cuenca de río Amacuzac en México, así mismo Ortega (2018) realizó la evaluación de la influencia de la incorporación de los datos de precipitación satelitales (TMPA-3B42R) para la simulación hidrológica a nivel diaria en la cuenca alta del río Guayllabamba en Ecuador.

3.6 Fuentes de información de datos variable de precipitación para la cuenca del río Lebrija medio

En el desarrollo de esta investigación se han seleccionado tres fuentes de información que tienen disponibilidad suficiente de datos, son de acceso libre, además manejan registros diarios de eventos que ocurrieron en épocas anteriores y se pueden usar para el análisis en el estudio de la cuenca media del río Lebrija, dichas fuentes se presentan a continuación:

- **Datos hidrometereológicos del IDEAM.**

El Instituto de Hidrología, Meteorología y Estudios Ambientales IDEAM tiene implantadas estaciones climatológicas en diferentes regiones del territorio nacional colombiano, cuenta con tres clases; meteorológica, hidrológica e hidrometereológica las cuales se caracterizan según el tipo de variables que mide cada una, se pueden definir a nivel de Agrometeorológica, Climatológica Ordinaria, Climatológica Principal, Limnigráfica, Limnimétrica, Mareográfica, Meteorológica especial, Pluviográfica, Pluviométrica, Radio Sonda, Sinóptica Principal y Sinóptica Secundaria, además las estaciones del IDEAM se pueden identificar según su tecnología, las cuales pueden ser de tipo Convencional o manual, Automática con telemetría y Automática sin telemetría.

La información de las estaciones requerida se puede examinar en el portal de datos abiertos DHIME (Datos Hidrológicos y Meteorológicos) que es una página web del IDEAM, la cual funciona como una plataforma de consulta y descarga de datos hidrometereológicos, además según el tipo de información de la estación a consultar se puede solicitar de manera directa, en la periodicidad del tiempo que se requiera, ya sea diaria, mensual o anual, en cualquier caso mediante análisis estadísticos se puede realizar la correlación entre los datos, complemento de las series faltantes y análisis que climático se desee.

- **Datos hidrometereológicos desde interface web Power Data Access Viewer de la (NASA Prediction of Worldwide Energy Resources).**

El POWER (Prediction Of Worldwide Energy Resources) Data Access Viewer es una herramienta proporcionada por La National Aeronautics and Space Administration NASA, para la visualización de datos climáticos y meteorológicos.

La NASA es una agencia del gobierno de los Estados Unidos de Norte América, que estudia diversos aspectos espaciales, una de las líneas de profundización a partir de la cual se

producen grandes investigaciones, es el departamento de meteorología que analiza las diferentes variables climatológicas en diferentes regiones del planeta tierra, cuenta con diversos sensores satélites que monitorean de manera diaria diferentes variables climatológicas. The Power Data Access Viewer de la NASA Prediction of Worldwide Energy Resources es un portal web donde se puede visualizar y descargar información de manera libre y gratuita, contiene información muy actualizada para América del sur. Según lo mencionado por Villena (2021), el sitio web Power de la NASA contiene información de parámetros relacionados con la meteorología y energía solar formulados para diseñar sistemas de energías renovables, basado en esto fue implementado en el análisis temporal de años hidrológicos en la cuenca del río Santa en Perú.

Según manifestó, National Aeronautics and Space Administration - NASA (2021)

The POWER Data Services consists of a series of restful Application Programming Interfaces (API), geospatial enabled image services, and Open-source Project for a Network Data Access Protocol (OPeNDAP) services. These three different service offerings support data discovery, access, and distribution to our user base as Analysis Ready Data (ARD) and as direct application inputs to decision to support tools.

(p. 1)

El portal web The Power Data Acces Viewer de la NASA se convierte en una buena fuente de información para la extracción de datos meteorológicos de diversas variables en las diferentes regiones del territorio colombiano, cuenta con datos que se pueden extraer definiendo las condiciones de entrada, tales como el tipo de variable y preciosidad de tiempo, localización geográfica determinando las coordenadas del punto de interés, esta fuente de información se considera muy útil para el desarrollo de la presente investigación por la interfaz que maneja en la extracción de datos.

Los datos de precipitación en el portal web The Power Data Acces Viewer de la NASA utiliza los sensores satelitales TRMM (Tropical Rainfall Measuring Mission), GPM (Global Precipitation Measurement), MERRA-2 (Modern-Era Retrospective analysis for Research and Applications), MODIS (Moderate Resolution Imaging Spectroradiometer) y VIIRS (Visible Infrared Imaging Radiometer Suite).

- **Datos hidrometereológicos desde la interface web GIOVANNI of The NASA. (The Earth Science Data Systems (ESDS))**

Giovanni es un sistema desarrollado por la National Aeronautics and Space Administration NASA para la visualización y análisis de datos de varias misiones satelitales y terrestres relacionadas con el medio ambiente y el clima. El programa Earth Science Data Systems (ESDS) de la NASA, cuenta con información de bases de datos completos y abiertos que se implementan a nivel mundial para estudios hidro-climáticos que permiten entender la dinámica del planeta tierra y la protección del mismo. Este portal web cuenta con información de bases de datos en toda la región del planeta tierra, es ideal para la predicción de escenarios climatológicos y meteorológicos en sectores de investigación como agricultura, agua, calidad del aire, diversidad biológica, previsión ecológica, desastres naturales, enfermedades, sistemas de información geográfica, gases de efecto invernadero, cambio climático, objetivos de desarrollo sostenible e incendios forestales, entre otros.

Los datos que presenta dicho portal son de carácter oficial y abiertos, su acceso es de forma libre y se pueden extraer bases de datos climatológicas de manera gratuita, tiene la ventaja que implementa varios sensores de la NASA que monitorean desde hace años diversas variables, en las ramas de estudio que se requiera y los datos se entregan de forma diaria, mensual o anual según se requiera.

Sagun, El portal web National Aeronautics and Space Administration - NASA (n.d.)“

Giovanni is a Web-based application developed by the Goddard Earth Sciences Data and Information Services Center (GES DISC) that provides a simple and intuitive way to visualize, analyze, and access vast amounts of Earth science remote sensing data without having to download the data (although data downloads are also supported). Giovanni is an acronym for the Geospatial Interactive Online Visualization and Analysis Infrastructure. **(p. 2)**

Al igual que el portal web The Power Data Acces Viewer, **GIOVANNI** utiliza los sensores satelitales TRMM , GPM, MERRA-2, MODIS y VIIRS , se concentra en datos atmosféricos y climáticos, incluyendo información sobre precipitación, nubes, humedad, viento y temperatura.

En particular ambos portales web, comparten algunas fuentes de datos comunes de sensores satelitales, pero los enfoques y las variables que se analizan son diferentes. Por una parte, GIOVANNI suele tener un enfoque más amplio en el análisis climático a diferencia de POWER Data Access Viewer el cual se enfoca en la energía renovable y las condiciones climáticas que afectan la producción de energía. Sin embargo, para la variable de precipitación ambas tienen información ideal para el análisis de estudios hidrológicos.

La presente investigación se realizó para las estaciones climatológicas que se encuentren dentro del área aferente de la cuenca hidrográfica del río Lebrija medio, únicamente para la variable de precipitación, cabe mencionar que en la zona se monitorean también diferentes variables, sin embargo, estas no forman parte del análisis de este estudio, adicionalmente se seleccionaran las estaciones que tengan la mejor calidad de las muestras, así como la disponibilidad de datos, cuyos periodos deben ser superiores a 20 años consecutivos.

Por otra parte, se recomienda que el nivel de aprobación que es la *(calificación de calidad de los datos obtenidos de las estaciones meteorológicas)* por parte del IDEAM, sea de

1200 lo que corresponde a una estación con datos de buena calidad, para aquellas muestras cuyos datos consecutivos de las estaciones son faltantes en un periodo de más de 5 días seguidos, se llenarán los datos faltantes y se realizaran los respectivos análisis, para los datos tomados de sensores satelitales no es necesario realizar relleno de datos dado que en su información se entregan valores diarios completos.

4 Marco teórico y conceptual.

El siguiente marco teórico y conceptual, muestra las teorías y las definiciones de los conceptos técnicos relacionados con la investigación de estudios hidro-climáticos en cuencas hidrográficas, los cuales se realizan a través de análisis estadísticos de predicción de eventos de lluvia que se producen en una temporalidad determinada, para establecer unos caudales que se producen en un cauce específico y cuál es la importancia de los datos o muestras de la variable de precipitación en dichos estudios.

En el desarrollo de estudios climatológicos, hidrológicos, meteorológicos y ambientales a nivel mundial se hace necesario contar con registros de bases de datos de acontecimientos que ocurrieron en un tiempo determinado en diferentes zonas del planeta, los cuales contienen información de variables climáticas, son monitoreadas por diferentes entidades en estaciones climatológicas convencionales o mediante sensores satelitales. Para entender la dinámica del ciclo hidrológico del agua se usan las variables de precipitación y evaporación.

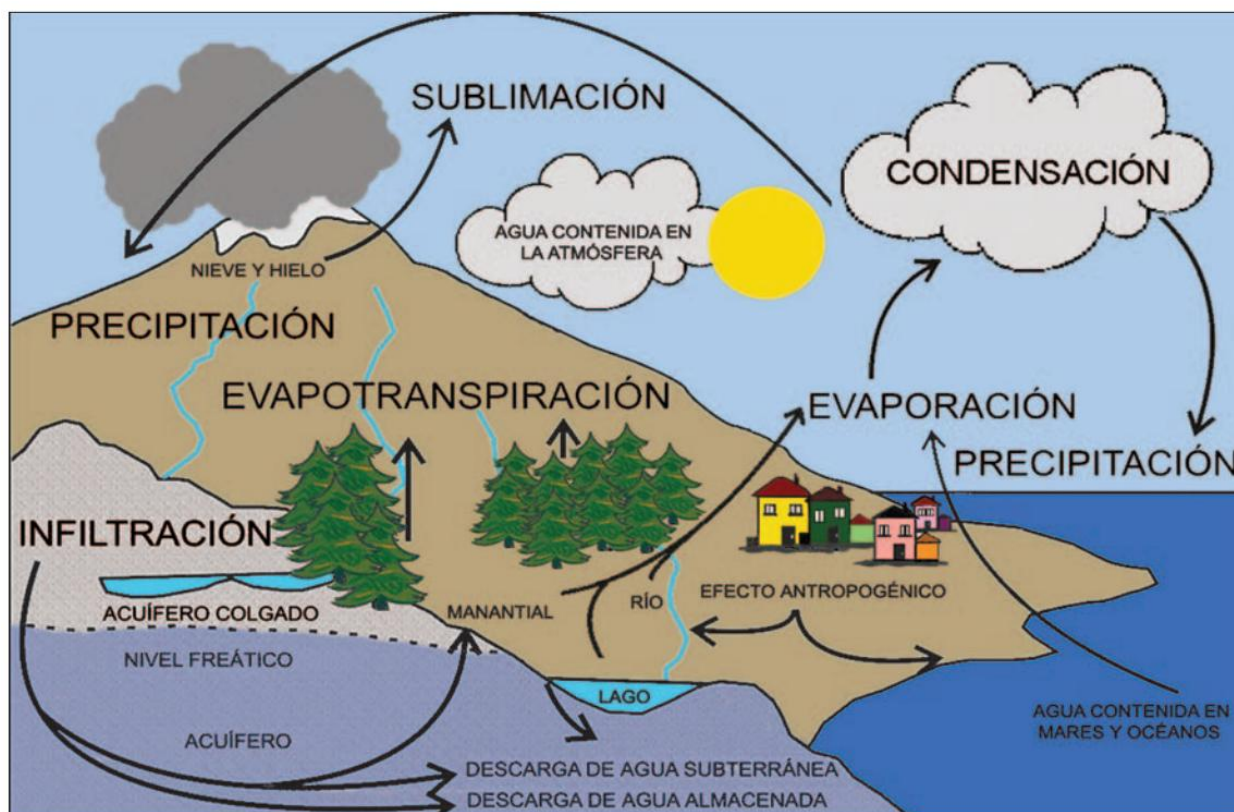
4.1 ¿Qué es el ciclo hidrológico?

El ciclo hidrológico describe la circulación del agua en sus tres estados (sólido, líquido y gaseoso) en la superficie de la Tierra. Se compone de seis etapas: evaporación, transpiración, condensación, precipitación, infiltración - escorrentía y almacenamiento, tal como lo indica Ángel et al. (2009) este ciclo supone un movimiento o transferencia de masas de agua como consecuencia de un flujo energético. Este proceso ocurre de manera cíclica, con variaciones según la época del año y la ubicación geográfica. El agua en estado líquido fluye sobre la Tierra, el hielo o nieve representan el estado sólido, y las nubes corresponden al estado gaseoso,

Las etapas del ciclo hidrológico se reconocen por ser la evaporación, la cual es la condición donde el agua de las fuentes (lagos, ríos, cauces y océanos, etc.), se evapora por el calor del sol, convirtiéndose en vapor de agua que asciende hacia la superficie de la atmosfera, por otra parte, la transpiración, es aquella condición donde las plantas liberan vapor de agua y se produce la evotranspiración, la condensación es donde el vapor de agua que está en la atmosfera, se enfría convirtiéndose en las gotas de agua que alimentan las nubes, cuando las gotas de agua se agrupan y descienden a la tierra se conoce como precipitación o lluvia, esas gotas de agua cuando tocan la superficie de la tierra transitan por la capa superficial del suelo se conoce como escorrentía, la cual alimenta los cauces o ríos también conocida como la fase de almacenamiento, mientras que esas gotas que ingresan y traspasan las capas del suelo se conoce como infiltración, la cual se caracteriza por recargar los acuíferos o flujos subsuperficiales.

Según Ángel et al. (2009), el ciclo hidrológico es un conjunto de procesos físico-químicos simultáneos interrelacionados entre sí, en el cual supone un movimiento de masas de agua en sus diferentes estados de forma continua como consecuencia de un flujo energético que se produce en el planeta tierra, dicho autor nos presenta una imagen gráfica en la cual se puede entender cada una de las fases que lo conforman (ver la Figura 4 y la importancia en el balance ecosistémico a nivel global, una de los temas que causa mayor relevancia por dicho autor se refiere a las condiciones del suelo que encuentra en la superficie de la tierra, ya que este puede permitir unas condiciones de vegetación e infiltración que de manera directa condicionan todas las fases del ciclo hidrológico.

Figura 4
Ilustración del ciclo Hidrológico



Nota. Tomado de “El ciclo hidrológico: experiencias prácticas para su comprensión 2009” , (p. 79), por M. Ángel et al., 2009, Revista Enseñanza de las Ciencias de la Tierra.

Conociendo el ciclo hidrológico y la importancia que tiene para la estabilidad ecosistémica del planeta tierra, los estudios que se desarrollan en esta investigación se caracterizan por realizar un énfasis especial en la variable de precipitación, la cual como se pudo definir en dicho ciclo, es aquella que permite que toda el agua que se almacena para los diferentes usos en la superficie terrestre viene directamente de la condensación en su estado gaseoso y descargada hacia la superficie en su condición de lluvia. Por otra parte dichas fases que conforman el ciclo hidrológico son conocidas como variables climáticas y se desarrollan en el área de una cuenca hidrográfica.

4.2 Cuenca Hidrográfica

Una Cuenca hidrográfica es un área definida por la parte más alta de las montañas también conocida como línea divisoria de agua, se caracteriza por captar todos los niveles de escurrimiento de la lluvia que alimenta los cauces principales y flujos de escorrentía superficial, según Aguirre (2007), son consideradas parte integrante de la región hídrica, la define **“como una área natural donde el agua de lluvia es captada y desalojada o depositada en un almacenamiento natural; por un sistema de drenaje, definido por medio de líneas divisorias, topográficamente llamadas parte-aguas (p. 9)”**.

4.3 Variables Climatológicas

Las variables climáticas son elementos que se encuentran entre la atmósfera y la superficie terrestre y se miden para comprender el ciclo hidrológico. Estas variables son fundamentales para el balance energético de la Tierra, afectan diversos fenómenos climáticos. Estas variables permiten estudiar los patrones atmosféricos y meteorológicos en distintas regiones y ambientes ecológicos, como forestales, desérticos, tropicales, entre otros. En Colombia, las variables que se miden incluyen precipitación, temperatura del aire, presión atmosférica, humedad relativa, brillo solar, nubosidad, radiación solar, evaporación y velocidad del viento, las cuales son cruciales para entender cómo las condiciones climáticas influyen en los ecosistemas de las cuencas hidrográficas

Según Poveda (2004) la variabilidad climática de Colombia en sus distintas escalas temporales, muestra una síntesis de diagnóstico de elementos físicos que se producen en las zonas del territorio y las evalúa en las escalas Inter-Decadal, escala inter Anual (ENSO) y Escala Anual, el cual concluye que la variación se produce de forma natural para las diferentes escalas temporales, asumiendo que el cambio ambiental por las acciones antropogénicas está aumentando la complejidad del sistema climático.

4.3.1 Precipitación

La variable de precipitación es aquella que muestra los datos de lluvia, los cuales cómo se ha manifestado anteriormente, dentro del ciclo hidrológico es cuando se produce la condición de transformación desde el estado gaseoso a convertirse en gotas de lluvia o granos de hielo. Estos niveles alimentan los drenajes superficiales que transitan por las diferentes superficies y capas del suelo, su medición se hace generalmente en unidades de medida (mm/h), los cuales permiten establecer cuantos milímetros de agua cayeron en un tiempo determinado para una área determinada.

En el desarrollo de esta investigación se usará únicamente la variable de precipitación que mide los niveles de lluvia que se han producido en diferentes épocas y en zonas establecidas en el territorio colombiano, estos datos se analizarán de manera diaria en un rango de periodo desde el 1 de enero de 2003 hasta junio del 2023, los datos de precipitación se entregan en unidad de medida milímetros por hora (mm/h), y los datos fueron tomados de estaciones convencionales del IDEAM y de los sensores satelitales Tropical Rainfall Measuring Mission (TRMM) y Global Precipitation Measurement (GPM) ambos de la NASA.

Con los datos de la variable de precipitación se pueden realizar predicciones mediante modelos estadísticos de las condiciones pluviográficas de una cuenca hidrográfica, los cuales son usados en la evaluación de estudios de eventos de lluvia que pueden ocurrir en determinadas zonas y en una periodicidad determinada; estos estudios son conocidos como Hidrológicos.

4.3.2 Instrumentos de medición

Para la variable de precipitación existen diferentes mecanismos que permiten medir la cantidad de agua en forma de lluvia que se produjo en un evento determinado, cabe resaltar que dichas mediciones las realizan diferentes entidades especializadas, las cuales alimentan la

información en bases de datos climatológicos de eventos que ocurrieron en una época específica, a nivel colombiano el IDEAM, es la principal entidad que monitorea todas las variables climatológicas.

Para las estaciones climatológicas de precipitación existen dos principales instrumentos de medición, los pluviómetros y pluviógrafos, los cuales miden la cantidad de agua que se produjo en determinado tiempo y localización. A continuación se hace una breve descripción de cada uno de manera específica.

- **Pluviómetros**

Es un instrumento de medición de volúmenes de agua proveniente de la lluvia que se recolecta en un recipiente de manera manual verificando cuantos fueron los milímetros de agua que se recolectaron en unas probetas que se llenan según la intensidad y duración del aguacero, para el IDEAM según lo manifiesta Vega. (2022), la toma de muestras se debe realizar de manera diaria recolectando las muestras en un periodo 10 minutos antes de las 07:00 horas colombianas y la muestra tiene como resultado el volumen acumulado en un horario de 24 horas contando desde las 07:00 horas del día anterior.

El procedimiento se realiza contabilizando el agua recolectada, si el aguacero es fuerte la probeta se rebaza y es almacenada por un cilindro el cual llena otras probetas, al final el resultado se produce por la contabilización de la suma de probetas que tienen un máximo de 10 mm de agua, esto produce un dato que es denominado “Lluvia PM mm” también conocida como Precipitación máxima en milímetros. En la Figura 5 se muestran los instrumentos utilizados.

Figura 5

Ilustración del pluviómetro y registro de agua en estaciones del IDEAM.



Nota. Tomado de “Protocolo de observaciones en las estaciones meteorológicas 2022” , (p. 24), por Vega. R. E. E., 2022, INSTITUTO DE HIDROLOGÍA, METEOROLOGÍA Y ESTUDIOS AMBIENTALES - IDEAM.

- **Pluviógrafos**

Es un instrumento de medición más avanzado que verifica los volúmenes de agua proveniente de la lluvia de manera automática y continua, verificando cuantos fueron los milímetros de agua que se recolectaron en un una cantidad de tiempo y en un aguacero específico, permite medir intensidades de lluvia en tiempos cortos., para el IDEAM según lo manifiesta Vega. (2022), la toma de muestras se realiza de manera continua mediante la toma en unas bandas o fajas registradoras, las cuales se deben cambiar según el mantenimiento que se programe a la estación.

Este instrumento permite obtener datos más precisos ya que levanta muestras de manera continua en el horario de las 24 horas y permite observar el comportamiento de los aguaceros que se produjo en un tiempo y evento establecido, los resultados también indican valores de “Lluvia PM mm”. En la se muestran los instrumentos utilizados.

Figura 6

Ilustración del pluviógrafo y extracción del tambor del pluviógrafo



Nota. Tomado de “Protocolo de observaciones en las estaciones meteorológicas 2022” , (p. 26), por Vega. R. E. E., 2022, INSTITUTO DE HIDROLOGÍA, METEOROLOGÍA Y ESTUDIOS AMBIENTALES - IDEAM.

Adicional a los mecanismos convencionales establecidos para estaciones climatológicas, se han desarrollado diferentes instrumentos de medición satelital los cuales toman muestras de valores de precipitación de manera continua mediante la emisión de ondas transmitidas hacia la tierra, dichos instrumentos son conocidos como sensores satelitales. A continuación se hace una breve descripción de estos de manera específica.

- **Sensores satelitales.**

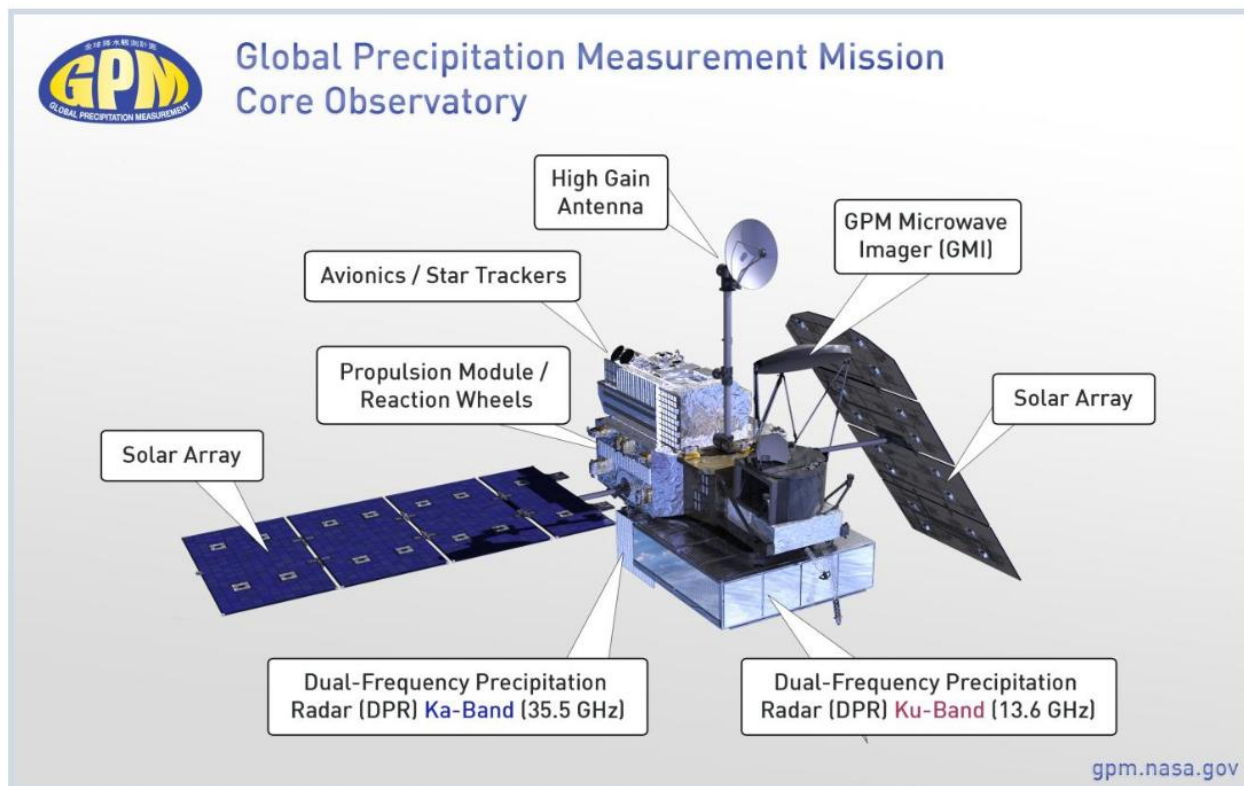
Son aquellos instrumentos que emplean sensores que detectan la radiación en longitudes de onda mediante radiómetros y radares, estos emiten señales de microondas hacia la tierra, cuando las ondas impactan la superficie de la atmosfera y se cruza con eventos de lluvia en condición liquida o sólida, la señal es reflejada hacia el satélite y dicha información proporciona datos de un evento que se está produciendo en un área y espacio determinado.

La NASA cuenta con los sensores TRMM y GPM los cuales son los más usados para la toma de información de eventos climatológicos de precipitación en diferentes estudios a nivel

mundial. En la Figura 7 se muestra la cómo se toman los datos en sensores de la nasa de manera automática.

Figura 7

Ilustración del pluviógrafo y extracción del tambor del pluviógrafo.



Nota. Tomado de "Global precipitation measurement", por GPM, 2014, <https://gpm.nasa.gov/missions/GPM/core-observatory>

4.4 Estudios Hidrológicos

Los estudios hidrológicos permiten entender la dinámica de las precipitaciones en diferentes zonas y definir los caudales que se pueden presentar en el área de una cuenca hidrográfica para diferentes periodos de retorno (T_r); los periodos de retorno corresponden a la probabilidad de ocurrencia que se puede producir de un evento, en un periodo de tiempo determinado.

Los estudios hidrológicos son fundamentales a nivel mundial, ya que proporcionan datos esenciales para investigaciones hidráulicas, ambientales y climatológicas. Estos estudios permiten identificar las condiciones geomorfológicas y climáticas de las cuencas hidrográficas, estimar caudales y predecir lluvias. Su aplicación incluye el diseño de obras hidráulicas, el análisis de ríos, la gestión de riesgos por inundaciones o inestabilidad en los suelos, la predicción de fenómenos climáticos como El Niño y La Niña, estudios hidrogeológicos, y la evaluación de la sostenibilidad ambiental frente al cambio climático.

Los estudios hidrológicos requieren una base sólida de datos confiables, ya que esto influye en la precisión de los resultados. Es importante tener en cuenta que, aunque las condiciones del entorno de una cuenca hidrográfica sean similares, las características topográficas pueden generar variaciones en los patrones de lluvia en áreas pequeñas dentro de la misma cuenca. Por lo tanto, los datos deben ser lo más cercanos posible a la realidad de la región analizada y presentar una buena correlación entre ellos.

En Colombia, algunas cuencas hidrográficas cuentan con varias estaciones climatológicas que permiten obtener una buena cantidad de información. Sin embargo, en otras regiones dentro de una misma cuenca, la falta de estaciones de monitoreo de las variables climáticas representa un desafío para el desarrollo de estudios hidrológicos.

4.5 Bases teóricas Hidrológicas

En hidrología, un dato o muestra se refiere a cualquier medición o registro que se tomó a un evento climatológico de precipitación que sucedió en un periodo determinado, son de vital importancia dado que las lluvias alimentan los cuerpos de agua, (ríos, quebradas, lagos, ciénagas, embalses, acuíferos y flujos de infiltración). Para la variable de precipitación, los datos permiten conocer las características de las lluvias que han sucedido en el área de una

cuenca hidrográfica y cuáles han sido los eventos máximos y mínimos que se han producido en una época específica.

En las estaciones climatológicas, los datos de pluviosidad incluyen mediciones del volumen de lluvia acumulada en una zona determinada, mientras que las limnigráficas determinan el caudal transitorio, que es el volumen de agua que fluye en un área durante un tiempo específico.

Para medir el nivel del agua en ríos o lagos, se utilizan limnímetros, los cuales son instrumentos que proporcionan mediciones directas. La velocidad y dirección del flujo del agua se registran con correntómetros. Por otra parte, los pluviómetros son utilizados para medir la cantidad de lluvia en una zona específica, además de estos parámetros, se registran datos como los niveles de agua, la temperatura del agua y las concentraciones de contaminantes o nutrientes, entre otros parámetros relevantes.

Las muestras o datos en hidrología son fundamentales para entender el comportamiento del agua en diferentes entornos, así como para realizar evaluaciones de recursos hídricos, modelado hidrológico, predicción de eventos extremos como inundaciones o sequías y estudios de calidad del agua. Los datos tomados de estaciones climatológicas y sensores satelitales deben ser analizados mediante procesos estadísticos, los cuales se definen según la información disponible, calidad y análisis a realizar.

Con los datos de las estaciones climatológicas del IDEAM y de los sensores satelitales de la NASA, se debe realizar diferentes validaciones y comprobaciones estadísticas que permiten definir la bondad de las muestras, la cantidad de datos faltantes, el llenado de las muestras que sean requeridas, correlación para las diferentes fuentes de información y así poder determinar la diferencias o similitudes entre las muestras obtenidas.

5 Estado del arte

Como se ha visto a lo largo del documento, esta investigación realizó el análisis de datos climatológicos con el fin de determinar la confiabilidad de la información de entidades nacionales respecto a información satelital recopilada con organismos internacionales, este estudio resulta destacable no solo para la región aledaña a la cuenca del río Lebrija, sino en general, para diferentes regiones del país, en consecuencia, y para tener una información detallada y veraz se han analizado estudios similares que se han realizado tanto a nivel local, nacional e internacional, los cuales servirán de base para el estudio aquí realizado, tales estudios se resumen a continuación:

5.1 Antecedentes

En la región Latinoamericana es muy común realizar modelos hidrológicos basado en información estaciones climatológicas (Pluviográficas) localizadas en el área de diferentes cuencas hidrográficas en sus diferentes territorios, sin embargo, la escases de información en algunas zonas, baja aceptación y correlación de las muestras obtenidas, han llevado a que se desarrollen diferentes estudios para implementar nuevas técnicas de extracción de datos climatológicos de sensores satelitales.

Diferentes autores han desarrollado investigaciones acerca de la disponibilidad de información de sus cuencas de estudio, cuáles son los problemas encontrados con base a la disponibilidad de información de las estaciones climatológicas y cuáles han sido las metodologías implementadas para la extracción de información de diferentes sensores satelitales aplicados a diferentes estudios de caso de diversas cuencas hidrográficas.

Tal es el caso de Conde (2018) quien desarrollo su tesis para obtención de su título de ingeniera agrícola, la cual se basó en el estudio de estimación de precipitación tomando

información de los sensores satelitales TRMM y GPM obtenidos del geoportal Giovanni de la NASA y los comparó con la información de 5 estaciones meteorológicas (INIA CANAAN, Cuchoquesera, La Quinoa, San Pedro de Cachi y Wayllapampa) localizadas en la provincia de Huamanga en Perú, los resultados del estudio presentan un ajuste aceptable, considera que su implementación permite tener una disponibilidad de información continua en una temporalidad diaria, el comportamiento es similar cualitativamente entre ambas fuentes de información, presentan coeficientes de correlación de Pearson cercanas a 1.

Por otra parte Vilchis et al. (2015) realizó un artículo con la modelación hidrológica de la cuenca del río Amacuzac en México, basado en información de estimaciones de precipitación con sensores satelitales hidrometeorológicos, el cual mediante esta investigación manifiesta que los radares (estaciones meteorológicas) que se encuentran en diferentes portales de bases de datos abiertos, al igual que los sensores satelitales proporcionan información valiosa en la estimación de estudios de la variable de precipitación.

A su vez indica que las técnicas satelitales más usadas, están basadas en dos tipos de sensores, el pasivo de microondas y el infrarrojo, siendo los sensores pasivos de microondas aquellos que proporcionaron estimaciones pluviométricas más exactas comparados con los datos de los sensores de luz visible o infrarroja, el resultado del análisis presento datos de estimación de caudales de la cuenca y afirma el autor que las estimaciones hidrológicas de la variable de precipitación basadas en datos por radar (Pluviógrafo) a nivel diario producen resultados satisfactorios comparados con los estimados por satélite; sin embargo, en la cuenca del río Amacuzac los datos de sensores satelitales no presentaron buena correlación, debido a la ausencia de información por el bloqueo orográfico de la señal del radar, este bloqueo se refiere a la interferencia o atenuación de una señal de radar cuando encuentra obstáculos geográficos, como nubes, montañas u ondulaciones del terreno a diferentes alturas.

5.2 Estudios locales

Dentro de los estudios que se han realizado a nivel local en la región de Santander, en Colombia está el documento denominado “Caracterización hidrológica y de calidad del agua de la quebrada Tonchalá de la ciudad de San José de Cúcuta (Norte de Santander)” en la que sus autores Díaz y Guevara (2020) obtuvieron los caudales máximos, medio y mínimos diarios del exutorio de la subcuenca de la quebrada Tonchalá desde 1989 hasta el 2019, las cuales presentaron dos fuertes crecidas por el fenómeno del niño en 1999 y 2010, así mismo, analizaron índice de vulnerabilidad al desabastecimiento hídrico (IVH) en el concluyeron que se encuentra dentro de la categoría de “Muy Alto”, lo que según los autores representa que es una cuenca altamente vulnerable ante eventos amenazantes o fenómenos como el Niño, finalmente, en términos medioambientales, ha presentado calidades “mala2 y “regular” conforme al índice ICA, niveles que se evidencian cierto grado de contaminación producto de los vertimientos municipales de la ciudad de San José de Cúcuta – Norte de Santander.

De tal forma, este estudio resulta representativo pues analiza las condiciones climáticas y los niveles de la cuenca que se encuentra en un región cercana al río Lebrija, en consecuencia, pese a que las condiciones geográficas de cada territorios varían, lo cierto es que la caracterización por los autores realizada arroja datos importantes sobre los niveles de precipitación y contaminación de un afluente importante de la región, los cuales puede dar información relevante para esta investigación.

Por otro lado, se encuentra el estudio denominado “Evaluación de la incidencia del cambio climático en la cuenca Lebrija alto aplicando Hydrobid” en la que sus autores Quiroga y Duarte (2021) realizaron una valoración de las afectaciones que genera el cambio climático, asociado a los periodos de altas precipitaciones, sequias y disponibilidad de agua en la cuenca alta del río Lebrija, que fue realizado a través de una herramienta hidrológica denominada “HydroBID”, que se encarga de realizar simulación de caudales efluentes generados por la

cuenca alta y la correlación con la cantidad de agua requerida por la cuenca teniendo en cuenta la demanda de los diferentes usos que se puedan presentar y su caudal ecológico, por medio de la información climatológica del IDEAM. A partir de este estudio y simulando condiciones climáticas adversas asociadas a los efectos del cambio climático los investigadores pudieron determinar que tiene una mayor complejidad, debido a que presentó un caudal menor en comparación con la demanda de la cuenca, siendo el mes de agosto el más bajo con un caudal de 5,38 (m³/s) mientras que la demanda de la cuenca obedece un caudal de 5,62 (m³/s).

Conforme lo anterior, es evidente que las condiciones climáticas variables, tal como las que se suelen presentar en los últimos años, originadas entre otros factores por el cambio climático, generan un aumento en el cauce de los ríos, por tanto, el estudio aquí presentado permite tener un panorama sobre las posibles condiciones que tendría la cuenca del río Lebrija y estos datos pueden servir de referencia para ser comparados con los datos recopilados y analizados en esta investigación.

5.3 Estudios nacionales

Ahora bien, en cuanto a los estudios que se han realizado en otras regiones del país, se encuentra el denominado: “Variabilidad climática en la planificación hídrica de la cuenca del río Cravo Sur (Casanare, Colombia)” por medio del cual sus autores Ruiz et al. (2022) buscaron evaluar la inclusión de la variabilidad climática natural en los proyectos del Plan de Ordenación y Manejo de Cuenca Hidrográfica (POMCA) del río Cravo Sur (Casanare, Colombia) a través del Índice de Evaluación del Sistema de Gestión (IESG). A partir de esta investigación los autores concluyeron que los proyectos del POMCA del río Cravo Sur en su mayoría incluyen la variabilidad climática, pero un porcentaje cercano al 30% requieren ser revisados desde su formulación y gestión en función de la una dinámica temporal del clima. Además, se requiere

ampliar la cobertura de los proyectos para mejorar las decisiones de corto, mediano y largo plazo para garantizar la gestión integral del recurso hídrico en un territorio con información escasa.

Este estudio resulta relevante para la investigación en particular aquí desarrollada pues a través del POMCA de esta cuenca hídrica de la región de los llanos orientales se permite analizar diferentes variables climatológicas en el marco de los procesos de ordenamiento y gestión de las cuencas, y los hallazgos aquí identificados permiten entender de qué forma los datos meteorológicos recolectados puede servir de base para la toma de decisiones por parte de autoridades locales.

Por su parte, se encuentra la tesis denominada “Modelación hidrológica e hidráulica acoplada de la cuenca media y baja del río Magdalena”, en la que su autor Sánchez (2017) realizó una modelación acoplada reproducir el comportamiento observado en la cuenca media y baja del río Magdalena durante el período de tiempo comprendido entre el 1/1/2010 y el 31/03/2011. De esta manera el autor pretende tener un modelo que permita facilitar la transición hacia un proceso de toma de decisiones basado en el conocimiento, teniendo en cuenta que esta toma de decisiones históricamente en el país se ha basado en observaciones en un número limitado de ubicaciones a lo largo de la cuenca hídrica y no en un conocimiento sólido del sistema fluvial en su conjunto, por tal razón, el modelo identificó que existen falencias en la información disponible y permite que pueda ser replicado en otros ríos y considerar la inclusión de otros elementos como el transporte de sedimentos, la calidad del agua y el modelamiento del hábitat acuático.

Este estudio resulta de vital importancia pues aborda un modelo hidrológico a partir de información histórica de la cuenca del río y estos datos pueden permitir generar comparaciones frente a los datos que se arrojan en la cuenca del río Lebrija, teniendo en cuenta a su vez que

este último es un afluente del Magdalena y en la cuenca media comparten similitudes geográficas que resultan relevantes para esta investigación.

5.4 Estudios internacionales

En torno a los estudios de análisis de información climatológica en cuencas hídricas realizada en otros países de la región, es posible encontrar la investigación denominada “Caracterización Hídrica de la Cuenca del Arroyo Vaquerías (Córdoba, Argentina) Mediante el Modelado Hidrológico e Información Satelital”, mediante la cual su autor Weber (2020) buscó realizar la caracterización mediante el modelado computacional, utilizando como fuente de información para la precipitación se recurrió al producto CHIRPS (Climate Hazards group InfraRed Precipitation with Station data), y para la evapotranspiración, al producto MOD16A2, de tal forma que el modelado se realizó a través del modelo semidistribuido TOPMODEL, que en forma simplificada considera la contribución al escurrimiento de los flujos superficial y subterráneo. Como conclusión del estudio el autor determinó que, pese a la falta de información previa existente, con los datos recolectados a través de las herramientas referidas fue posible realizar el modelo y completar la caracterización de la cuenca hídrica que resulta de vital importancia para la Reserva Natural Vaquerías.

En este sentido, esta investigación resulta de importancia pues a partir de datos de origen satelital se realizaron las comparaciones necesarias para aplicar un modelo que permitió obtener una caracterización completa que puede servir como fuente de información para las autoridades locales, y en tal sentido, la aplicación de modelos y comparación de datos históricos como los analizados en esta investigación pueden resultar en información de vital importancia para la toma de decisiones en torno a la cuenca hídrica de la región santandereana.

Finalmente, se tuvo en cuenta la investigación realizada en la región centro de Tamaulipas, México, nombrada por sus autores Vázquez et al. (2021) como “Modelo hidrológico, calidad del agua y cambio climático: soporte para la gestión hídrica de la cuenca del río Soto la Marina” en la que se analizó el comportamiento de la cuenca, a efectos de simular la vulnerabilidad hídrica, evaluando la calidad del sistema espacio-temporal: avenidas y sequías, mediante la determinación de demanda bioquímica de oxígeno, demanda química de oxígeno y oxígeno disuelto, además de temperatura, en 15 sitios de la cuenca del río Soto la Marina, con modelado del abatimiento del oxígeno disuelto. Este modelo hidrológico de disponibilidad y calidad del agua realizado por los autores resulta de vital importancia para la región pues esta red fluvial esta alimentada por tres ríos, además de tener incidencia directa con la presa Vicente Guerrero, en consecuencia, contar con datos reales y aproximaciones a las condiciones y cambios climatológicos de esta cuenca permute tener control de los efectos y riesgos para las poblaciones cercanas, del mismo modo que se pueden ver afectados los habitantes de la cuenca del río Lebrija en este estudio analizada, por tanto, los hallazgos encontrados en este estudio realizado en México pueden servir de referencia para el caso colombiano en objeto de estudio.

6 Aspectos metodológicos

En el marco del desarrollo de esta investigación se relacionan y abordan diferentes aspectos con el uso y disponibilidad de los datos para la variable de precipitación en la cuenca media del río Lebrija, utilizando bases de datos de diferentes fuentes de información, las cuales fueron analizadas estadísticamente para determinar cuál es la correlación entre los datos obtenidos.

Esto permite al lector identificar las dinámicas encontradas entre las muestras obtenidas, para el desarrollo de estudios en cuencas hidrológicas, determinando la posibilidad de usar información climatológica que no provenga de estaciones convencionales en zonas donde no hayan sido instrumentadas y aquellas donde no se tiene buena confiabilidad entre los datos de la información disponible de las estaciones existentes. Por otra parte, a continuación, se abordan los principales aspectos metodológicos que se llevaron a cabo, para identificar los resultados de los análisis realizados, los cuales se orientaron en este proceso investigativo.

6.1 Paradigma de investigación

La presente investigación se enmarca dentro del paradigma positivista, el cual es una corriente de pensamiento filosófico basado en el métodos científico el cual fue considerado por primera vez, en el siglo XIX por August Comte, según Meza (2003) no admite cómo validos otros conocimientos sino todos aquellos que proceden de ciencias empíricas, se considera una alternativa epistemología dialéctico-crítica, la cual se basa en muestras de operaciones observables de eventos sociales, desarrolladas en la evolución de las ciencias modernas de la naturaleza humana, que para el caso de esta investigación forman parte de las variables climatológicas que se monitorean para determinar condiciones meteorológicas que se desarrollan en las estaciones seleccionadas para el estudio.

El paradigma positivista, tiene una naturaleza de estudio cuantitativo, la cual se caracteriza, según Hernández et al. (2014) por considerar una precisión en los resultados, busca comprender la causa de fenómenos sociales, independiente de los individuos que se relacionen, utiliza análisis matemáticos y estadísticos, en la obtención de resultados para obtener respuestas a problemas identificados.

Las muestras utilizadas para el desarrollo de esta investigación se analizaron mediante procesos estadísticos para definir la correlación entre los datos de la variable de precipitación y definir qué condiciones conceptuales pueden llegar a existir en el planteamiento del problema identificado, hasta presentar una visión científica de los procesos analizados que permitan definir la viabilidad en la implementación de las fuentes de información usadas.

6.2 Enfoque de la investigación

Tal como lo ha establecido Hernández et al. (2014) la presente investigación tiene un enfoque cuantitativo el cual permite recolectar datos y analizarlos sobre variables de fenómenos cuantitativos, sus principales características están dadas por planteamientos acotados, mide fenómenos, utiliza estadísticas, y prueba teorías e hipótesis, se basa en el desarrollo de procesos deductivo, secuenciales, probatorios, y analiza la realidad objetiva, sus principales bondades se enmarcan en la generalización de resultados, control sobre fenómenos, precisión, replica y predicciones.

Para desarrollar el enfoque se debe identificar el problema, en el caso de esta investigación es la incertidumbre entre la información disponible de los parámetros climatológicos de la variable de precipitación de las estaciones del IDEAM y su escasez o baja disponibilidad en algunas zonas del país, basado en que el enfoque cuantitativo se enfoca en la medición y análisis de parámetros identificados en una muestra, se pretende plantear resultados de forma objetiva y coherente con la problemática identificada, que den soluciones concretas para la implementación de nuevas fuentes de información.

6.3 Temporalidad

Esta investigación se desarrolla en el marco de un estudio observacional, para una cuenca hidrográfica existente, es de corte longitudinal, considerando que se han realizado la toma de datos de información climatológica de precipitación en el área establecida por la localización de las estaciones climatológicas en la zona de estudio, con lo cual se ha determinado una temporalidad de análisis a nivel diario y mensual multianual con un rango de tiempo de mínimo 20 años de las muestras obtenidas.

Los estudios de corte longitudinal permiten hacer seguimiento de procesos en un tiempo concreto o establecido, en el caso de esta investigación se estableció que para obtener un análisis ideal es necesario que las estaciones tengan información diaria de datos de precipitación y que su periodo tenga al menos 20 años de información, los datos faltantes se llenaran mediante procesos de validación estadísticos.

6.4 Análisis de datos

Los datos de las muestras obtenidas de las tres fuentes de información (1- Portal de datos abiertos DHIME del Instituto de Hidrología, Meteorología y Estudios Ambientales IDEAM, 2-portal de bases de datos de Power Data Access Viewer de (NASA Prediction of Worldwide Energy Resources) y 3- GIOVANNI of The NASA. (The Earth Science Data Systems (ESDS)) se analizaron mediante el desarrollaron scripts personalizados en lenguaje de programación Python que permiten realizar el manejo de bases de datos, este proceso de validación de datos permitió definir la localización geográfica de las estaciones, los datos disponibles para cada año, en caso de no contar con una continuidad de información en años seguidos o que existan más de un año sin información la estación se descarta. Para las estaciones seleccionadas se hace el llenado de los datos faltantes mediante la interpolación lineal, una vez llenos los datos se hace la validación de la correlación entre las muestras de las tres fuentes de información

mediante procesos estadísticos, definiendo cual es la diferencia entre los datos y determinando cual es la correlación entre los mismos.

Para los análisis de datos se han implementado rutinas en el lenguaje de programación Python, mediante la plataforma de ejecución Jupyter Notebook, con las cuales se han realizado los procesos estadísticos de validación de los datos. Estas rutinas permitieron identificar los años, meses y días sin información en las seis estaciones seleccionadas para la investigación, determinando con precisión los periodos que presentan vacíos en la serie.

El llenado de estos datos faltantes no solo es necesario para permitir una adecuada comparación con las series continuas provenientes de sensores satelitales (como los datos de la NASA), sino también para garantizar la integridad y consistencia de los análisis estadísticos posteriores. Una serie con datos completos es esencial para calcular indicadores representativos, aplicar modelos de regresión o predicción, y evaluar tendencias o estacionalidades de forma fiable. Por tanto, el proceso de completado de datos constituye una etapa crítica dentro del flujo de análisis, y fue abordado con criterios técnicos orientados a preservar la representatividad climática de las series corregidas.

El análisis se estructuró en tres bloques desarrollados en lenguaje de programación Python. El primero correspondió al llenado de datos faltantes, utilizando bibliotecas como pandas, numpy, scipy y statsmodels, aplicando una metodología híbrida de interpolación: se empleó inicialmente interpolación lineal, seguida de una imputación ajustada por estacionalidad mediante la descomposición STL (*Seasonal-Trend decomposition using Loess*), lo que permitió preservar la estructura temporal de la serie. El segundo bloque implementó análisis estadísticos de la precipitación, incluyendo estadísticas descriptivas y coeficientes de correlación de Pearson (r), Spearman (ρ) y Kendall (τ), con el fin de evaluar la relación entre las series de datos satelitales y las observaciones locales. Finalmente, el tercer bloque se basó en dos pilares fundamentales: la validación cruzada mediante el cálculo del error cuadrático medio

(RMSE), utilizando el método *K-Fold*, y la modelación de series temporales a través del modelo SARIMA, evaluando la estacionariedad de la serie con la prueba Augmented Dickey-Fuller (ADF).

6.5 Metodología propuesta.

La presente investigación se enfoca en analizar la información de los datos climatológicos de la variable de precipitación dentro del área de la cuenca media del río Lebrija, para comparar entre sí, con información obtenida tanto de los datos abiertos las fuentes del IDEAM, como de los sensores POWER Data Access Viewer Prediction of Worldwide Energy Resources y Giovanni EARTH DATA de la NASA. A través de la combinación de los conjuntos de datos, se pretende obtener una visión integral sobre las dinámicas de precipitación en la región de estudio y evaluar la confiabilidad de las fuentes de datos.

Las tres fuentes de información mencionadas corresponden a bases de datos abiertos de variables climatológicas a nivel nacional y global, todas cuentan con una temporalidad multianual diaria y cuentan con datos para otras variables, sin embargo, para efectos de esta investigación solo se analizarán los datos de precipitación, los cuales previamente se identificaron según su localización, disponibilidad de años y nivel de aceptación.

Los estudios que se desarrollan en la presente investigación surgen de la necesidad de encontrar diferentes fuentes de información para el desarrollo de análisis hidrológicos, climatológicos y relacionados con el cambio climático y el medio ambiente, debido a que una de las principales limitaciones en este campo es la falta de estaciones meteorológicas en diferentes regiones de Colombia y otros países a nivel global, lo que limita la capacidad de monitorear y comprender el clima en diferentes regiones del mundo.

Esto puede tener consecuencias alarmantes para la gestión de los recursos hídricos y la adaptación al cambio climático. En este contexto, la utilización de tecnologías de teledetección y la integración de datos de múltiples fuentes puede proporcionar una visión más completa y precisa del clima en áreas que no cuentan con registros de estaciones meteorológicas terrestres.

Esta investigación, cómo se ha mencionado previamente, se desarrolla dentro del área geográfica de la cuenca medía del río Lebrija, localizada entre los departamentos de Cesar, Santander y Norte de Santander en Colombia, con el fin de delimitar de manera precisa la información a analizar, se ha definido un marco espacial claro que permitirá abordar adecuadamente los datos climatológicos disponibles, a continuación se realiza una descripción detallada de la metodología propuesta para el desarrollo de la identificación de datos, selección de estaciones, comprobación de datos faltantes, llenado de datos, análisis estadístico, correlación de las muestras, validación y comparación entre las tres fuentes de información propuestas.

6.6 Selección de información de las estaciones climatológicas que se encuentren localizadas en el área de la cuenca medía del Río Lebrija.

Tal cómo se estableció en el primer objetivo específico se debe realizar el proceso identificación de los datos que se encuentran en el área de la cuenca hidrográfica del río Lebrija medio, los cuales corresponden a la variable de precipitación, dichos datos según la fuente de información de donde se extrajeron, deben pasar por un proceso de comprobación y validación de las muestras, para posteriormente llenado los datos faltantes. Dado lo anterior se hace la descripción de todos los elementos que se han considerado en la presente investigación para el desarrollo de dicho análisis.

6.6.1 Identificación de estaciones climatológicas.

Para el desarrollo de estudios climatológicos e hidrológicos, especialmente en el territorio colombiano, lo primero que se debe realizar es la identificación de las condiciones geográficas y geomorfológicas del área de la cuenca, de tal manera que se pueda establecer los límites del área que conforman la cuenca hidrográfica.

Una vez definidos los límites mediante las líneas divisorias de agua, se identifican las estaciones ubicadas dentro del área de la cuenca o en sus proximidades. Dado que existen regiones con escasa cobertura de estaciones climatológicas, se procede a identificar la totalidad de estaciones disponibles y, posteriormente, se depuran aquellas que presentan una mejor calidad y continuidad en la información. Para cualquier análisis, es fundamental contar con registros consecutivos de al menos 20 años; en esta investigación, se han descargado series de datos correspondientes a ese periodo con fines de validación.

Al identificar la zona de estudio, se realiza la verificación de las estaciones climatológicas que se encuentran instrumentadas dentro del área de la cuenca hidrográfica. En caso de tratarse de una cuenca de menor extensión, se consideran también las estaciones más cercanas y la información disponible en cada una. Tras la selección inicial, y de acuerdo con el parámetro de interés; en este caso, la precipitación, se revisa la disponibilidad temporal de los datos, así como el estado operativo de la estación o, en su defecto, hasta qué fecha estuvo en funcionamiento.

En cuanto a las estaciones climatológicas del IDEAM se ingresa al portal web DHIME, se seleccionan las variables, en este caso precipitación, seguido se establece la temporalidad de años que se van a descargar (se debe tener en cuenta que la aplicación no permite descargar más de 30 años de información), adicionalmente se seleccionan los datos a descargar (se pueden descargar datos medios, máximos, mínimos y totales), para efectos de la

investigación se descargan los datos de PTPM que corresponde al Promedio Diario de Precipitación el cual lo suministra de manera diaria horaria, es decir en valores que en caso de esta investigación se tomaron desde las 00:00 horas del 01 de enero de 2004 hasta las 24:00 horas del 31 de diciembre de 2024.

El IDEAM establece que para realizar un buen análisis los datos suministrados deben tener un nivel de aprobación que oscila entre 1200 a 800 entendiéndose que 1200 es una muy buena y 800 una muy baja aprobación, esto quiere decir que de manera inicial uno puede depurar que tan buenos pueden estar tomados los datos, solo descargando la información sin ninguna validación, esto permite seleccionar de manera oportuna todas aquellas que estén con valores cercanos a 1200.

En la Figura 8 se muestran la selección de todas las estaciones que se identificaron localizadas dentro del área de la cuenca hidrográfica y en la Figura 9 todas aquellas que fueron seleccionadas, ya sea porque tienen datos consecutivos de al menos 20 años, porque la estación está funcionando o vigente y su nivel de aprobación es cercano a los 1200.

Una vez definidas las estaciones climatológicas se identifica su localización geográfica en coordenadas planas, si llegaron a estar en geográficas, se debe relajar la conversión. Con la localización se pretende ubicar espacialmente dentro de los portales web de los sensores de la NASA la ubicación exacta para descargar los datos de las dos fuentes de información, cabe resaltar que se pueden descargar los datos de las estaciones depuradas o finales, sin embargo, en esta investigación se descargaron la misma cantidad de datos para las tres fuentes de información y solo se analizó la que se muestra en Figura 9.

El portal web The Power Data Access Viewer de la NASA permite ubicar dentro del mapa global el punto de interés o introducir las coordenadas exactas, en nuestro caso las estaciones tienen su ubicación en coordenadas planas (Longitud y Latitud) por ende fue preciso

que estén en la misma localización, se selecciona el nivel temporal (Diario, mensual o anual), para efectos de esta investigación diarios y el tiempo de descarga de las muestras, el cual es el mismo de las estaciones del IDEAM (00:00 horas del 01 de enero de 2004 hasta las 24:00 horas del 31 de diciembre de 2024), se selecciona el parámetro de precipitación y se descarga en formato (.csv) que es el equivalente para abrir e libros de Excel.

Por otra parte para los sensores de Giovanni de la NASA, la descarga de información es similar a la anterior, solo que previamente para ingresar al portal se debe realizar el registro y autenticación en el EARTH DATA, una vez se ingresa con su usuario y contraseña, se hace el procedimiento con la misma secuencia anterior de localización, nivel temporal, parámetro, solo que se adiciona la selección del sensor que se va a descargar en nuestro caso TRMM y se descarga la información en el mismo formato (.csv.)

Cabe resaltar que para los datos de los sensores satelitales ya sean The Power Data Access Viewer y Giovanni ambos de la NASA no se deben llenar los datos faltantes, dado que las muestras las toman diarias de señales de microondas y no se presentan datos faltantes.

6.6.2 Definición y llenado de los datos faltantes

Una vez identificados los datos faltantes, mediante diferentes técnicas estadísticas es posible llenar los datos diarios que no tienen registros, esto solo se debe realizar para los datos obtenidos de información primaria de estaciones climatológicas, no aplica para los datos de sensores satelitales, existen varias técnicas de llenado, la más común es utilizar la media de los datos, puede ser ponderando entre los datos vecinos o la media total, sin embargo, existen otros métodos, más analíticos que se aplican de una manera más efectiva para las condiciones del estudio.

Según, Valencia (2020), dentro de su proyecto hidrológico Balance hídrico para la plantación agrícola del arroz en el Valle del Cauca, propone tres métodos para el llenado de los datos faltantes, el primero es la Razón normal (RN) el cual interpola datos tomando información de estaciones vecinas, detecta datos erróneos, saltos en la media, tendencias, rellena las muestras y presenta la corrección de los errores aislados.

Por otra parte, realiza el análisis armónico de las series de Fourier (SF), el cual mediante análisis trigonométricos descompone los senos y cosenos de las muestras para obtener el dato faltante y por ultimo realiza la interpolación de la distancia inversa al cuadrado (IDW) el cual busca estimar los valores de una variable en puntos desconocidos a partir de valores conocidos en puntos cercanos. El llenado de los datos faltantes se realizará para efectos de la presente investigación mediante el método de la Razón Normal (RN).

El proceso inicia con la revisión y limpieza de datos, el cual es el primer análisis que se realiza, una vez se seleccionadas las estaciones climatológicas y descargado los datos, se puede realizar mediante la implementación de diferentes herramientas ofimáticas, identifica posibles errores o valores atípicos (outliers) de las muestras que se salen de la tendencia de los mismos.

Una vez seleccionadas las estaciones climatológicas localizadas dentro del área de la cuenca hidrográfica que se van a usar y después descargar los datos de las tres fuentes de información, se procede con la revisión de los datos faltantes de las estaciones y llenado de los mismos, cabe mencionar que este segundo proceso se debe realizar para poder realizar el análisis de correlación entre las muestras de los datos de la NASA y del IDEAM, dado que los datos de los sensores satelitales no tienen datos faltantes y por ende se debe armonizar para tener la misma cantidad de datos en los análisis posteriores.

Existen diferentes formas de analizar la información de datos faltantes, cabe mencionar que los datos descargados de las tres fuentes de información son de tipo diaria horaria, con lo cual se toman los valores de precipitación máximos presentados por horas para cada día de estudio, esto implica que por cada día se tienen muestras de 24 datos, los valores que son 0.00 son aquellos que muestran tendencias de que no hubo lluvias en ese rango, los valores que estén vacíos corresponden a muestras que no se tomaron y los que estén con valores superiores a 0.00 significan que ocurrió un evento con una intensidad específica.

Tanto para el llenado de los datos, como la identificación de los datos faltantes, se utilizaron herramientas ofimáticas, mediante la implementación de análisis de datos los cuales se desarrollaron mediante el uso de códigos de lenguaje de programación tales como visual Basic VBA y Python, los cuales se describen a continuación.

6.6.2.1 Validación de Datos faltantes mediante Python.

Adicionalmente se desarrolló un código mediante la aplicación del lenguaje de programación de Python el cual es utilizado principalmente en aplicaciones web, para el desarrollo de softwares, análisis de la ciencia de datos y aprendizaje automático (machine learning) o también conocida como inteligencia artificial.

A diferencia del código de VBA, éste elimina los datos y deja las celdas en blanco, cabe mencionar que cuando no se tienen registros de una fecha u otra, los datos de las estaciones climatológicas provenientes del IDEAM, se saltan al siguiente día que tiene registros, por ende se pierde la secuencia diaria horaria, con lo cual se pretende crear una celda con el día faltante, al igual que en el análisis de VBA se crea una hoja nueva y se incluyen los valores totales con información y las nuevas celdas.

Cabe mencionar que este código por sí solo no tiene casi ninguna validez, para la identificación de los datos faltantes y llenado de los mismos se deben completar otros dos

códigos y los resultados los muestra en el análisis final, se debe prestar atención en el Título 6.6.2.2 para visualizar los resultados finales de la validación y llenado de los datos faltantes.

6.6.2.2 Asignación de las celdas faltantes en blanco mediante Python.

Seguido del proceso de identificación de los datos faltantes el proceso de análisis mediante el uso de Python, se complementó mediante la elaboración de otro código que después de validar los datos faltantes y crear unas celdas nuevas procede a la creación de todas las celdas dejarlas vacías para de esta manera en un siguiente código realizar el proceso de llenado de los datos faltantes. Al igual que con el código anterior los resultados se pueden ver una vez se culmine el último proceso que se puede identificar en el Título 6.6.2.2.

6.6.2.3 Llenado de datos faltantes mediante Python.

El proceso se culmina una vez realizado los procesos anteriores, a través del llenado de los datos faltantes mediante el uso de Python, se complementó mediante la elaboración de otro código el cual usa dentro de sus análisis de datos una secuencia para validar, completar y analizar estadísticamente las muestras, para presentarlas de manera gráfica y detallada comparando los datos sin rellenar y los datos completos. El proceso tiene en cuenta la siguiente secuencia lógica.

- **Generación de Rango Completo de Fechas y Relleno de Datos Faltantes;** lo realiza mediante el análisis de **Interpolación Lineal**, la cual consiste en trazar una recta que pasa por dos puntos conocidos, en este caso datos de precipitación (utiliza estaciones vecinas) y calcula los valores intermedios que pasan cercanos a la línea trazada.
- **Identificación y Marcado de Valores Anómalos;** se procede con el cálculo de los cuartiles (Q1) y (Q3), de la distribución de los valores de precipitación y se calcula **Rango Intercuartílico (IQR)** que corresponde a la resta de $Q3 - Q1$,

posteriormente define los límites de los valores atípicos que estén fuera de los rangos ($Q1 - 3 * (IQR)$) y ($Q3 + 3 * (IQR)$), el valor (3) de la fórmula corresponde a un factor multiplicador para detectar valores extremos.

- **Descomposición STL (Seasonal and Trend decomposition using Loess);** este análisis determina la **tendencia** que corresponde a la dirección general a largo plazo de la serie, además define la **estacionalidad** la cual analiza las fluctuaciones repetitivas en intervalos regulares (anuales) y los **residuos** de los datos que se producen después de eliminar la tendencia y la estacionalidad.
- **Visualización de Componentes y Análisis de Residuos;** una vez analizado los tres análisis anteriores el código muestra de manera grafica la identificación de las muestras y presenta los resultados de la Interpolación Lineal, Rango Intercuartílico (IQR) y Descomposición STL.
- **Exportación de Datos;** El proceso de llenado y validación de datos culmina con que los datos procesados, incluyendo la identificación de los valores atípicos, se exportan a un archivo Excel para su almacenamiento o análisis posterior.

6.7 Contrastar la información disponible de las tres fuentes seleccionadas

Conforme a lo que se estableció en el segundo objetivo específico y en base en los análisis anteriores, se identificaron las estaciones que se utilizarán en el análisis, las cuales pasaron por un riguroso proceso de depuración. Este proceso permitió seleccionar las estaciones con la mejor disponibilidad de datos, según el mayor nivel de aprobación. Posteriormente, se validaron los datos recopilados y se procedió a completar aquellos que estaban incompletos.

Por otra parte, conforme al segundo objetivo específico, es necesario realizar las comprobaciones de los valores obtenidos, tomando como referencia los datos proporcionados

por el IDEAM y los sensores satelitales de la NASA, con el fin de asegurar la precisión y fiabilidad de los mismos. En este sentido, a continuación, se describe detalladamente todos los elementos y procedimientos considerados en la presente investigación para llevar a cabo este análisis, con el objetivo de garantizar que se utilicen los mejores datos disponibles y se sigan las metodologías adecuadas para la correcta interpretación de los resultados.

6.7.1 Análisis de estadística descriptiva de datos del IDEAM completos y de la NASA.

Los análisis de la estadística descriptiva son un conjunto de técnicas utilizados para resumir las características principales de un conjunto de datos de manera simple, se reconoce por la proyección de las mediciones de tendencia central, en la cual se destacan la media, mediana y la moda, adicionalmente las medidas de dispersión, como lo son la desviación estándar, rango y varianza, y las medidas de forma conocidas como asimetría y curtosis. Se realizará usando los datos de precipitación de ambas fuentes de información de estaciones climatológicas del IDEAM y sensores satelitales de la NASA.

Por otra parte las pruebas de normalidad son procedimientos que se utilizan para verificar si un conjunto de datos sigue una distribución normal o gaussiana, es una de las distribuciones más comunes en estadística. Se estima que si los datos no siguen una distribución normal puede afectar la validez de los resultados de ciertos análisis estadísticos, su interpretación se hace mediante tablas de valores y graficas. Por otra parte se realizará la validación para los datos de ambas fuentes de información.

Adicionalmente se realiza el análisis de tendencias y variabilidad. Este análisis estudia la tendencia temporal de los datos y su variabilidad a lo largo del tiempo, se enfoca en la determinación de cuánto puede ser la diferenciación en base a una media central, busca identificar patrones a largo plazo y entender las fluctuaciones o inconsistencias dentro de las

series de datos, se realizará usando los datos de precipitación de ambas fuentes de información de estaciones climatológicas del IDEAM y sensores satelitales de la NASA.

Para el análisis de la estadística descriptiva y validación de datos, se desarrolló un código en lenguaje de programación Python, ejecutado mediante el entorno Jupyter Notebook. En este proceso se utilizaron principalmente las bibliotecas pandas Mckinney (2010) numpy y scipy.stats, Virtanen et al. (2020) junto con herramientas de visualización como (matplotlib.pyplot y seaborn). Estas permitieron automatizar el procesamiento, análisis y representación gráfica de la información de precipitación proveniente tanto de las estaciones del IDEAM como de los sensores satelitales (NASA – POWER y Giovanni). Para el modelamiento estadístico y predictivo se incorporaron statsmodels, especialmente para análisis de series temporales, y sklearn, usada para validación cruzada y evaluación de modelos Pedregosa (2011). Estas bibliotecas son ampliamente reconocidas por su robustez y versatilidad en el tratamiento de datos científicos.

El análisis estadístico incluyó el cálculo de medidas de tendencia central (media, mediana, moda), de dispersión (desviación estándar, varianza y rango), y de forma (asimetría y curtosis). Esto permitió identificar patrones y comportamientos generales de la serie de datos.

También se aplicaron pruebas de normalidad sobre los conjuntos de datos, con el fin de evaluar si seguían una distribución normal, lo que es relevante para validar el uso de ciertos procedimientos estadísticos. Estas pruebas se interpretaron a partir de los valores p y gráficamente mediante histogramas y curvas de densidad Virtanen et al. (2020).

Para evaluar la similitud entre las fuentes de información (IDEAM y NASA), se calcularon los coeficientes de correlación de Pearson (r), Spearman (ρ) y Kendall (τ), utilizando funciones del módulo (scipy.stats) Virtanen et al. (2020). Esto permitió establecer el grado de concordancia entre las series para cada estación.

Además, se implementó un procedimiento de validación cruzada utilizando KFold de la biblioteca (`sklearn.model_selection`), con cinco particiones ($K=5$). En cada iteración se calcularon los errores cuadráticos medios (MSE) y su raíz (RMSE), con el fin de estimar la precisión del ajuste entre los datos satelitales y los observados Pedregosa (2011).

Como parte del análisis de series temporales, se utilizó el modelo SARIMAX (`statsmodels.tsa.statespace.sarimax`) para evaluar la capacidad predictiva de los datos satelitales. Previamente, se aplicó la prueba de Dickey-Fuller Aumentada (ADF), mediante (`adfuller`), para verificar la estacionariedad de las series. El modelo se configuró con parámetros estacionales (1,1,1,12) y su resumen se exportó automáticamente a un archivo .txt para su interpretación posterior McKinney (2010).

Por otra parte, se realizó un proceso de limpieza y completitud de los datos de precipitación. Inicialmente, se generó un rango completo de fechas a partir del primer y último día de la serie, identificando y llenando los vacíos mediante interpolación lineal. Este método estima los valores faltantes con base en los registros inmediatos anterior y posterior, garantizando continuidad en la serie McKinney (2010).

Todas las salidas, incluyendo tablas de resultados, métricas estadísticas y gráficas (barras comparativas, líneas de tendencia, diagramas de dispersión, descomposición estacional STL), fueron exportadas automáticamente a hojas de cálculo Excel mediante (`pandas.ExcelWriter`) y visualizadas usando (`matplotlib` y `seaborn`). Esto permitió consolidar toda la evidencia de manera ordenada y comprensible.

6.7.2 Análisis de correlación entre las fuentes de información

Estas pruebas son herramientas para explorar las relaciones entre dos o más variables hidrológicas, son de vital importancia especialmente en este estudio, ya que permiten entender

si existe una asociación entre ambas variables, por otra parte, determinan si los datos cambian de manera lineal o no lineal.

En esta investigación las pruebas de correlación son uno de los análisis que causa mayor relevancia, dado que compara las variables de precipitación con los datos de los sensores de la NASA y de las estaciones del IDEAM, por otra parte, determina si existe una asociación entre sí.

Las pruebas de correlación se expresan en términos que van desde el -1 al 1, donde si el resultado es **-1** quiere decir que tiene una correlación negativa perfecta, **1** correlación positiva perfecta y **0** no hay correlación. Según Forero y Rodríguez (2016), el coeficiente de correlación proporciona información sobre la relación lineal existente entre dos variables cualesquiera y propone en su estudio las metodologías de Combinación Lineal Ponderada (CLP) y Distancia Inversa Ponderada (IDW), por otra parte Robledo (2020) presenta su metodología mediante el análisis de Coeficiente de correlación de Pearson (r).

En esta investigación se propone el análisis de coeficiente de correlación mediante la metodología de Pearson, Spearman y Kendall, para las variables de precipitación entre los datos de los sensores satelitales y de estaciones climatológicas.

Para el análisis de la estadística descriptiva y validación de datos, se desarrolló un código en lenguaje de programación Python, lo que sirvió como base para, posteriormente, aplicar los procedimientos de correlación entre las fuentes IDEAM y NASA. ejecutado mediante el entorno (Jupyter Notebook). En este proceso se utilizaron principalmente las bibliotecas pandasMcKinney (2010), (numpy y scipy.stats) Virtanen et al. (2020), junto con herramientas de visualización como (matplotlib.pyplot y seaborn). Estas permitieron automatizar el procesamiento, análisis y representación gráfica de la información de precipitación proveniente tanto de las estaciones del IDEAM como de los sensores satelitales (NASA – POWER y

Giovanni). Para el modelamiento estadístico y predictivo se incorporaron (statsmodels), especialmente para análisis de series temporales, y sklearn, usada para validación cruzada y evaluación de modelos Pedregosa (2011).. Estas bibliotecas son ampliamente reconocidas por su validez y flexibilidad en el tratamiento de datos científicos.

El análisis estadístico incluyó el cálculo de medidas de tendencia central (media, mediana, moda), de dispersión (desviación estándar, varianza y rango), y de forma (asimetría y curtosis). Esto permitió identificar patrones y comportamientos generales de la serie de datos.

También se aplicaron pruebas de normalidad sobre los conjuntos de datos, con el fin de evaluar si seguían una distribución normal, lo que es importante para validar el uso de determinados procedimientos estadísticos. Estas pruebas se interpretaron a partir de los valores p y gráficamente mediante histogramas y curvas de densidad Virtanen et al. (2020).

En cuanto a la evaluación de la similitud entre las fuentes de información (IDEAM y NASA), se diseñó un bloque específico de análisis en Python, utilizando las funciones de correlación del módulo (scipy.stats), lo que permitió comparar de manera directa las series de precipitación. Posteriormente se calcularon los coeficientes de Pearson (r), Spearman (ρ) y Kendall (τ), debido a que permiten evaluar tanto relaciones lineales como no lineales, dependiendo de la naturaleza de los datos Virtanen et al. (2020).

Estos indicadores se complementaron con gráficos de dispersión, líneas de ajuste y etiquetas del coeficiente correspondiente para visualizar el grado de conexión. La implementación de este procedimiento permitió evaluar la concordancia entre las mediciones satelitales y las observadas por estaciones del IDEAM, consolidando así una base comparativa robusta para posteriores validaciones. de Pearson (r), Spearman (ρ) y Kendall (τ), utilizando funciones del módulo (scipy.stats) Virtanen et al. (2020).

Estas métricas permitieron establecer el grado de conexión entre las variables de precipitación satelitales y las observadas por estaciones del IDEAM. Adicionalmente se integraron gráficos de dispersión, líneas de regresión y anotaciones del coeficiente de correlación para facilitar la visualización del grado de concordancia. Dicho procedimiento permitió evaluar tanto relaciones lineales como no lineales, fortaleciendo la interpretación sobre el comportamiento conjunto de las dos fuentes de datos.

Además, se implementó un procedimiento de validación cruzada utilizando KFold de la biblioteca (`sklearn.model_selection`), con cinco particiones ($K=5$). En cada iteración se calcularon los errores cuadráticos medios (MSE) y su raíz (RMSE), con el fin de estimar la precisión del ajuste entre los datos satelitales y los observados Pedregosa (2011).

Como parte del análisis de series temporales, se utilizó el modelo SARIMAX (`statsmodels.tsa.statespace.sarimax`) para evaluar la capacidad predictiva de los datos satelitales. Previamente, se aplicó la prueba de Dickey-Fuller Aumentada (ADF), mediante `adfuller`, para verificar la estacionariedad de las series. El modelo se configuró con parámetros estacionales (1,1,1,12) y su resumen se exportó automáticamente a un archivo .txt para su interpretación posterior. McKinney (2010).

6.7.3 Modelo de predicción y estudio de tendencias.

6.7.3.1 Validación cruzada

Los análisis de validación cruzada evalúan la precisión de los modelos hidrológicos o de predicción y compara los resultados esperados con los datos observados, busca determinar que tan bien se comportan los datos en un modelo estadístico, este proceso se debe realizar para ambas fuentes de información IDEAM y NASA.

La validación cruzada se utiliza para evaluar la precisión de los datos de precipitación de la NASA en comparación con los del IDEAM. Se divide el conjunto de datos en subconjuntos utilizando Validación cruzada estratificada (Stratified K-Fold Cross-Validation) y se calcula el error cuadrático medio (RMSE) entre los datos de la NASA y los del IDEAM. Un RMSE bajo indica una alta precisión de los datos de la NASA.

6.7.3.2 Pruebas de significancia estadística:

Estas pruebas permiten establecer si los resultados obtenidos a partir del análisis de series de datos son estadísticamente representativos o podrían deberse al azar. Aunque no se aplicaron pruebas paramétricas clásicas como *t de Student* o ANOVA, se realizaron análisis de correlación (Pearson, Spearman y Kendall) y validación cruzada con RMSE como herramientas complementarias para evaluar la consistencia entre los datos de precipitación del IDEAM y los sensores satelitales de la NASA. Estas técnicas permitieron cuantificar el nivel de concordancia y verificar si las diferencias observadas entre las fuentes son significativas desde un punto de vista práctico y estadístico..

6.7.3.3 Análisis de incertidumbre

Evalúa la incertidumbre asociada con los datos y los resultados obtenidos, es un proceso que se usa para cuantificar, entender y gestionar las incertidumbres, las cuales se refieren a la variabilidad que puede existir en los resultados obtenidos, puede ser por diferentes factores que no dependen del dato en sí, sino la forma en como fue tomado, esta incertidumbre permite identificar que tan reales o cercanos a la realidad están la totalidad de los datos.

Permite establecer si las muestras pueden ser veraces o cuentan con anomalías en sus resultados, en el desarrollo de esta investigación son de vital importancia dado que se deben comparar de manera individual ente ambas fuentes de información y definir qué tan reales pueden llegar a ser los datos.

6.7.3.4 Validación y comparación.

En esta investigación, la validación y comparación de los datos de precipitación se ha centrado en el análisis de la concordancia entre las series observadas por estaciones del IDEAM y los registros estimados por sensores satelitales de la NASA. Para ello, se emplearon métodos estadísticos como la correlación y la validación cruzada, los cuales permitieron establecer el grado de similitud entre ambas fuentes.

Si bien no se aplicaron análisis de probabilidad clásicos para eventos extremos, como los modelos de Gumbel o Log-Pearson, el enfoque se mantuvo en el análisis comparativo de la información diaria y mensual, evaluando la consistencia de los datos y su comportamiento general. Esta comparación permitió determinar la viabilidad de utilizar información satelital como complemento o alternativa en zonas donde no se cuenta con estaciones meteorológicas.

7 Resultados y discusión

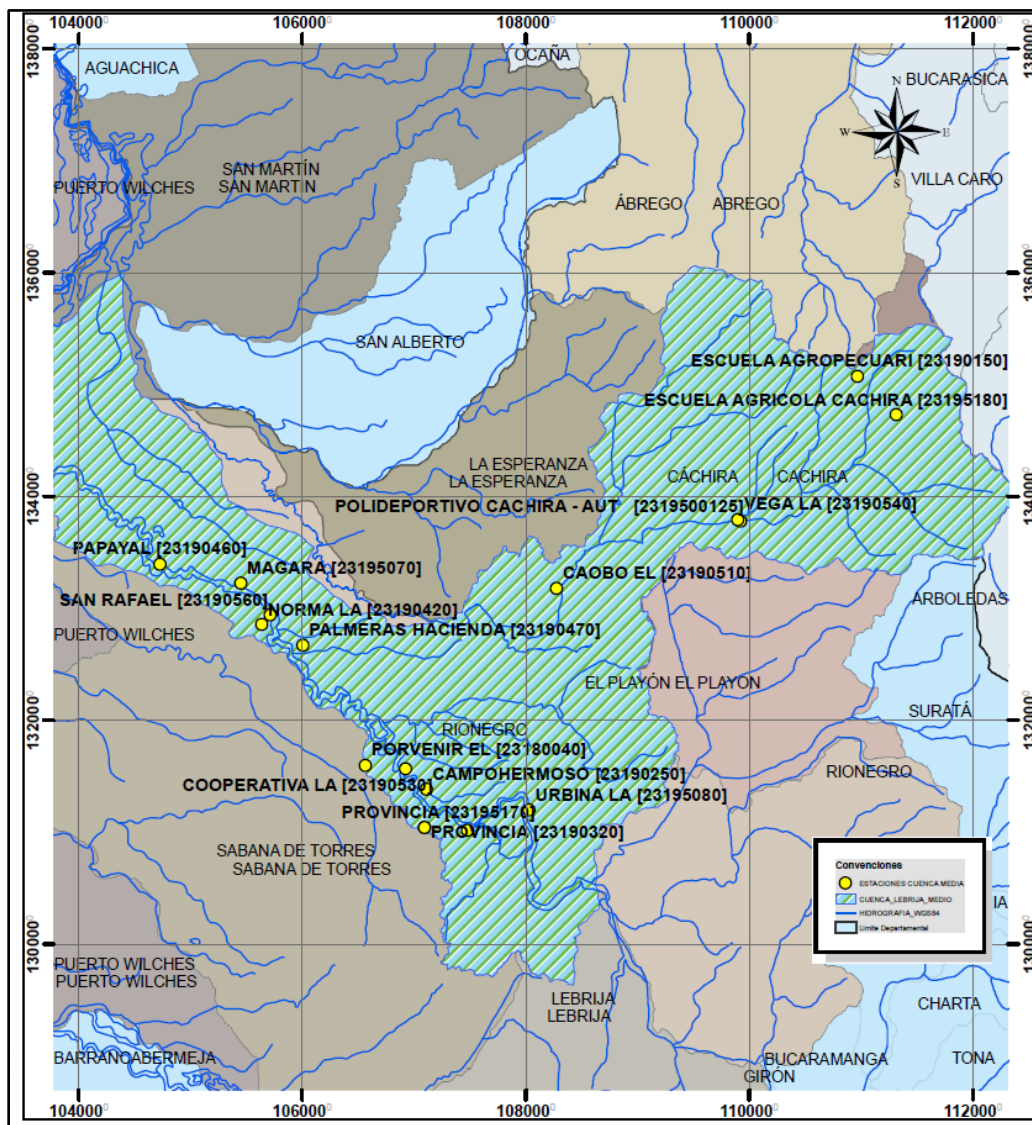
7.1 Resultado de Identificación de estaciones climatológicas.

Se recopilaron los datos climatológicos de las estaciones pluviométricas ubicadas en el área de estudio, proporcionados por el IDEAM en el portal de datos abiertos DHIME (Datos Hidrológicos y meteorológicos). Se hizo la selección de todas las estaciones que están localizadas en el área de la cuenca hidrográfica, en total se descargó la información diaria a nivel multianual de precipitaciones de 25 estaciones, se verificó el nivel de aprobación de la estación, según recomendaciones del IDEAM ideal 1200, lo que significa que los datos tienen un nivel de aceptación bueno, se realizó la verificación de la disponibilidad de información que tuvieran al menos 20 años consecutivos con tomas de datos, además que la cantidad de datos faltantes no fuera muy extensa y que cuente con estaciones cercanas en caso de ser requerido para complementar datos faltantes.

En la (Figura 8) se muestra la localización de las 25 estaciones climatológicas seleccionadas inicialmente con la variable de precipitación para la investigación en el área de la cuenca hidrográfica.

Figura 8

Localización de las estaciones climatológicas de Cuenca medía río Lebrija.



Nota. Fuente elaboración propia.

Una vez definidas las estaciones climatológicas a implementar distribuidas espacialmente en el área de la cuenca hidrográfica, se identificaron 25 estaciones que se localizaron en el área de la cuenca (ver Figura 8) de las cuales tan solo 6 lograron cumplir con una buena disponibilidad los datos, además de una confiabilidad de los mismos (ver Figura 9), las cuales se utilizaron en los análisis realizados en la presente investigación.

Tabla 1*Información descargada de las estaciones existentes en la cuenca*

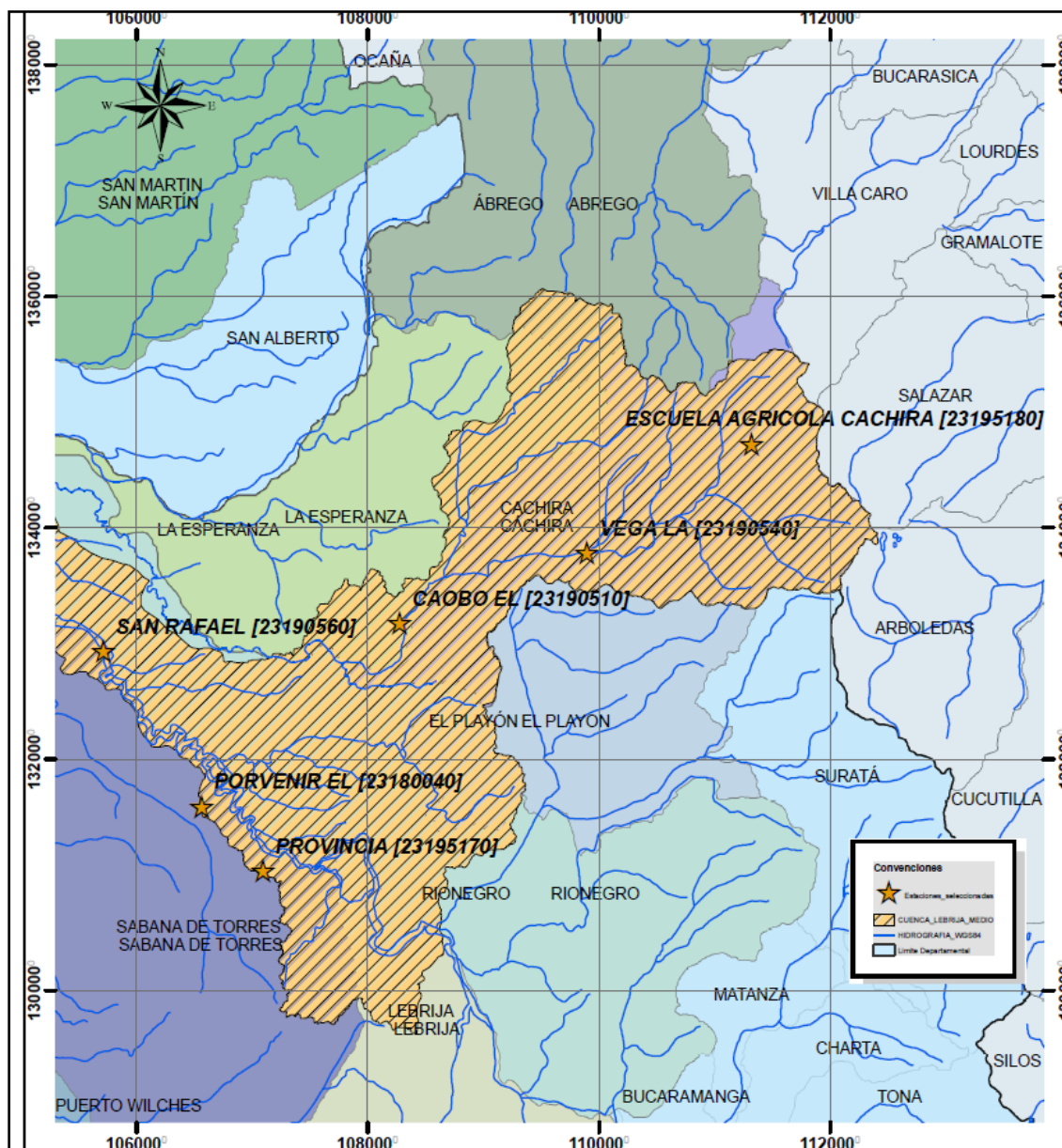
ID No.	Código Estación	Nombre Estación	Latitud	Longitud	Altitud	Municipio	Fecha inicio	Fecha fin
1	23180040	PORVENIR EL [23180040]	7.45277778	-73.48277778	110	Sabana Torre	1/01/2003 0:00	31/03/2023 0:00
2	23180050	ESTABLO EL [23180050]	7.54138889	-73.59805556	98	Sabana Torre	1/01/2003 0:00	30/04/2023 0:00
3	23180070	SABANA DE TORRES [23180070]	7.39	-73.48944444	144	Sabana Torre	1/01/2003 0:00	26/06/2023 0:00
4	23180080	ELOY VALENZUELA [23180080]	7.48944444	-73.67833333	90	Sabana Torre	1/01/2003 0:00	31/05/2023 0:00
5	23185010	VILLA LEIVA [23185010]	7.45611111	-73.53722222	328	Sabana Torre	1/01/2003 0:00	26/06/2023 0:00
6	23190210	MAGARA [23190210]	7.59333333	-73.68138889	89	Sabana Torre	1/01/2003 0:00	31/05/2023 0:00
7	23190320	PROVINCIA [23190320]	7.4025	-73.435	109	Sabana Torre	1/01/2003 0:00	30/04/2023 0:00
8	23190560	SAN RAFAEL [23190560]	7.57444444	-73.55972222	84	Sabana Torre	1/01/2003 0:00	26/06/2023 0:00
9	24060080	PAYOA 5 [24060080]	7.27027778	-73.49083333	161	Sabana Torre	1/01/2003 0:00	30/03/2023 0:00
10	23190350	LLANO DE PALMAS [23190350]	7.24013889	-73.19536111	778	Rionegro	1/01/2003 0:00	26/06/2023 0:00
11	23190360	PORTACHUELO [23190360]	7.32805556	-73.165	800	Rionegro	1/01/2003 0:00	30/04/2023 0:00
12	23190460	PAPAYAL [23190460]	7.61555556	-73.64888889	100	Rionegro	1/01/2003 0:00	30/04/2013 0:00
13	23190260	LAGUNA LA [23190260]	7.07972222	-73.21361111	150	Lebrija	1/01/2003 0:00	28/02/2023 0:00
14	23190380	PALMAS [23190380]	7.21108333	-73.21788889	855	Lebrija	1/01/2003 0:00	30/04/2023 0:00
15	23190440	NARANJO EL [23190440]	7.205	-73.29972222	825	Lebrija	1/01/2003 0:00	26/06/2023 0:00
16	23195502	AEROPUERTO PALONEGRO AUT [23195502]	7.12147222	-73.18452778	1189	Lebrija	1/01/2003 0:00	26/06/2023 0:00
17	23190140	PLAYON EL [23190140]	7.46472222	-73.20138889	500	El Playon	1/01/2003 0:00	26/06/2023 0:00
18	23190600	PANTANO EL [23190600]	6.9975	-73.23027778	1280	Girón	1/01/2003 0:00	4/12/2023 0:00
19	23190540	VEGA LA [23190540]	7.65083333	-73.18055556	710	Cachira	1/01/2003 0:00	30/09/2023 0:00
20	23195180	ESCUELA AGRÍCOLA CACHIRA [23195180]	7.73527778	-73.05166667	1882	Cachira	1/01/2003 0:00	4/12/2023 0:00
21	24060070	PARROQUIA LA [24060070]	7.07638889	-73.32777778	267	Girón	1/01/2003 0:00	28/02/2011 0:00
22	23195110	LLANO GRANDE - AUT [23195110]	7.02555556	-73.16722222	777	Girón	1/01/2003 0:00	4/12/2023 0:00
23	23190280	PALO GORDO [23190280]	6.9675	-73.13305556	950	Girón	1/01/2003 0:00	31/12/2022 0:00
24	24060060	PUENTE LA PAZ [24060060]	7.10861111	-73.41944444	180	Girón	1/01/2003 0:00	30/09/2010 0:00
25	24065030	TRIGUEROS HACIENDA [24065030]	7.07666667	-73.35805556	400	Girón	1/01/2003 0:00	31/08/2010 0:00

Nota. Fuente elaboración propia.

Conforme se mencionó anteriormente de las 25 estaciones climatológicas seleccionadas solo 6 lograron cumplir con los parámetros de aceptación mínimos para realizar un estudio hidrológico de manera adecuada, tanto en tiempo de los datos registrados, como en la calidad de las muestras, así como para el proceso de validación con los datos obtenidos de los sensores de la NASA, las estaciones con las cuales se ha desarrollado la presente investigación fueron **PORVENIR EL [23180040]**, **ESCUELA AGRÍCOLA CACHIRA [23195180]**, **CAOBO EL [23190510]**, **PROVINCIA [23190320]**, **SAN RAFAEL [23190560]** y **VEGA LA [23190540]**, su información se puede ver en Tabla 1 y su localización se puede observar en la Tabla 2.

Figura 9

Localización de las estaciones climatológicas de Cuenca media río Lebrija.



Nota. Fuente elaboración propia.

Tabla 2*Información Seleccionada para el estudio de las estaciones existentes en la cuenca*

ID No	Código Estación	Nombre Estación	Latitud	Longitud	Altitud	Municipio	Fecha inicio	Fecha fin
1	23180040	PORVENIR EL [23180040]	7.45277778	-73.48277778	110	Sabana Torre	1/01/2003 0:00	31/03/2023 0:00
2	23190510	CAOBO EL [23190510]	7.59555556	-73.3275	300	La Esperanza	1/01/2003 0:00	30/04/2023 0:00
7	23190320	PROVINCIA [23190320]	7.4025	-73.435	109	Sabana Torre	1/01/2003 0:00	30/04/2023 0:00
8	23190560	SAN RAFAEL [23190560]	7.57444444	-73.55972222	84	Sabana Torre	1/01/2003 0:00	26/06/2023 0:00
19	23190540	VEGA LA [23190540]	7.65083333	-73.18055556	710	Cachira	1/01/2003 0:00	30/09/2023 0:00
20	23195180	ESCUELA AGRÍCOLA CACHIRA [23195180]	7.73527778	-73.05166667	1882	Cachira	1/01/2003 0:00	4/12/2023 0:00

Nota. Fuente elaboración propia.

7.2 Datos de sensores satelitales implementados.

Se hizo la descarga de la información de parámetros de precipitación a nivel diario multianual de los portales abiertos de bases de datos climatológicos de POWER Data Access Viewer del departamento Prediction of Worldwide Energy Resources el cual se basa en la información del sensor (MERRA-2) Modern-Era Retrospective analysis for Research and Applications, versión 2, y Giovanni de la EARTH DATA quien a su vez se basa en el sensor Tropical Rainfall Measuring Mission (TRMM) y Global Precipitation Measurement (GPM). ambas fuentes de la NASA para las mismas coordenadas geográficas de las 25 estaciones identificadas en la (ver Tabla 1 localizadas en el área de la cuenca, sin embargo en los análisis de información disponible se implementaron únicamente en las seis estaciones mencionadas en la (ver Tabla 2)

7.3 Comparación de series temporales y validación estadística

7.3.1 Evaluación del Sensor GIOVANNI frente a Estaciones de Referencia (datos diarios)

Se realizó un análisis comparativo entre los datos de precipitación diaria obtenidos del sensor satelital Giovanni y los registros tomados por las estaciones del IDEAM, utilizando

herramientas estadísticas y gráficas. El propósito fue identificar el grado de correspondencia entre ambas fuentes y evaluar qué tan representativos pueden ser los datos satelitales en el contexto local de la cuenca del río Lebrija y en general en el uso para estudio, hidrológicos, de cambio climático y ambientales.

Para ello se calcularon los coeficientes de correlación de Pearson, Spearman y Kendall. Los resultados mostraron diferencias notables entre estaciones. La Escuela Agrícola Cachirá presentó los mejores niveles de ajuste entre las series, mientras que estaciones como Porvenir y Provincia reflejaron correlaciones considerablemente bajas ver (**Anexo A, Anexo I, Anexo Q, Anexo Y, Anexo GG y Anexo OO**).

Los valores de correlación para Escuela Agrícola Cachirá alcanzaron un Pearson de 0.3648, un Spearman de 0.4884 y un Kendall de 0.3807, lo cual aunque no representa una relación fuerte, sí sugiere una coherencia moderada entre los datos satelitales y los observados. En contraste, estaciones como Caobo o Porvenir no superaron el 0.1 en los coeficientes, reflejando que en esos puntos la correspondencia entre fuentes es débil (ver **Tabla 5, Tabla 7, Tabla 9, Tabla 11, Tabla 13 y Tabla 15**)

Por otra parte, también se calcularon los indicadores de error como RMSE y MAE, así como el coeficiente de determinación R^2 . Estos valores permitieron cuantificar las diferencias entre las fuentes. Escuela Agrícola Cachirá nuevamente presentó los mejores resultados con un RMSE de 4.26 mm, un MAE de 2.80 mm y un R^2 de 0.1268 (ver **Anexo I, Anexo J, Anexo K y Anexo L**). En contraste, la estación Porvenir mostró un RMSE superior a 16 mm y un R^2 prácticamente nulo.

Lo anterior evidencia que el sensor Giovanni, aunque útil en zonas con pocas estaciones, presenta limitaciones importantes en términos de precisión a nivel diario de

análisis. En varias estaciones las diferencias son amplias, lo cual podría estar relacionado con la forma en que estos sensores captan y procesan la información diaria de lluvia

En cuanto al comportamiento estacional, se observaron patrones similares en ambas fuentes, especialmente el régimen bimodal de lluvias típico de la región. No obstante, en los datos de Giovanni se evidenció una menor intensidad en los picos de precipitación, lo que sugiere una posible subestimación de los eventos más intensos. Esta diferencia puede observarse en los registros diarios de las estaciones, tal como se presenta en los anexos gráficos correspondientes (ver **Anexo C, Anexo L, Anexo W, Anexo EE, Anexo MM y Anexo UU**).

Al revisar las gráficas de dispersión, se observó que la mayoría de los registros se agrupan en los valores bajos de precipitación, lo que indica que ambas fuentes coinciden con mayor frecuencia en días sin lluvia o con lluvias ligeras. Sin embargo, a medida que aumentan los valores, las diferencias también se hacen más evidentes. En varios casos, Giovanni no registra eventos de lluvia intensa que sí aparecen en los datos del IDEAM, lo cual sugiere una subestimación sistemática de los eventos fuertes. Esta situación es visible en los anexos de dispersión por estación (ver **Anexo B, Anexo J, Anexo R, Anexo Z, Anexo HH y Anexo PP**).

En el análisis de distribución, se evidenció que los datos del IDEAM presentan una mayor asimetría positiva, con una frecuencia más alta de valores bajos y algunos eventos extremos bien marcados. En cambio, los datos de Giovanni tienden a estar más concentrados, con curvas más planas y menos dispersión, lo que indica una posible homogeneización en la estimación de la precipitación. Esta diferencia puede afectar el análisis en zonas donde los valores extremos tienen mayor relevancia, como en estudios de eventos intensos o escurridos. (ver **Tabla 3**)

Entre las estaciones con mayor coherencia entre fuentes se destacan Escuela Agrícola Cachirá y La Vega. En estas estaciones, tanto los valores como la forma de las distribuciones fueron más similares. En cambio, en estaciones como Porvenir o Provincia, las diferencias se mantienen en todos los aspectos analizados

Este comportamiento podría explicarse por condiciones locales que influyen en cómo se capta la lluvia en cada estación. Factores como la altitud, la cobertura del suelo, la topografía y la ubicación del punto de observación pueden afectar la precisión de los sensores satelitales. Aunque Giovanni ofrece una cobertura espacial continua, su precisión no es uniforme en todas las zonas, lo que hace necesario validar sus datos considerando las particularidades del entorno de cada estación Virtanen et al. (2012).

Es importante resaltar que los datos satelitales tienen un valor significativo para zonas donde no existen estaciones operativas, ya que permiten contar con una estimación base. Sin embargo, para estudios más detallados o calibraciones de modelos hidrológicos, es recomendable apoyarse en estaciones locales para mejorar la precisión. Esta necesidad de validación y ajuste con datos de campo ha sido destacada en revisiones recientes sobre productos satelitales, que advierten sobre sus limitaciones en zonas complejas y la importancia de su calibración local Qiaohong et al. (2017).

Con base en lo anterior, se concluye que el sensor Giovanni puede ser una herramienta útil para complementar estudios en zonas poco instrumentadas, pero que requiere una validación previa en campo. La estacionalidad general sí es bien representada, pero los valores absolutos y los eventos extremos necesitan ser analizados con precaución.

La información presentada en la Tabla 3 resume los principales indicadores obtenidos en la comparación entre los datos diarios del sensor Giovanni y los registros de las estaciones del IDEAM. Como se ha mencionado previamente, la estación Escuela Agrícola Cachirá mostró

el mejor ajuste, con coeficientes de correlación más altos y los menores valores de error. En contraste, las estaciones Porvenir y Provincia reflejaron los desempeños más bajos, con correlaciones débiles y errores elevados, lo que refuerza las observaciones hechas en los gráficos de dispersión y distribución.

Tabla 3

Indicadores estadísticos de comparación entre datos diarios del sensor Giovanni y estaciones IDEAM.

Sensor	Estación	Pearson	Spearman	Kendall	RMSE	MAE	R2
GIOVANNI	CAOBO	0,136454	0,250385	0,194354	10,08494	6,563611	0,009272
GIOVANNI	ESCUELA AGRÍCOLA CACHIRA	0,364863	0,488403	0,380705	4,258785	2,80012	0,126843
GIOVANNI	PORVENIR	0,084817	0,183593	0,149568	16,01912	10,57541	0,001975
GIOVANNI	PROVINCIA	0,079944	0,178061	0,144839	15,06697	9,610866	0,010064
GIOVANNI	SAN RAFAEL	0,106415	0,260133	0,211091	13,37838	8,353286	0,007321
GIOVANNI	LA VEGA	0,249971	0,34803	0,280619	7,508669	4,888689	0,044913

Fuente: Elaboración propia con base en los resultados del análisis estadístico entre los datos diarios de precipitación del sensor Giovanni y las estaciones IDEAM.

7.3.2 Comparación entre PowerData y datos observados del IDEAM (datos diarios)

Se realizó una comparación entre los datos diarios de precipitación de PowerData y los registros de las estaciones IDEAM, aplicando los mismos criterios que en el caso del sensor Giovanni. Para ello se calcularon los coeficientes de correlación (Pearson, Spearman y Kendall), los errores RMSE y MAE, y el coeficiente de determinación R^2 , buscando identificar qué tan bien se ajusta esta fuente a los datos de campo en cada estación.

En la Tabla 4 se presentan los principales indicadores para cada estación. La Escuela Agrícola Cachirá vuelve a destacar como la de mejor desempeño, con valores aceptables de correlación y los errores más bajos (RMSE de 4,13 mm y MAE de 2,79 mm). En contraste, estaciones como Porvenir y Provincia siguen mostrando una baja correspondencia con PowerData, con correlaciones por debajo de 0,20 y errores superiores a los 14 mm.

En cuanto a los patrones estacionales, PowerData también refleja el régimen bimodal típico de la región, aunque con una menor amplitud en los picos. Esta característica se puede observar en los gráficos de estacionalidad incluidos en los (**Anexo E, Anexo M, Anexo U, Anexo CC, Anexo KK y Anexo RR**), correspondientes a las estaciones Caobo, Escuela Agrícola Cachirá, Porvenir, Provincia, San Rafael y La Vega, respectivamente, donde se nota cierta suavización de los valores máximos.

Respecto a la dispersión, se encontró que PowerData coincide con mayor frecuencia con los días de lluvia ligera, pero subestima varios eventos de lluvia intensa. Esto se refleja en las gráficas por estación (ver **Anexo F, Anexo N, Anexo V, Anexo DD, Anexo KK y Anexo SS** para PowerData), donde se nota cómo aumentan las diferencias en los días con valores altos.

En el análisis de distribución, los histogramas y las curvas de densidad muestran que PowerData tiende a concentrar los datos en valores medios, con menor dispersión y menos asimetría que los datos del IDEAM. Esta característica puede ser útil en análisis generales, pero limita su uso en la evaluación de eventos extremos (ver **Anexo G, Anexo N, Anexo V, Anexo DD, Anexo LL y Anexo TT**).

De manera general, se observó que PowerData tiene un comportamiento más regular que Giovanni, con menor variabilidad entre estaciones. Sin embargo, los niveles de correlación siguen siendo bajos y los errores relativamente altos, por lo cual sigue siendo necesario validar su uso dependiendo del tipo de aplicación.

Tabla 4
Indicadores estadísticos de comparación entre datos diarios de PowerData y estaciones IDEAM.

Sensor	Estación	Pearson	Spearman	Kendall	RMSE	MAE	R2
POWERDATA	CAOBO	0,189501	0,31408	0,23445	10,12422	6,627361	0,058612
POWERDATA	ESCUELA AGRÍCOLA CACHIRA	0,236098	0,329524	0,242674	4,126706	2,790591	0,03018
POWERDATA	PORVENIR	0,132551	0,19978	0,154249	15,6408	10,77134	0,004829

Sensor	Estación	Pearson	Spearman	Kendall	RMSE	MAE	R2
POWERDATA	PROVINCIA	0,187706	0,248477	0,191854	14,97936	9,386818	0,038502
POWERDATA	SAN RAFAEL	0,20247	0,301804	0,230499	12,97296	7,93969	0,015939
POWERDATA	LA VEGA	0,185364	0,249247	0,191758	7,702653	5,000518	0,031526

Nota. Fuente elaboración propia.

7.3.3 Análisis mensual de precipitación: IDEAM vs Giovanni

En este apartado se analiza el comportamiento mensual de la precipitación estimada por el sensor Giovanni, en comparación con los datos de las estaciones del IDEAM. Para ello se aplicó una validación cruzada por estación, utilizando el promedio mensual multianual de cada fuente, y se incluyeron modelos de ajuste tipo SARIMA para observar la respuesta ante la estacionalidad y la tendencia general.

A diferencia del análisis diario, a nivel mensual se obtuvo una mejor correspondencia entre las fuentes. Las correlaciones de Pearson muestran valores altos en todas las estaciones, con registros que superan el 0,86 en Porvenir y alcanzan hasta 0,97 en San Rafael. Esto indica que el sensor sí logra capturar con bastante precisión el patrón estacional mensual de la región.

En la estación Escuela Agrícola Cachirá, por ejemplo, se observó una correlación de 0,91 entre IDEAM y Giovanni. Las gráficas de barras y líneas (ver **Anexo YY**) muestran que, aunque existe una subestimación leve en algunos meses, el comportamiento general es coherente, y el modelo SARIMA ajustado (**Anexo KKK** y **Anexo LLL**) reflejó una serie estable en ambos casos.

La estación Caobo presentó una correlación aún mayor, con un valor de 0,92. La forma del ciclo bimodal se conserva claramente, y las diferencias mensuales son menores en comparación con las observadas en los datos diarios (ver **Anexo WW**). En este caso, el ajuste del modelo SARIMA mostró una menor dispersión del error, lo que da más confianza al uso de estos datos para análisis estacionales.

En el caso de Porvenir, Provincia y La Vega, los valores de correlación se mantuvieron entre 0,86 y 0,94, lo cual indica un buen desempeño general del sensor Giovanni en escalas agregadas. Aunque se evidencian pequeñas diferencias en los picos de lluvia mensual, el comportamiento general sigue la misma lógica de distribución

La estación San Rafael arrojó la mejor correlación de todas las estaciones, con un Pearson de 0,97. Esto refuerza la idea de que el producto Giovanni es capaz de representar adecuadamente el ciclo mensual de lluvias, al menos en cuanto a forma y momento de ocurrencia. Esta estación, además, mostró un ajuste bastante limpio en el modelo SARIMA (ver **Anexo QQQ** y **Anexo RRR**).

En general, se puede decir que el sensor Giovanni ofrece una buena aproximación a los datos mensuales, y que su uso puede considerarse confiable cuando se requieren estudios de mediano plazo o caracterización estacional. Sin embargo, siempre es recomendable acompañar estos productos con validación local y comparaciones con registros observados, especialmente en regiones de alta variabilidad como la cuenca del río Lebrija.

Es importante señalar que los resultados presentados en este apartado se basan en datos mensuales agregados para la cuenca media del río Lebrija, y están condicionados por las características locales de cada estación. Las diferencias observadas entre IDEAM y Giovanni fueron cuantificadas mediante estadísticos como la media, desviación estándar y percentiles, además de la correlación de Pearson. Por ejemplo, en la estación San Rafael, el promedio mensual registrado por IDEAM fue de 2.786 mm mientras que Giovanni (NASA) estimó 3.552 mm, con una diferencia absoluta promedio de 766 mm. Esta sobrestimación es consistente en varios meses, aunque la correlación de Pearson fue de 0,97, indicando un buen seguimiento de la forma del régimen estacional. Estos resultados reflejan que, aunque el sensor puede capturar la estacionalidad mensual, los valores absolutos deben interpretarse con cautela y

validados localmente. Por tanto, se recomienda no generalizar los resultados sin realizar previamente una validación adicional en contextos distintos.

Como complemento a la validación cruzada y el análisis estadístico mensual, se ajustaron modelos SARIMA a las series mensuales estimadas por Giovanni para cada estación. El modelo SARIMAX(1,1,1)(1,1,1)[12] permitió representar bien la forma del comportamiento estacional. Estaciones como Escuela Agrícola Cachirá y La Vega presentaron valores bajos de AIC (38840) y varianzas residuales estables (σ^2 alrededor de 45), lo que refleja un ajuste razonable del modelo. En otras como San Rafael y Porvenir la varianza fue más alta (superior a 120), pero aun así el modelo logró captar la tendencia general de las lluvias. Esto confirma que Giovanni puede seguir bien el patrón mensual, pero es necesario tener cuidado con los valores absolutos. Aunque la forma de la serie se ajusta, los volúmenes pueden diferir bastante frente a los datos reales del IDEAM, por eso es clave validar localmente antes de usar estos datos en análisis más específicos.

7.3.4 Análisis mensual de precipitación: IDEAM vs PowerData

Se llevó a cabo el análisis mensual de PowerData frente a los registros IDEAM en cada estación, utilizando la misma metodología aplicada con Giovanni. Se trabajó con promedios mensuales multianuales, se aplicó validación cruzada y se ajustaron modelos SARIMA por estación, con el fin de evaluar cómo se comporta este producto satelital a escala mensual.

Los resultados muestran correlaciones de Pearson superiores a 0,92 en todas las estaciones, destacándose Provincia con una correlación de 0,99 y San Rafael con 0,98. Esto indica que PowerData también sigue correctamente la forma del ciclo estacional mensual en la región, con diferencias menos marcadas que las observadas en la escala diaria.

En la estación Escuela Agrícola Cachirá, se obtuvo una correlación de 0,92. Aunque se presentaron pequeñas subestimaciones en algunos meses, las gráficas comparativas muestran coherencia en la estructura mensual (ver **Anexo ZZ**). Además, el modelo SARIMA aplicado reflejó un ajuste estable (Anexo P3), lo que respalda el uso de estos datos para análisis estacionales.

En el caso de Caobo, la correlación fue de 0,96. Se observó un patrón bimodal claro, y las diferencias mensuales fueron menores respecto a lo observado con Giovanni. El comportamiento general es más uniforme y el modelo SARIMA mostró un buen ajuste con errores bajos (ver **Anexo III** y **Anexo JJJ**).

Estaciones como Porvenir y La Vega mostraron correlaciones de 0,95 y 0,95 respectivamente. Aunque en ciertos meses se presentaron diferencias en volumen, la coincidencia en los picos de lluvia mensual fue evidente (ver Anexos P6 a P9).

Provincia fue la estación con mejor resultado general, con una correlación de 0,99. El sensor logró reproducir con precisión el patrón mensual observado, lo que se refleja también en el comportamiento del modelo SARIMA ajustado (**Anexo PPP**).

Por último, San Rafael, con una correlación de 0,98, confirmó el buen rendimiento de PowerData para este tipo de análisis. Las diferencias absolutas fueron mínimas y el patrón mensual se mantuvo estable en toda la serie observada (ver **Anexo RRR**).

En resumen, PowerData presenta un comportamiento sólido para análisis mensuales. Aunque también tiende a suavizar valores extremos, su ajuste al patrón estacional es alto y consistente. Estos resultados respaldan su uso como fuente complementaria en zonas sin cobertura instrumental, siempre que se acompañe de validación local.

Este análisis mensual se realizó con datos en milímetros, agregados por estación y validados mediante estadísticos como media, desviación estándar y correlación de Pearson. Se identificaron diferencias importantes entre PowerData e IDEAM, tanto en magnitud como en dispersión. Por ejemplo, en la estación Porvenir, el promedio mensual estimado por IDEAM fue de 4.804 mm, mientras que PowerData registró 2.628 mm, con una diferencia media de -2.176 mm y una desviación de 1.113 mm. A pesar de esa diferencia en volumen, la correlación fue alta (0,95), lo que indica que el sensor sigue correctamente la forma del régimen estacional. Por lo tanto, aunque PowerData puede ser una fuente útil en zonas sin estaciones, siempre se debe validar con datos locales para evitar errores en estudios sensibles al volumen absoluto.

En paralelo con el análisis aplicado a Giovanni, se ajustaron modelos SARIMA a las series mensuales estimadas por PowerData en cada estación. El modelo utilizado fue SARIMAX(1,1,1)(1,1,1)[12], que logró representar de forma estable el patrón estacional mensual. En las estaciones Caobo, Cachirá y La Vega, el modelo presentó AIC de 46669 y varianzas residuales cercanas a 25, lo que indica un buen ajuste. Por el contrario, en Porvenir, Provincia y San Rafael, la varianza fue mayor (σ^2 alrededor de 45), aunque el modelo siguió siendo útil para representar la tendencia general. Esto sugiere que PowerData, al igual que Giovanni, puede ajustarse bien a nivel mensual, pero que las diferencias en volumen requieren revisión caso por caso. Por eso es importante no asumir que el buen comportamiento del modelo garantiza precisión en los valores absolutos.

7.4 Recomendaciones según los resultados obtenidos

En futuros trabajos que incluyan sensores satelitales, se recomienda realizar una validación local previa con estaciones disponibles, ya que los valores absolutos pueden variar considerablemente dependiendo del lugar. Es importante complementar los análisis con ajustes

estadísticos o modelos de corrección, como regresiones lineales o calibraciones por cuenca, especialmente si los datos se usarán para modelaciones hidrológicas o balances hídricos.

Cuando no se cuente con información de estaciones, los sensores satelitales pueden servir como insumo base, pero siempre deben ser revisados frente a comportamientos extremos y estacionales. Se sugiere usar escalas mensuales o superiores para aprovechar mejor la continuidad y cobertura espacial de los sensores, ya que a nivel diario se observan mayores errores.

En regiones con alta variabilidad climática, es clave usar más de una fuente de información y comparar los resultados entre sensores y estaciones, de modo que se pueda evaluar la coherencia y ajustar según sea necesario. Para trabajos similares, se recomienda documentar detalladamente los procesos de depuración, validación cruzada y ajuste de modelos, ya que estos pasos son clave para interpretar correctamente los resultados y evitar errores de análisis. Finalmente, es conveniente mantener actualizados los registros y explorar nuevas plataformas o productos satelitales que puedan ofrecer mejores resoluciones temporales y espaciales según los objetivos de cada investigación.

8 Conclusiones

Conforme al objeto general y después de realizar el análisis comparativo de los datos de precipitación para la cuenca media del río Lebrija, se logró establecer que los sensores satelitales Giovanni y PowerData pueden seguir el comportamiento general del régimen estacional de lluvias, aunque con limitaciones importantes en cuanto a la precisión de los valores diarios y mensuales frente a los datos del IDEAM. Por otra parte, PowerData mostró un comportamiento más parejo entre estaciones, con errores más estables, mientras que Giovanni tuvo mejor ajuste en estaciones puntuales. Esto sugiere que la elección de la fuente depende del área y del tipo de análisis que se quiera hacer.

Según lo establecido en el objetivo general, en el análisis diario de los datos, las correlaciones fueron bajas en la mayoría de las estaciones, especialmente cuando se presentaron lluvias intensas. Esto evidencia que los sensores no captan bien los extremos, y por tanto, su uso a este nivel debe hacerse con precaución y siempre con validación local. A nivel mensual, el comportamiento de los sensores mejora, se encontraron correlaciones de Pearson superiores a 0,90, lo que indica que logran seguir la forma del ciclo de lluvias. Sin embargo, las diferencias en los valores acumulados persisten, por lo cual también se recomienda validarlos localmente.

Conforme al primer objetivo específico, se determinó que para el análisis de climatológico de la variable de precipitación, se establecieron 25 estaciones localizadas en la cuenca media del río Lebrija, de las cuales se seleccionaron 6 que cumplen con los parámetros de análisis mínimos, el tiempo años usado va desde el 1 de enero de 2004 hasta el 31 de diciembre de 2024, y se analizaron comparándolos estadísticamente entre los datos Giovanni vs IDEAM y PowerData vs IDEAM.

Por otra parte también después de analizar el primer objetivo específico y una vez validada la información disponible de los sensores satelitales, se observó que para los datos de Giovanni se encontró que se tienen registros para años hasta el 2019, con lo cual se considera mejor fuente de información satelital PowerData.

Según el segundo objetivo específico y después de realizar el análisis se puede concluir que los resultados obtenidos respaldan el uso combinado de datos, utilizando tanto estaciones del IDEAM como sensores satelitales. De esta manera se aprovecha la continuidad espacial de Giovanni y la precisión puntual del IDEAM. Además se logró determinar que los modelos SARIMA aplicados a las series mensuales confirmaron que es posible ajustar la forma del régimen estacional, aunque las varianzas residuales cambian según la estación. En estaciones como Escuela Agrícola Cachirá y La Vega el ajuste fue mejor, mientras que en Porvenir o San Rafael se encontraron mayores diferencias.

Adicionalmente a la conclusión anterior para el segundo objetivo específico, se comprobó que, en zonas sin estaciones, los sensores pueden servir como insumo base. No obstante, su uso debe ir acompañado de análisis complementarios que ayuden a corregir o ajustar los datos, especialmente si se va a usar esta información en estudios hidrológicos o ambientales. En resumen, los datos del IDEAM siguen siendo la fuente más confiable siempre que estén disponibles. Los sensores satelitales son una alternativa válida, pero su uso requiere revisión caso por caso, y no deben reemplazar de forma automática los registros de estaciones.

Después de realizar el análisis y respondiendo al tercer objetivo específico se recomienda que en futuros trabajos se recomienda evaluar ajustes locales de los datos satelitales mediante regresiones u otros métodos de corrección, de modo que se pueda reducir el error sin perder la ventaja de cobertura amplia que ofrece el satélite.

Adicionalmente este tipo de comparaciones es clave para estudios en cuencas como la del río Lebrija, donde la instrumentación es limitada y donde se hace necesario buscar nuevas fuentes de información confiables para el desarrollo de proyectos ambientales, hidrológicos y de gestión del recurso hídrico.

9 Referencias bibliográficas

- Aguirre N. M. (2007). *Manual para el manejo sustentable de cuencas hidrográficas*.
- Bernal R., Gradstein S. R., & Celis M. (2016). *Catálogo de plantas y líquenes de Colombia volumen i*.
- Cárdenas G. J. M., Sáenz H. G. D., Díaz A. D. A., Pirazan V.K. N., & Barrera M. S. F. (2024). Análisis de partes interesadas para la misión satelital FACSAT-3. *Revista UIS Ingenierías*, 23(3). <https://doi.org/10.18273/revuin.v23n3-2024006>
- Conde-Carrión L. E. (2018). *Estimación de precipitación por imagen satelital en la provincia de Huamanga, 2014 - 2016*. Universidad Nacional de San Cristóbal de Huamanga.
- Corporación Autónoma Regional para la Defensa de la Meseta de Bucaramanga -CDMB-. (2019). *Consultoría para el ajuste (actualización) del plan de ordenación y manejo de la cuenca hidrográfica del Rio Lebrija medio (código 2319-03) de conformidad con los términos establecidos por el Fondo de adaptación dentro del convenio interadministrativo no. 021 de 2014*.
- Díaz Z. M. P., & Guevara S. J. I. (2020). *Caracterización hidrológica y de calidad del agua de la quebrada Tonchalá de la ciudad de San José de Cúcuta (Norte de Santander)*. Universidad Santo Tomás.
- Eslava R. J. A. (1993). Climatología y Diversidad Climática de Colombia. *REV. ACAD. COLOMBIA. CIENCI. VOL. XVIII, NUMMERO 71, MARZO DE 1993*, 507–538.
- Forero L. J. A., & Rodríguez G. M.:S. (2016). *Evaluación de métodos convencionales para estimación de datos faltantes de precipitación de las estaciones hidro-meteorológicas de influencia en la microcuenca El Cune en Villeta Cundinamarca*. https://ciencia.lasalle.edu.co/ing_civil
- Hernández S. R., Fernández C. C., & Baptista L. P. (2014). *Metodología de la investigación* (6th ed.). Mc Graw Hill.
- Luis Gerardo Meza Cascante, M. (n.d.). *El paradigma positivista y la concepción dialéctica del conocimiento*.
- M. Ángel, Lado J. J., V. Martínez, L. María, & R. García. (2009). *El ciclo hidrológico: experiencias prácticas para su comprensión*.
- Mckinney, W. (2010). *Data Structures for Statistical Computing in Python*.
- Mejía F., Mesa O., Poveda G., Vélez J., Hoyos C., Mantilla R., Barco J., Cuartas A., Montoya M., & Botero B. (1999). Distribución espacial y ciclos anual y semianual de la precipitación en Colombia. *REVISTA DE LA FACULTAD DE MINAS, UNIVERSIDAD NACIONAL DE COLOMBIA. DEDE MEDELLÍN.* , 6–15.

- National Aeronautics and Space Administration - NASA. (n.d.). *Giovanni - Global Precipitation Measurement*. Giovanni. Retrieved May 16, 2025, from <https://gpm.nasa.gov/data/sources/giovanni>
- National Aeronautics and Space Administration - NASA. (2021). *NASA POWER Docs*. <https://power.larc.nasa.gov/docs/services/>
- Ortega R. H. M. (2018). *Evaluación de la influencia de la incorporación de datos de precipitación satelital (TMPA-3B42R) sobre la simulación hidrológica a escala diaria en la cuenca alta del Río Guayllabamba*. <http://bibdigital.epn.edu.ec/handle/15000/20072>
- Pineda C. G., & De Piñerez N. G. (2020). *Comparación de un balance hidrológico de la cuenca del río Checua haciendo uso de métodos tradicionales (estaciones meteorológicas) y de sensores remotos (teledetección), a partir de análisis de sensibilidad*. 1–262.
- Poveda, G. (2004). La hidroclimatología de Colombia: una síntesis desde la escala inter-decadal hasta la escala diaria por ciencias de la tierra. *Escuela de Geociencias y Medio Ambiente, Facultad de Minas*, 201–222.
- Quiroga R. N. S., & Duarte C. A. (2021). *Evaluación de la incidencia del cambio climático en la cuenca Lebrija alto aplicando Hydrobid*. Unidades Tecnológicas de Santander.
- Robledo D. V. (2020). *Validación de datos satelitales de irradiancia solar de GOES-EAST sobre Medellín-Colombia*. Universidad de Antioquia.
- Rodríguez P. G. W., Sofrony E. J., Cortés G. E. D., & Rueda K. (2020). Diseño de misión, síntesis de factores operacionales y representaciones del segmento espacial, caso FACSAT y EMFF. *Ciencia y Poder Aéreo*, 15(2), 143–165. <https://doi.org/10.18667/cienciaypoderaereo.678>
- Ruiz-Ochoa, M. A., Vargas-Corredor, Y. A., Orduz-Amaya, L. P., & Torres-Corredor, J. S. (2022). Variabilidad climática en la planificación hídrica de la cuenca del río Cravo Sur (Casanare, Colombia). *Información Tecnológica*, 33(4), 117–124. <https://doi.org/10.4067/s0718-07642022000400117>
- Sánchez L. J. L. (2017). *Modelación hidrológica e hidráulica acoplada de la cuenca media y baja del río Magdalena*. Universidad Nacional de Colombia.
- Sánchez L- T. (2024). *Comparación entre Datos de programas satelitales CHIRPS, ERA 5 y del IDEAM para generar información climatológica. Caso de estudio Boyacá y Casanare*.
- Valencia G. J. N. (2020). *Balance hídrico para la planeación agrícola del arroz en el Valle del Cauca, Colombia*. Universidad Autónoma Chapingo.
- Vázquez-Ochoa, L. A., Correa-Sandoval, A., Vargas-Castilleja, R. C., Vázquez-Sauceda, M. L., & Rodríguez-Castro, J. H. (2021). Modelo hidrológico, calidad del agua y cambio climático:

soporte para la gestión hídrica de la cuenca del río Soto la Marina. *CienciaUAT*, 16, 20–41. <https://doi.org/10.29059/cienciauat.v16i1.1498>

Vega. R. E. E. (2022). *Protocolo de observaciones en las estaciones meteorológicas*. (Vol. 1).

Vilchis-Mata, I., Bâ, K. M., Franco-Plata, R., & Díaz-Delgado, C. (2015). *Modelación hidrológica con base en estimaciones de precipitación con sensores hidrometeorológicos*. VI, 45–60. <http://www.isac>.

Villena T. P. A. (2021). *Análisis temporal de los años hidrológicos de la Cuenca del Río Santa, debido a la incidencia del cambio climático*.

Weber J. F. (2020). Caracterización Hídrica de la Cuenca del Arroyo Vaquerías (Córdoba, Argentina) Mediante el Modelado Hidrológico e Información Satelital. *Revista Tecnología y Ciencia*, 39, 118–139. <https://doi.org/10.33414/rtyc.39.118-139.2020>

10 Anexos

Anexo 1: Análisis comparativo datos a nivel diario estación Caobo.

Anexo A

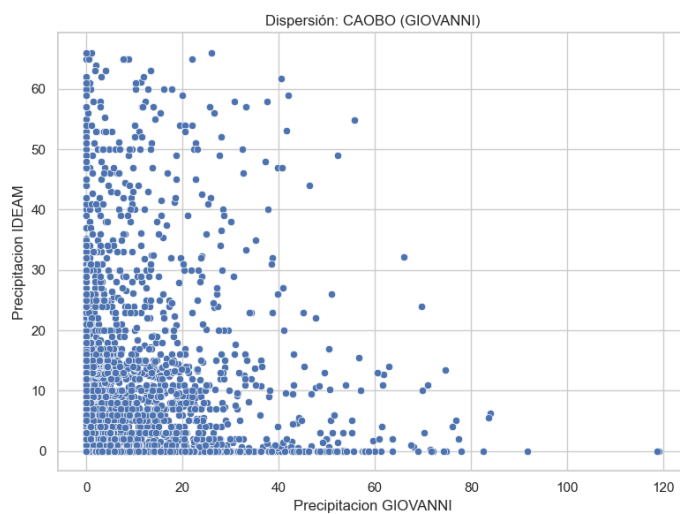
Análisis comparativo datos IDEAM vs datos Giovanni diario estación Caobo.



Nota. Fuente elaboración propia.

Anexo B

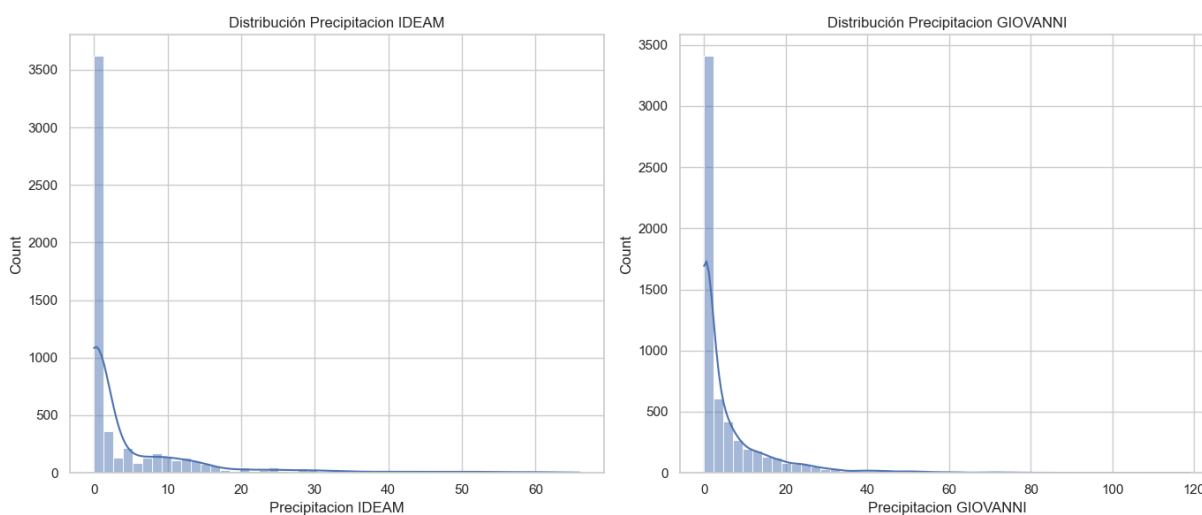
Análisis comparativo de dispersión datos IDEAM vs datos Giovanni diario estación Caobo.



Nota. Fuente elaboración propia.

Anexo C

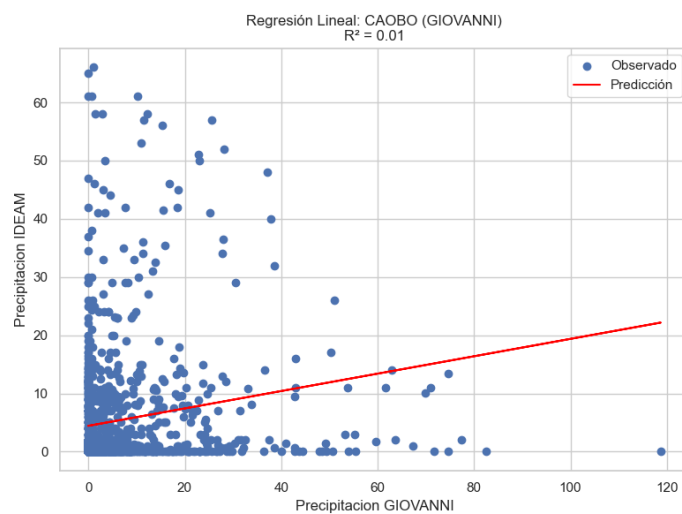
Análisis comparativo de distribución de precipitación datos IDEAM vs datos Giovanni diario estación Caobo.



Nota. Fuente elaboración propia.

Anexo D

Análisis comparativo de distribución de precipitación datos IDEAM vs datos Giovanni diario estación Caobo.



Nota. Fuente elaboración propia.

Tabla 5

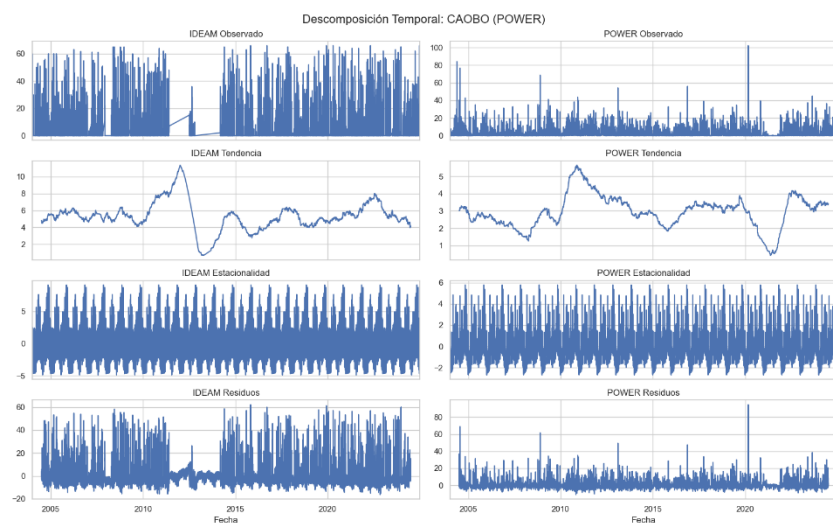
Distribuciones probabilísticas comparados datos IDEAM vs datos Giovanni diario estación Caobo.

Pearson	Spearman	Kendall	RMSE	MAE	R2
0.136454	0.250385	0.194354	10.08494	6.563611	0.009272

Nota. Fuente elaboración propia.

Anexo E

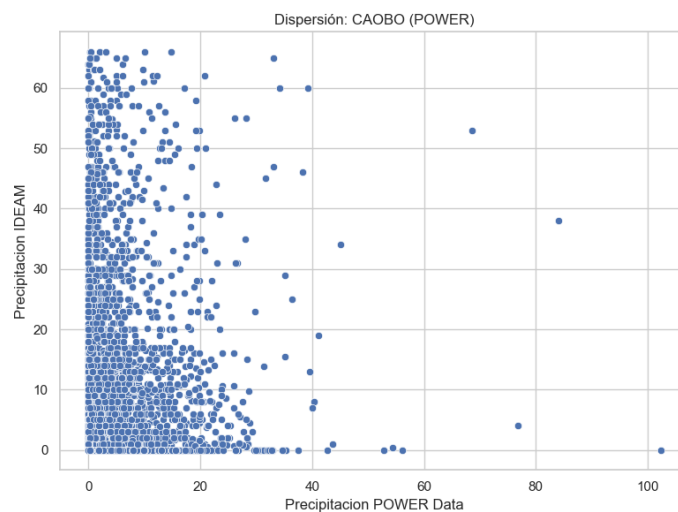
Análisis comparativo datos IDEAM vs datos Power Data diario estación Caobo.



Nota. Fuente elaboración propia.

Anexo F

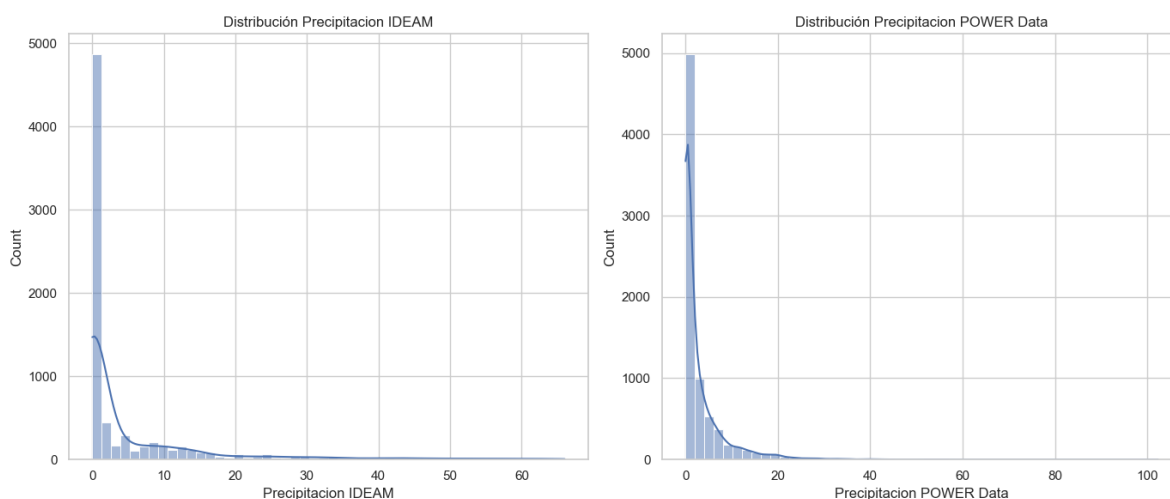
Análisis comparativo de dispersión datos IDEAM vs datos Power Data i diario estación Caobo.



Nota. Fuente elaboración propia.

Anexo G

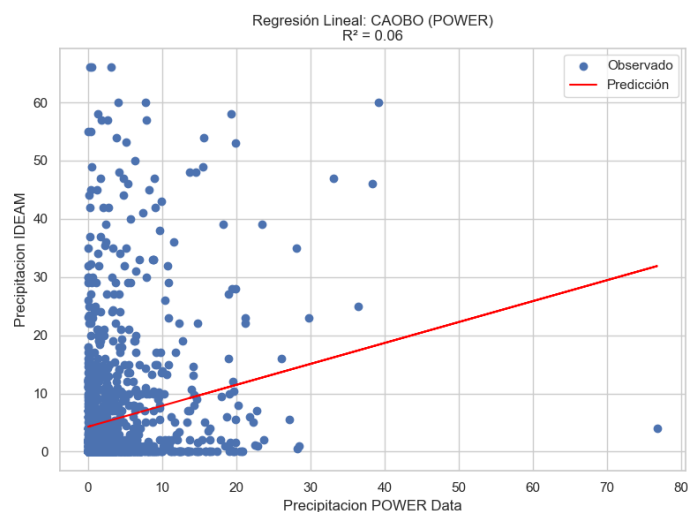
Análisis comparativo de distribución de precipitación datos IDEAM vs datos Power Data diario estación Caobo.



Nota. Fuente elaboración propia.

Anexo H

Análisis comparativo de distribución de precipitación datos IDEAM vs datos Power Data diario estación Caobo.



Nota. Fuente elaboración propia.

Tabla 6

Distribuciones probabilísticas comparados datos IDEAM vs datos Power Data diario estación Caobo.

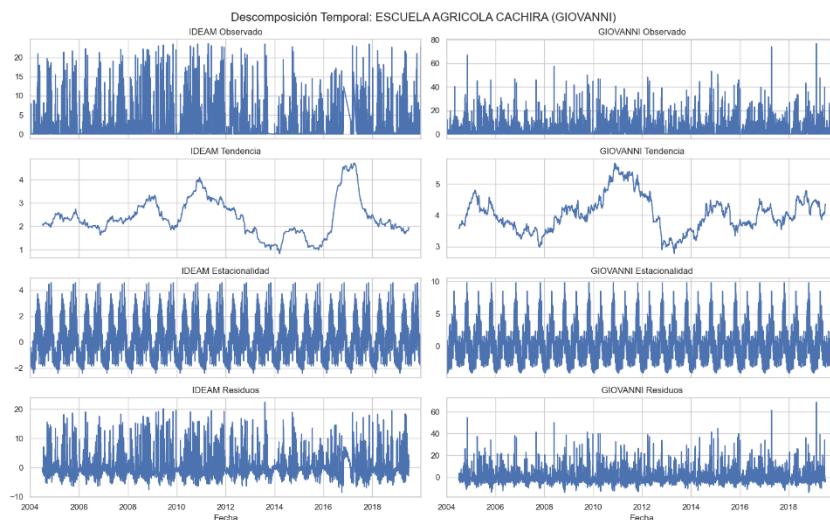
Pearson	Spearman	Kendall	RMSE	MAE	R2
0.189501	0.31408	0.23445	10.12422	6.627361	0.058612

Nota. Fuente elaboración propia.

Anexo 2: Análisis comparativo datos a nivel diario estación Escuela Agrícola Cachirá.

Anexo I

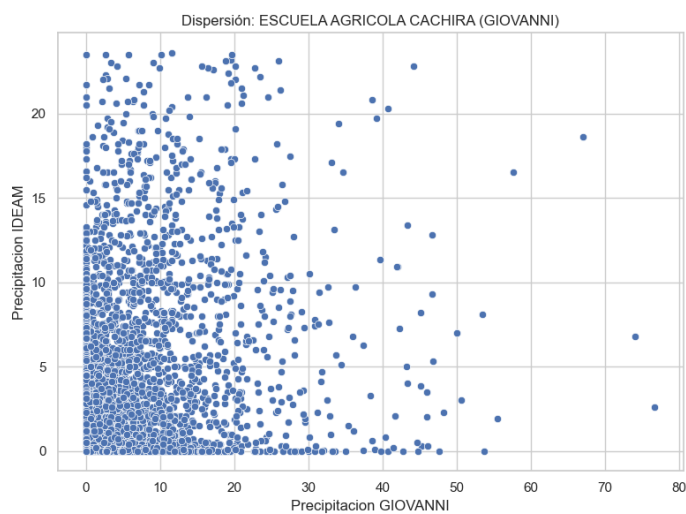
Análisis comparativo datos IDEAM vs datos Giovanni diario estación Escuela Agrícola Cachirá.



Nota. Fuente elaboración propia.

Anexo J

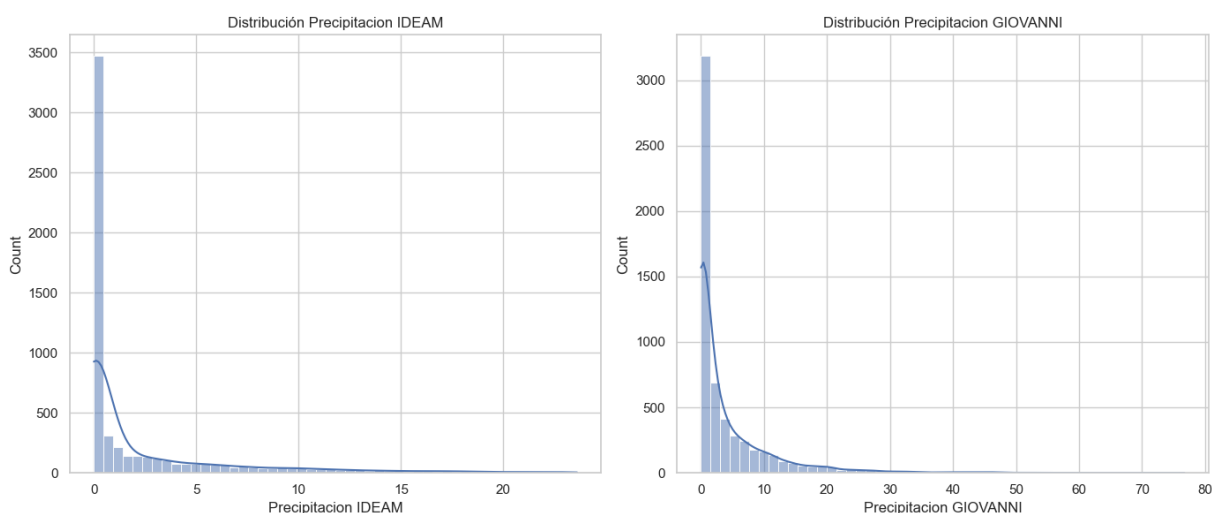
Análisis comparativo de dispersión datos IDEAM vs datos Giovanni diario estación Escuela Agrícola Cachirá..



Nota. Fuente elaboración propia.

Anexo K

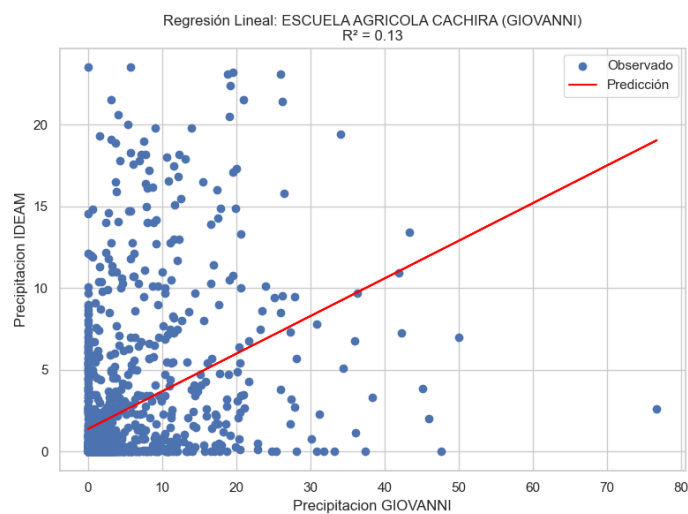
Análisis comparativo de distribución de precipitación datos IDEAM vs datos Giovanni diario estación Escuela Agrícola Cachirá.



Nota. Fuente elaboración propia.

Anexo L

Análisis comparativo de distribución de precipitación datos IDEAM vs datos Giovanni diario estación Escuela Agrícola Cachirá.



Nota. Fuente elaboración propia.

Tabla 7

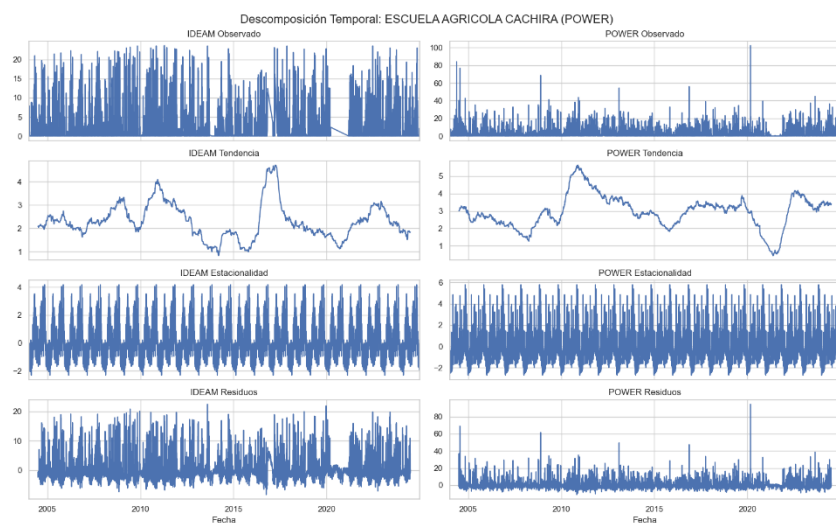
Distribuciones probabilísticas comparados datos IDEAM vs datos Giovanni diario estación Escuela Agrícola Cachirá.

Pearson	Spearman	Kendall	RMSE	MAE	R2
0.364863	0.488403	0.380705	4.258785	2.80012	0.126843

Nota. Fuente elaboración propia.

Anexo M

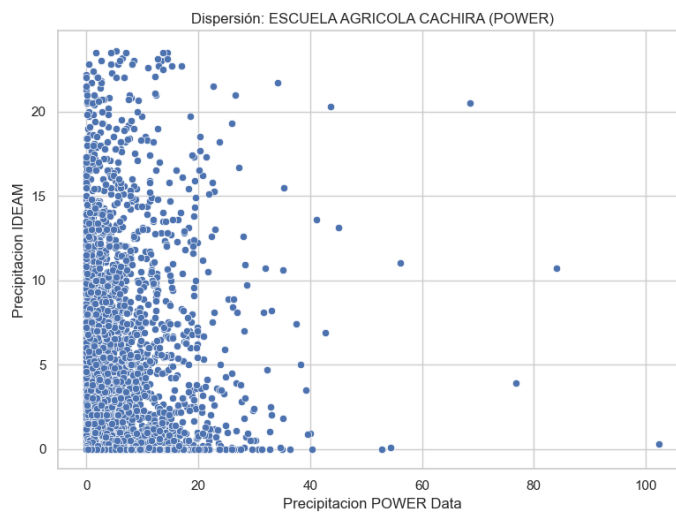
Análisis comparativo datos IDEAM vs datos Power Data diario estación Escuela Agrícola Cachirá.



Nota. Fuente elaboración propia.

Anexo N

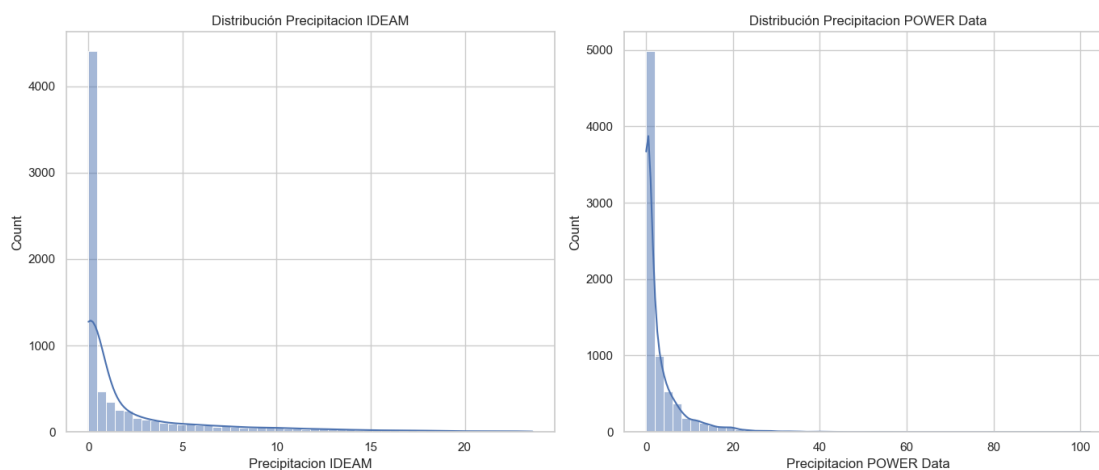
Análisis comparativo de dispersión datos IDEAM vs datos Power Data diario estación Escuela Agrícola Cachirá.



Nota. Fuente elaboración propia.

Anexo O

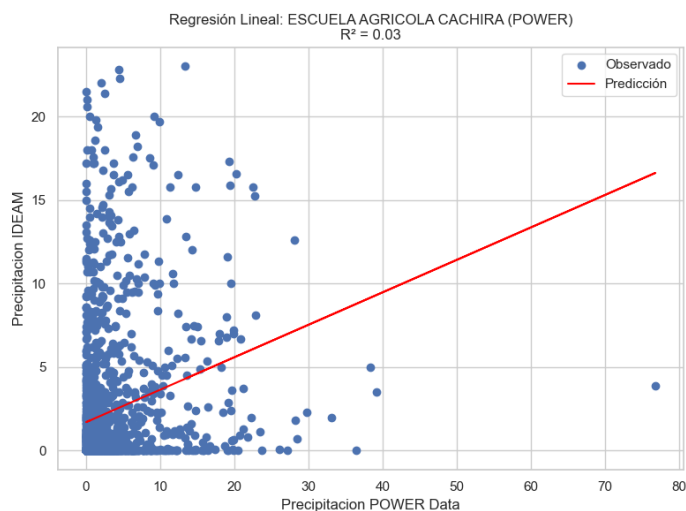
Análisis comparativo de distribución de precipitación datos IDEAM vs datos Power Data diario estación Escuela Agrícola Cachirá.



Nota. Fuente elaboración propia.

Anexo P

Análisis comparativo de distribución de precipitación datos IDEAM vs datos Power Data diario estación Escuela Agrícola Cachirá.



Nota. Fuente elaboración propia.

Tabla 8

Distribuciones probabilísticas comparados datos IDEAM vs datos Power Data diario estación Escuela Agrícola Cachirá.

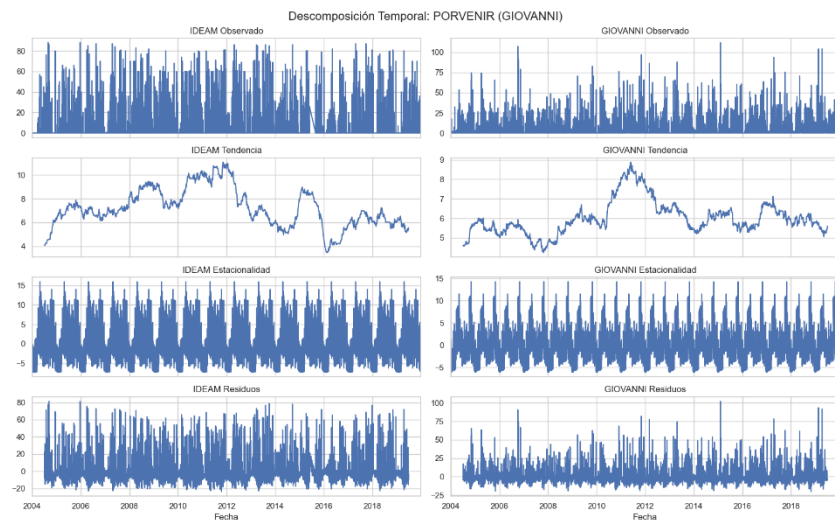
Pearson	Spearman	Kendall	RMSE	MAE	R2
0.189501	0.31408	0.23445	10.12422	6.627361	0.058612

Nota. Fuente elaboración propia.

Anexo 3: Análisis comparativo datos a nivel diario estación Porvenir.

Anexo Q

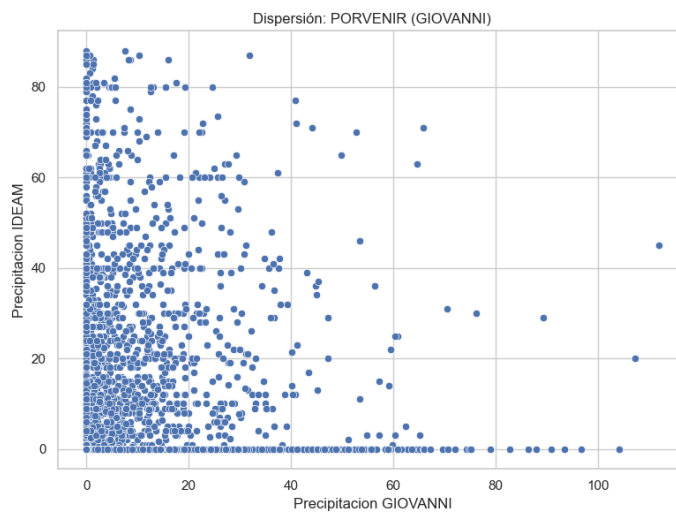
Análisis comparativo datos IDEAM vs datos Giovanni diario estación Porvenir..



Nota. Fuente elaboración propia.

Anexo R

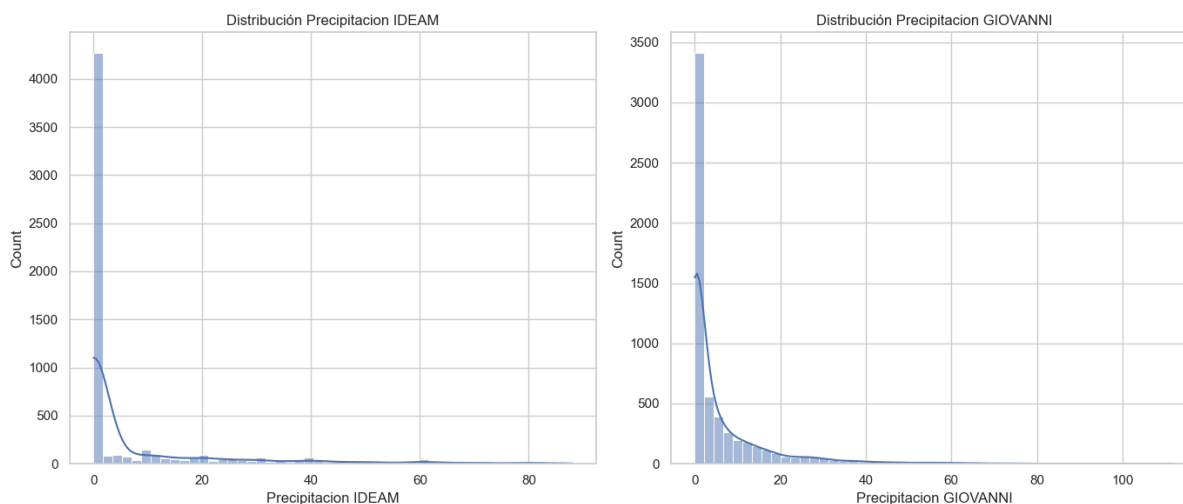
Análisis comparativo de dispersión datos IDEAM vs datos Giovanni diario estación Porvenir.



Nota. Fuente elaboración propia.

Anexo S

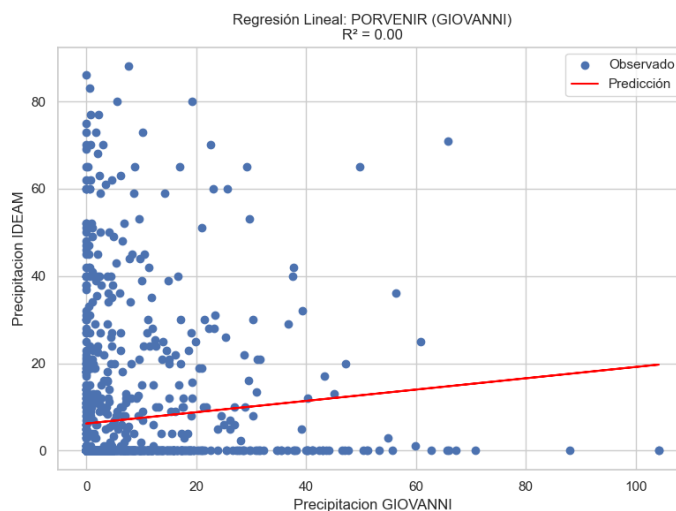
Análisis comparativo de distribución de precipitación datos IDEAM vs datos Giovanni diario estación Porvenir.



Nota. Fuente elaboración propia.

Anexo T

Análisis comparativo de distribución de precipitación datos IDEAM vs datos Giovanni diario estación Porvenir..



Nota. Fuente elaboración propia.

Tabla 9

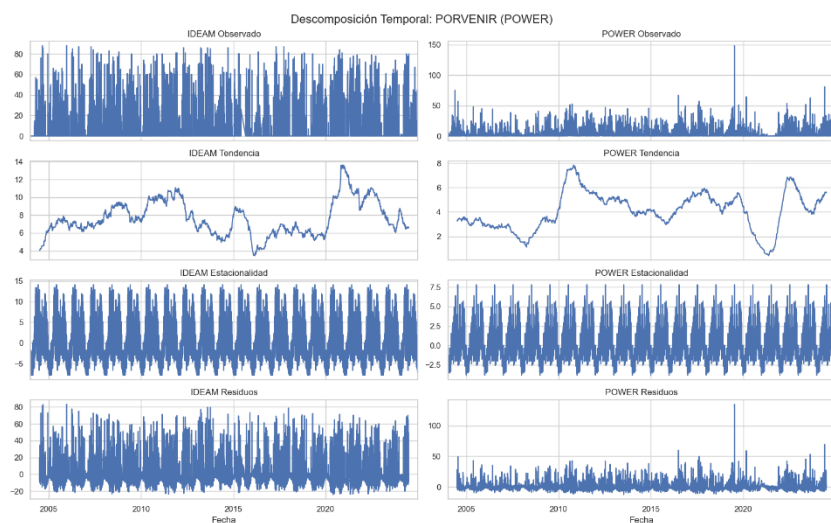
Distribuciones probabilísticas comparados datos IDEAM vs datos Giovanni diario estación Porvenir..

Pearson	Spearman	Kendall	RMSE	MAE	R2
0.084817	0.183593	0.149568	16.01912	10.57541	0.001975

Nota. Fuente elaboración propia.

Anexo U

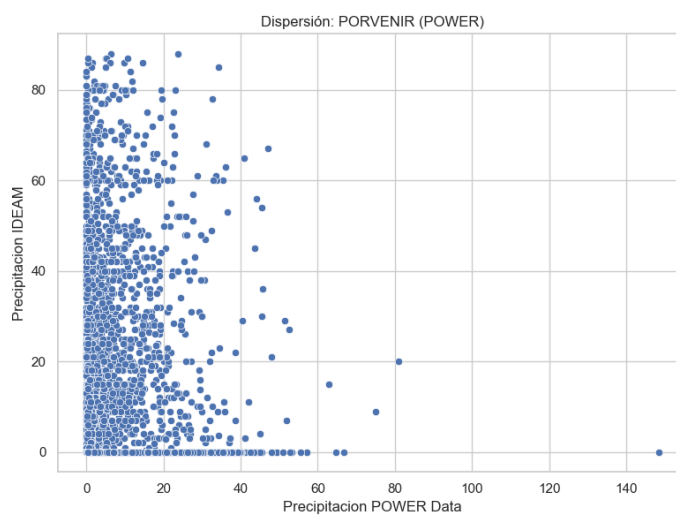
Análisis comparativo datos IDEAM vs datos Power Data diario estación Porvenir.



Nota. Fuente elaboración propia.

Anexo V

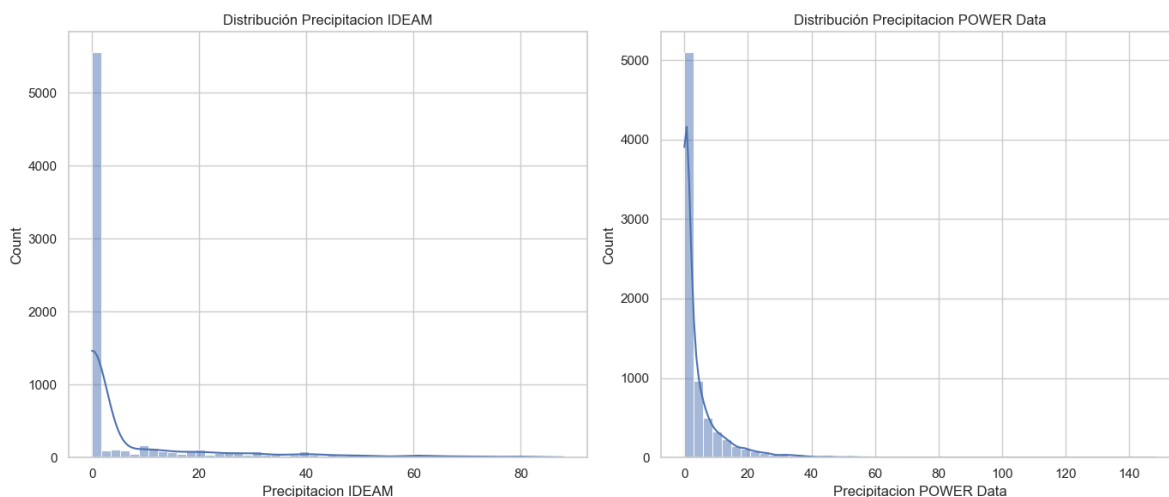
Análisis comparativo de dispersión datos IDEAM vs datos Power Data diario estación Porvenir.



Nota. Fuente elaboración propia.

Anexo W

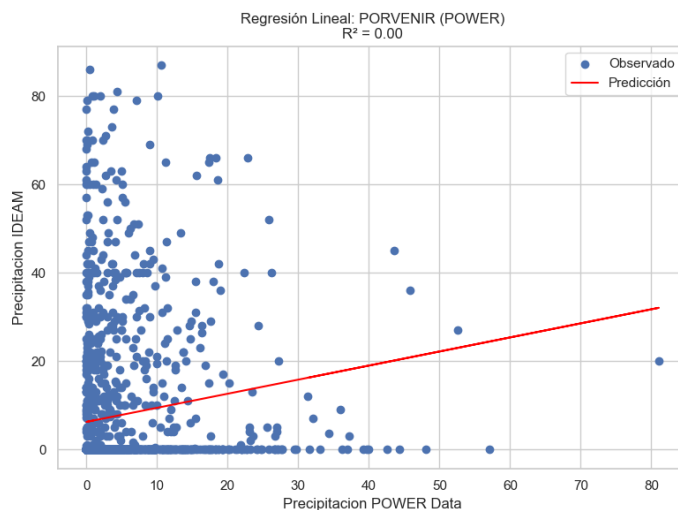
Análisis comparativo de distribución de precipitación datos IDEAM vs datos Power Data diario estación Porvenir.



Nota. Fuente elaboración propia.

Anexo X

Análisis comparativo de distribución de precipitación datos IDEAM vs datos Power Data diario estación Porvenir.



Nota. Fuente elaboración propia.

Tabla 10

Distribuciones probabilísticas comparados datos IDEAM vs datos Power Data diario estación Porvenir.

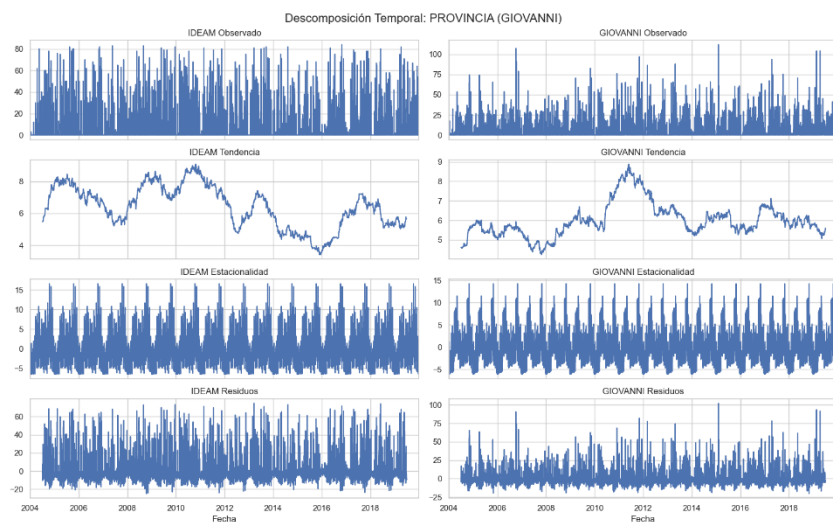
Pearson	Spearman	Kendall	RMSE	MAE	R2
0.132551	0.19978	0.154249	15.6408	10.77134	0.004829

Nota. Fuente elaboración propia.

Anexo 4: Análisis comparativo datos a nivel diario estación Provincia

Anexo Y

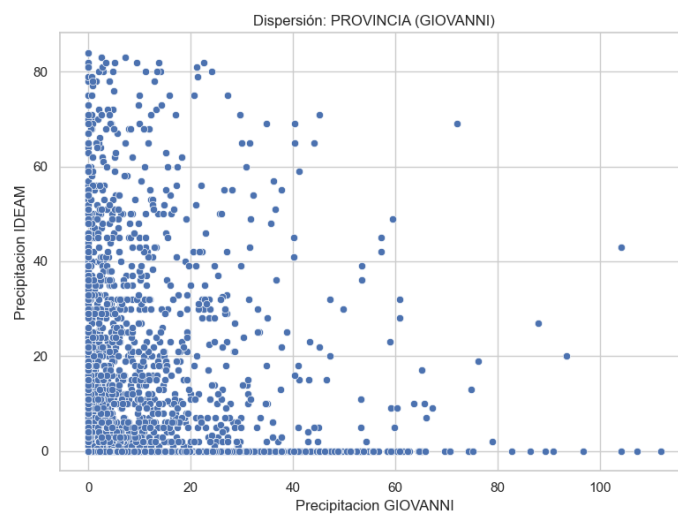
Análisis comparativo datos IDEAM vs datos Giovanni diario estación Provincia



Nota. Fuente elaboración propia.

Anexo Z

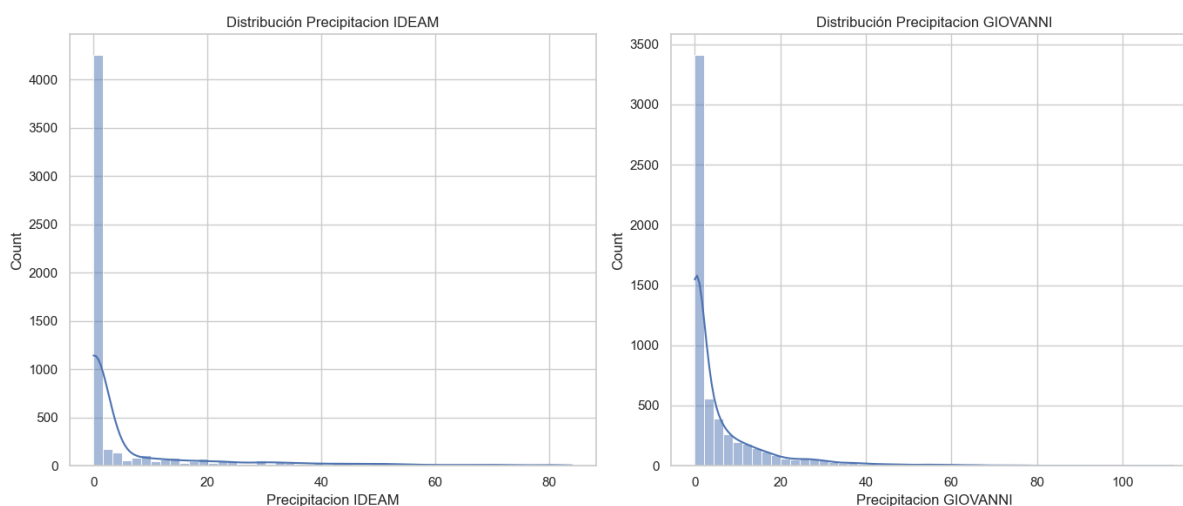
Análisis comparativo de dispersión datos IDEAM vs datos Giovanni diario estación Provincia



Nota. Fuente elaboración propia.

Anexo AA

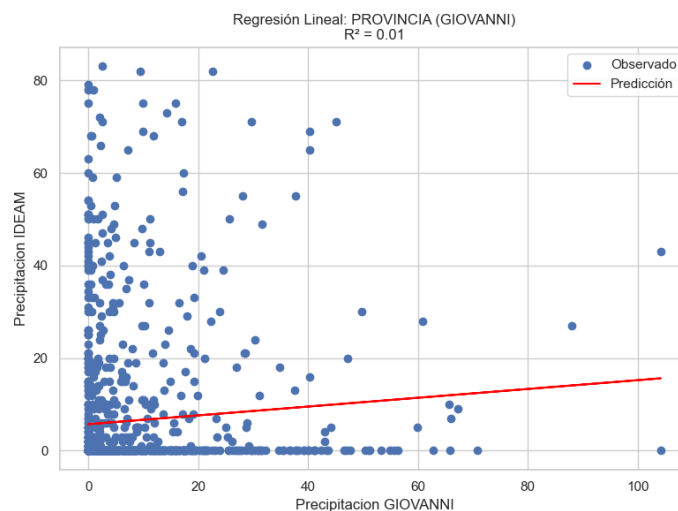
Análisis comparativo de distribución de precipitación datos IDEAM vs datos Giovanni diario estación Provincia



Nota. Fuente elaboración propia.

Anexo BB

Análisis comparativo de distribución de precipitación datos IDEAM vs datos Giovanni diario estación Provincia



Nota. Fuente elaboración propia.

Tabla 11

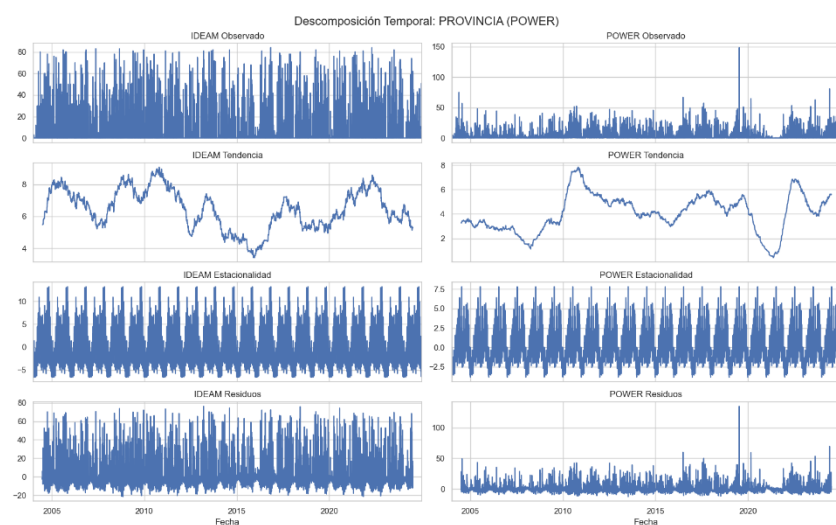
Distribuciones probabilísticas comparados datos IDEAM vs datos Giovanni diario estación Provincia

Pearson	Spearman	Kendall	RMSE	MAE	R2
0.079944	0.178061	0.144839	15.06697	9.610866	0.010064

Nota. Fuente elaboración propia.

Anexo CC

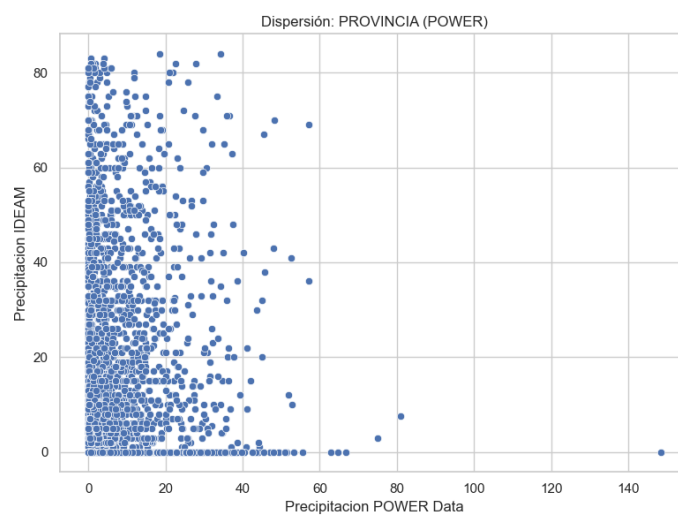
Análisis comparativo datos IDEAM vs datos Power Data diario estación Provincia



Nota. Fuente elaboración propia.

Anexo DD

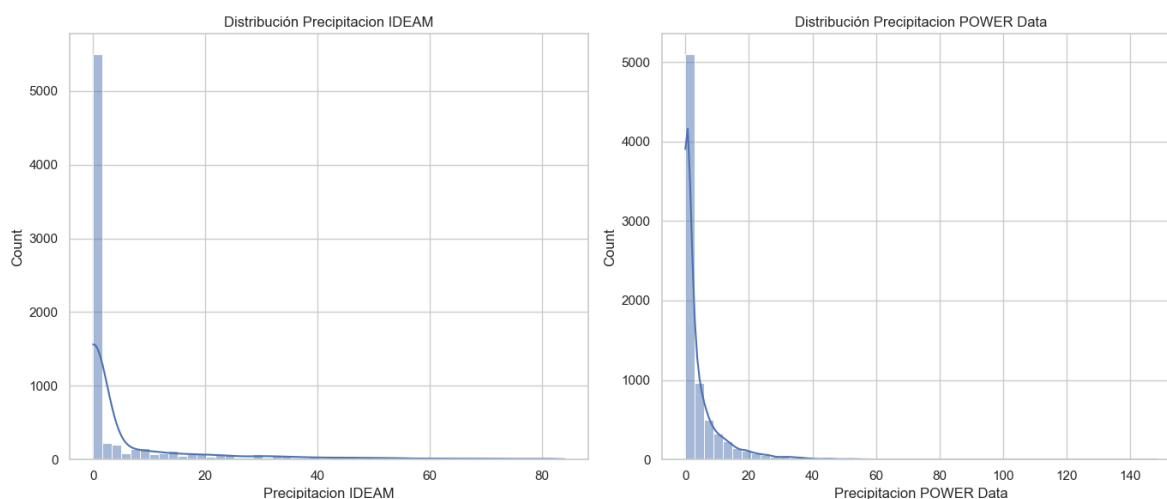
Análisis comparativo de dispersión datos IDEAM vs datos Power Data diario estación Provincia



Nota. Fuente elaboración propia.

Anexo EE

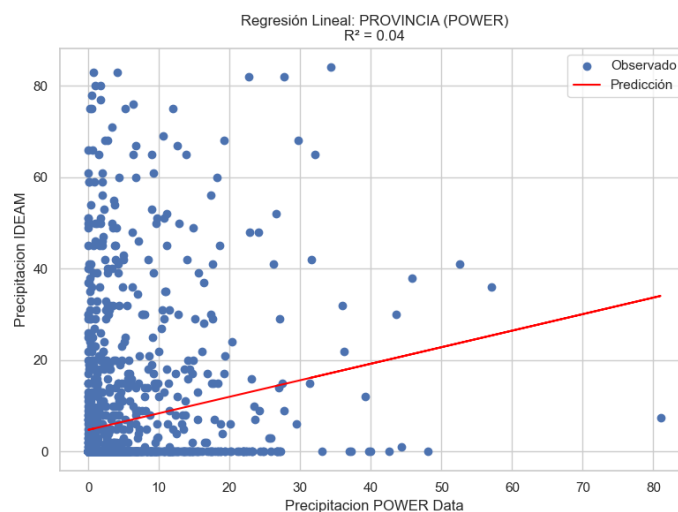
Análisis comparativo de distribución de precipitación datos IDEAM vs datos Power Data diario estación Provincia



Nota. Fuente elaboración propia.

Anexo FF

Análisis comparativo de distribución de precipitación datos IDEAM vs datos Power Data diario estación Provincia



Nota. Fuente elaboración propia.

Tabla 12

Distribuciones probabilísticas comparados datos IDEAM vs datos Power Data diario estación Provincia

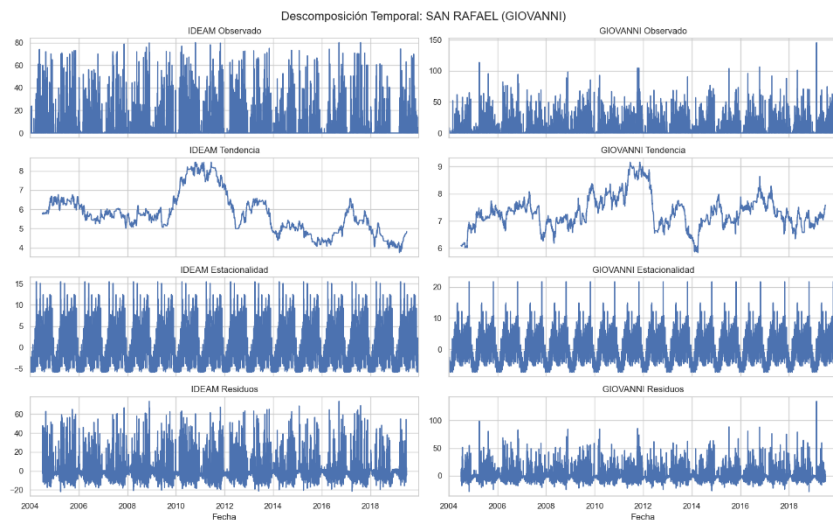
Pearson	Spearman	Kendall	RMSE	MAE	R2
0.187706	0.248477	0.191854	14.97936	9.386818	0.038502

Nota. Fuente elaboración propia.

Anexo 5: Análisis comparativo datos a nivel diario estación San Rafael.

Anexo GG

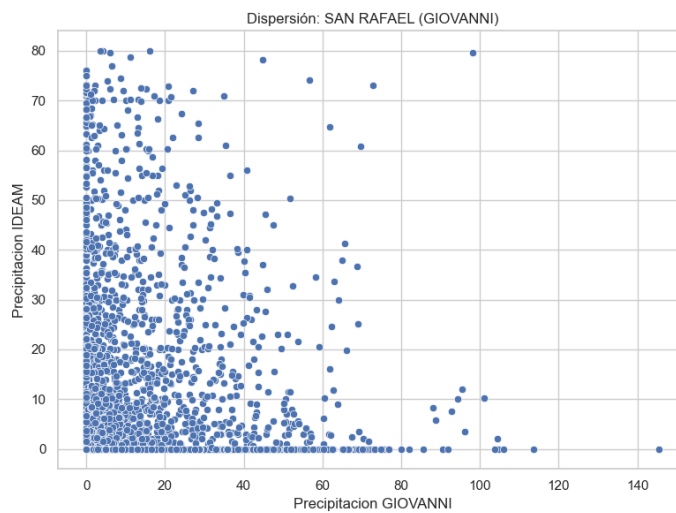
Análisis comparativo datos IDEAM vs datos Giovanni diario estación San Rafael.



Nota. Fuente elaboración propia.

Anexo HH

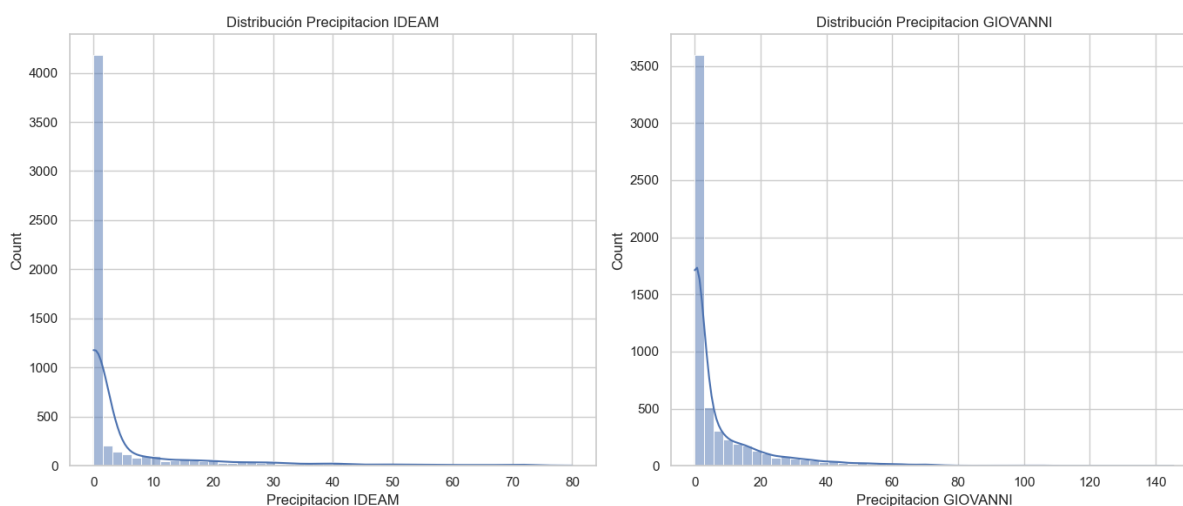
Análisis comparativo de dispersión datos IDEAM vs datos Giovanni diario estación San Rafael.



Nota. Fuente elaboración propia.

Anexo II

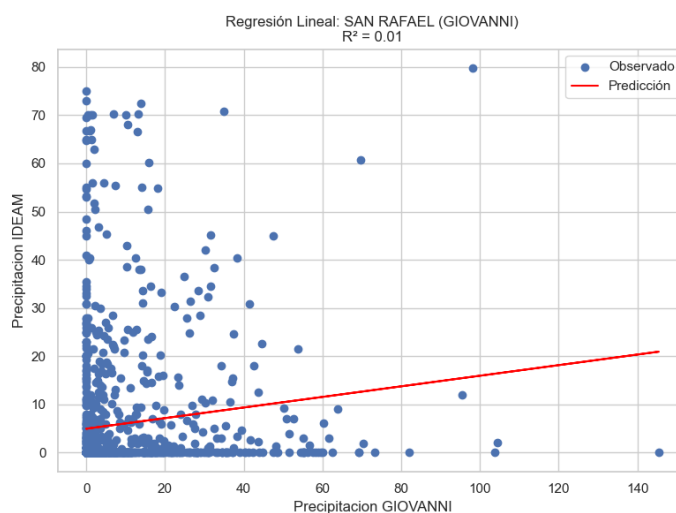
Análisis comparativo de distribución de precipitación datos IDEAM vs datos Giovanni diario estación San Rafael.



Nota. Fuente elaboración propia.

Anexo JJ

Análisis comparativo de distribución de precipitación datos IDEAM vs datos Giovanni diario estación San Rafael.



Nota. Fuente elaboración propia.

Tabla 13

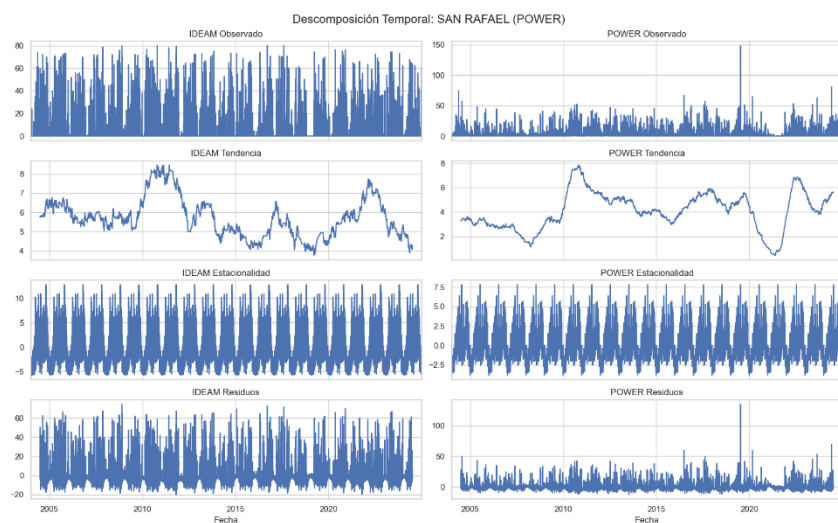
Distribuciones probabilísticas comparados datos IDEAM vs datos Giovanni diario estación San Rafael.

Pearson	Spearman	Kendall	RMSE	MAE	R2
0.106415	0.260133	0.211091	13.37838	8.353286	0.007321

Nota. Fuente elaboración propia.

Anexo KK

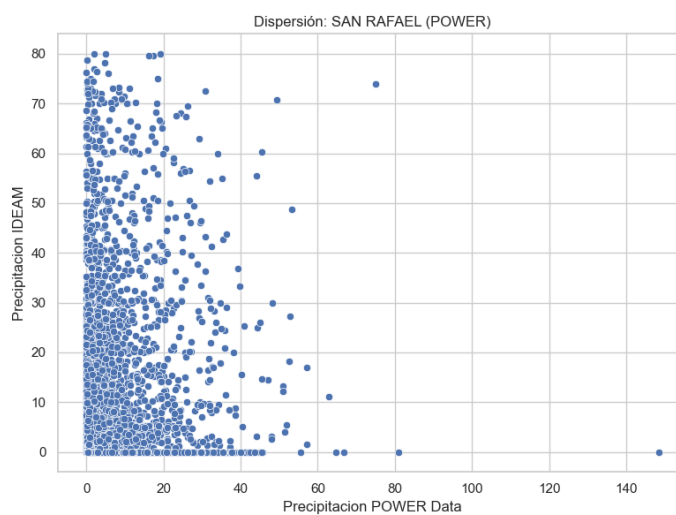
Análisis comparativo datos IDEAM vs datos Power Data diario estación San Rafael.



Nota. Fuente elaboración propia.

Anexo LL

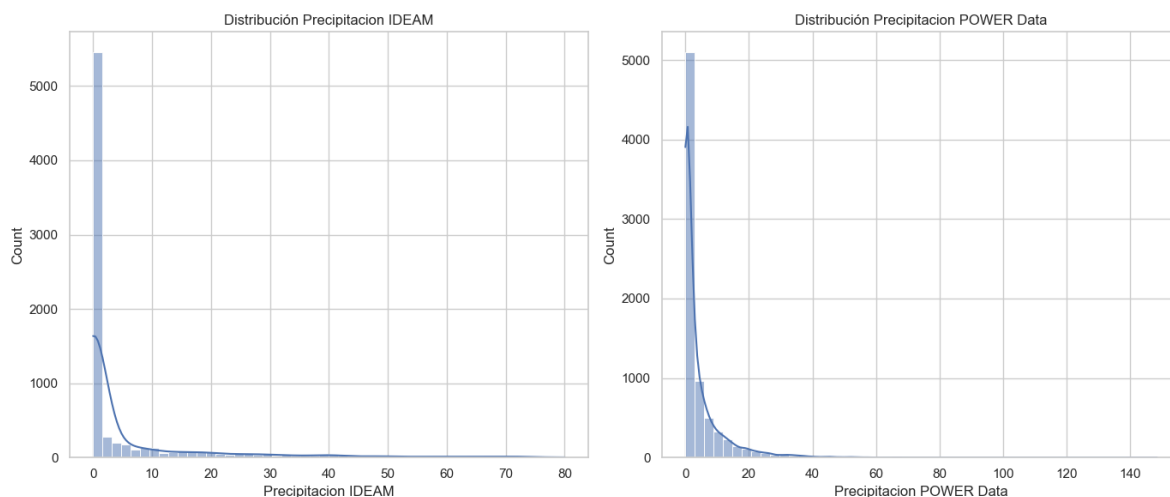
Análisis comparativo de dispersión datos IDEAM vs datos Power Data diario estación San Rafael.



Nota. Fuente elaboración propia.

Anexo MM

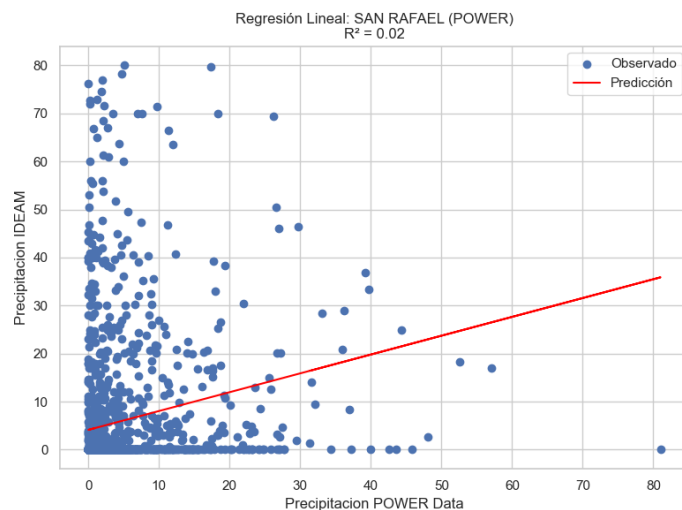
Análisis comparativo de distribución de precipitación datos IDEAM vs datos Power Data diario estación San Rafael.



Nota. Fuente elaboración propia.

Anexo NN

Análisis comparativo de distribución de precipitación datos IDEAM vs datos Power Data diario estación San Rafael.



Nota. Fuente elaboración propia.

Tabla 14

Distribuciones probabilísticas comparados datos IDEAM vs datos Power Data diario estación San Rafael.

Pearson	Spearman	Kendall	RMSE	MAE	R2
0.20247	0.301804	0.230499	12.97296	7.93969	0.015939

Nota. Fuente elaboración propia.

Anexo 6: Análisis comparativo datos a nivel diario estación La Vega.

Anexo OO

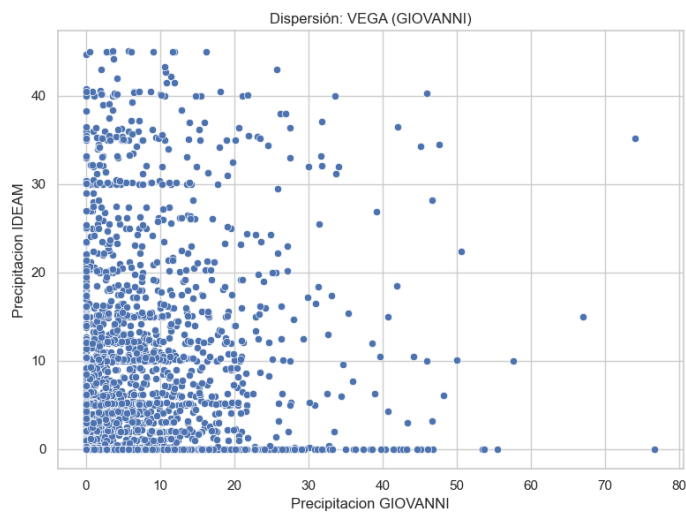
Análisis comparativo datos IDEAM vs datos Giovanni diario estación La Vega .



Nota. Fuente elaboración propia.

Anexo PP

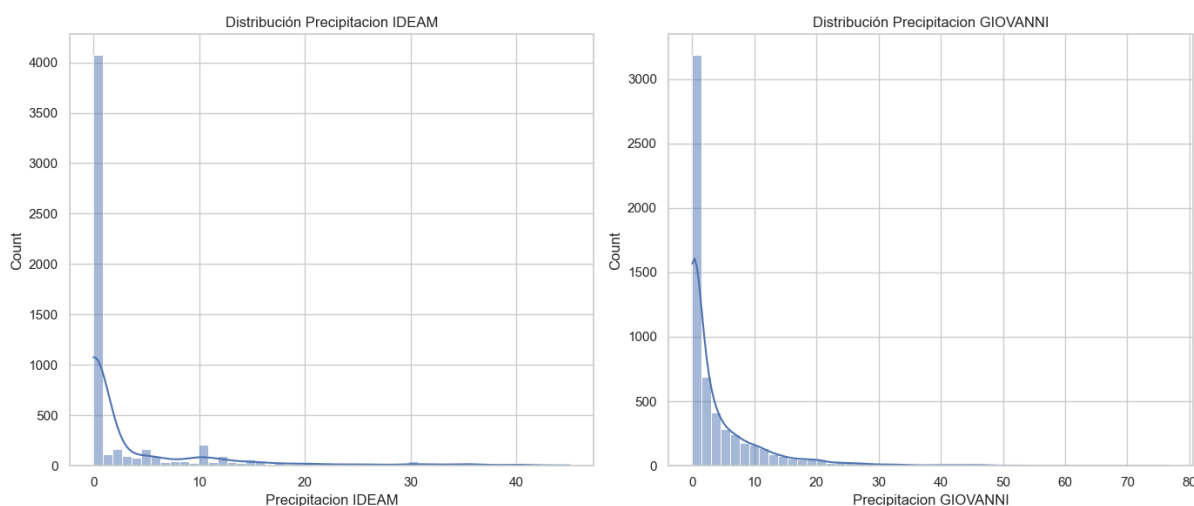
Análisis comparativo de dispersión datos IDEAM vs datos Giovanni diario estación La Vega.



Nota. Fuente elaboración propia.

Anexo QQ

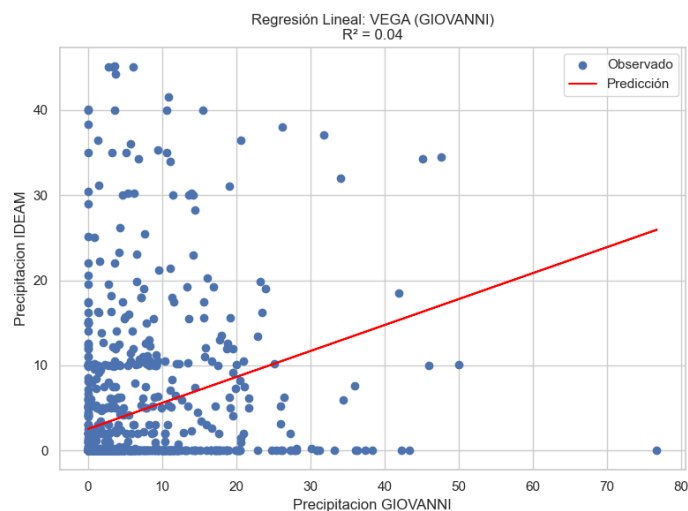
Análisis comparativo de distribución de precipitación datos IDEAM vs datos Giovanni diario estación La Vega.



Nota. Fuente elaboración propia.

Anexo RR

Análisis comparativo de distribución de precipitación datos IDEAM vs datos Giovanni diario estación La Vega.



Nota. Fuente elaboración propia.

Tabla 15

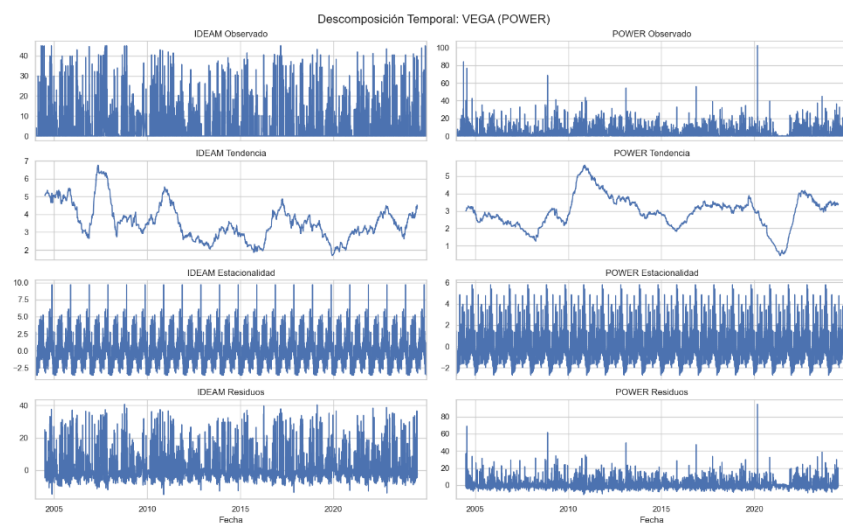
Distribuciones probabilísticas comparados datos IDEAM vs datos Giovanni diario estación La Vega.

Pearson	Spearman	Kendall	RMSE	MAE	R2
0.249971	0.34803	0.280619	7.508669	4.888689	0.044913

Nota. Fuente elaboración propia.

Anexo SS

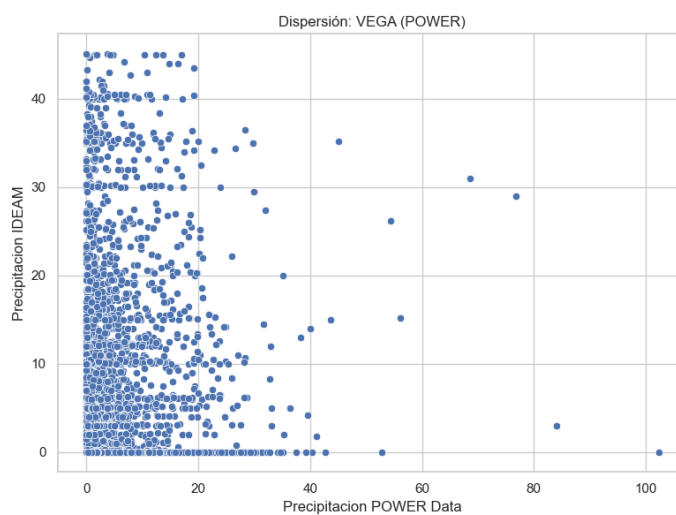
Análisis comparativo datos IDEAM vs datos Power Data diario estación La Vega.



Nota. Fuente elaboración propia.

Anexo TT

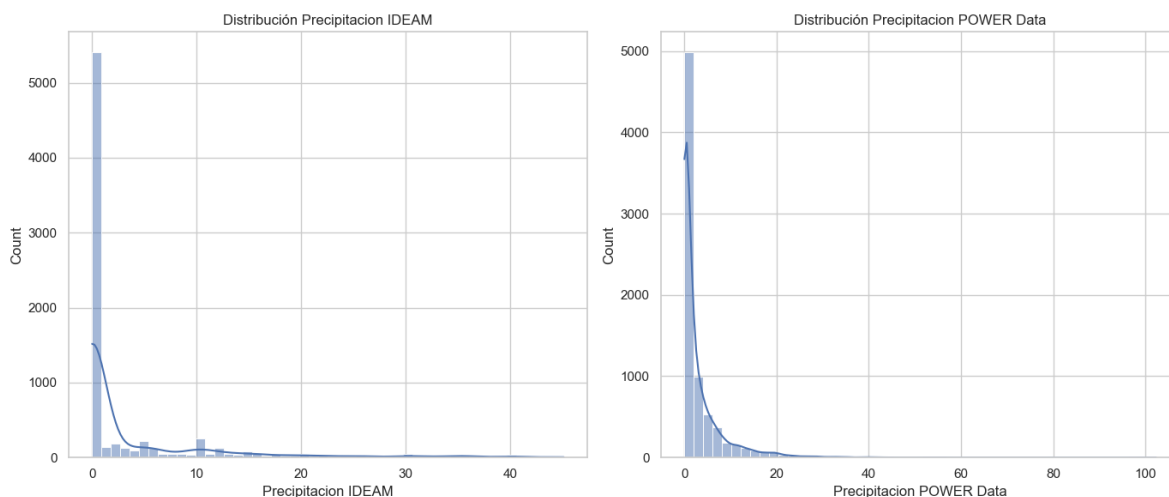
Análisis comparativo de dispersión datos IDEAM vs datos Power Data diario estación La Vega.



Nota. Fuente elaboración propia.

Anexo UU

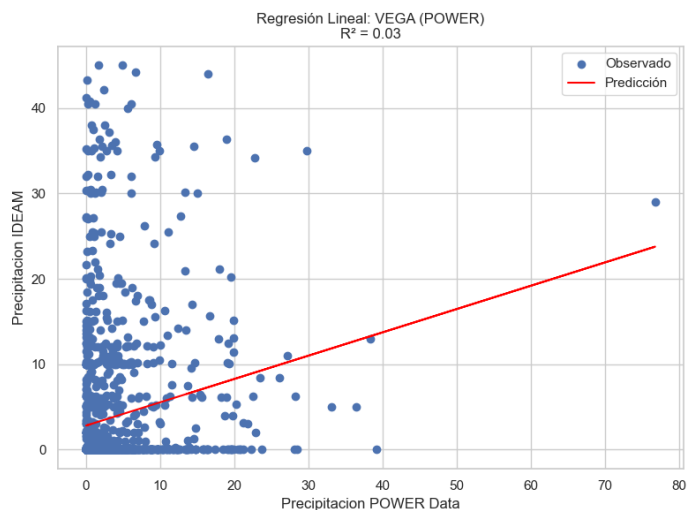
Análisis comparativo de distribución de precipitación datos IDEAM vs datos Power Data diario estación La Vega.



Nota. Fuente elaboración propia.

Anexo VV

Análisis comparativo de distribución de precipitación datos IDEAM vs datos Power Data diario estación La Vega.



Nota. Fuente elaboración propia.

Tabla 16

Distribuciones probabilísticas comparados datos IDEAM vs datos Power Data diario estación La Vega.

Pearson	Spearman	Kendall	RMSE	MAE	R2
0.185364	0.249247	0.191758	7.702653	5.000518	0.031526

Nota. Fuente elaboración propia.

Anexo 7: Análisis comparativo datos a nivel Mensual estación Caobo..

Tabla 17

Valores Medios Mensuales de precipitación IDEAM estación Caobo..

Año	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
2004	150.00	90.00	86.00	182.50	135.00	84.00	137.00	188.00	238.00	189.00	223.50	47.50
2005	106.00	103.00	145.00	333.50	207.00	111.00	96.00	106.00	143.00	349.00	221.00	203.00
2006	124.50	61.00	156.90	279.90	176.50	197.80	132.90	233.80	70.45	249.40	182.70	173.60
2007	6.90	10.00	155.90	223.50	262.25	36.60	153.60	305.60	219.75	295.90	202.50	177.60
2008	0.00	0.00	0.00	82.00	404.50	159.00	181.00	204.00	107.00	245.00	239.00	202.50
2009	131.00	139.00	189.50	75.00	150.00	180.00	151.00	134.00	47.00	268.00	179.00	63.50
2010	63.00	40.00	207.50	142.00	227.50	160.00	208.00	180.00	375.00	276.00	383.50	202.50
2011	91.00	199.00	305.50	245.00	196.00	219.30	245.52	264.74	274.50	302.56	311.10	340.38
2012	359.60	353.80	396.80	402.30	434.62	438.90	152.88	158.00	36.00	33.00	1.89	5.81
2013	9.72	12.15	17.17	20.35	24.88	27.80	32.58	36.49	39.04	44.20	46.50	51.90
2014	55.81	53.78	139.94	86.00	255.00	83.00	74.00	227.00	235.00	306.50	335.00	65.00
2015	118.00	148.50	135.50	148.00	196.00	13.00	39.00	123.00	103.00	169.00	218.00	54.00
2016	8.50	11.00	57.00	134.50	166.00	80.50	77.00	76.00	241.00	317.00	121.50	34.00
2017	97.00	13.00	178.50	154.00	270.00	182.00	132.00	88.00	174.00	244.00	319.50	260.00
2018	247.50	2.00	183.90	141.00	331.00	91.00	45.00	94.00	266.00	137.00	181.00	63.00
2019	41.00	28.00	110.00	213.00	206.00	95.00	137.00	150.00	142.00	221.00	226.00	240.00
2020	104.00	48.00	45.00	131.00	213.00	78.00	65.00	226.00	163.00	308.50	251.00	191.00
2021	11.00	116.00	71.00	349.00	162.00	166.00	118.00	158.00	191.00	260.00	222.00	46.00
2022	62.00	182.00	357.00	229.00	348.50	285.00	178.00	229.00	242.00	272.00	274.00	34.00
2023	238.00	59.00	225.00	217.50	218.18	31.82	88.00	100.00	106.00	294.00	285.00	338.00
2024	0.00	85.00	105.00	81.00	204.00	120.00	90.50	181.00	144.00	118.00	210.50	150.00

Nota. Fuente elaboración propia.

Tabla 18

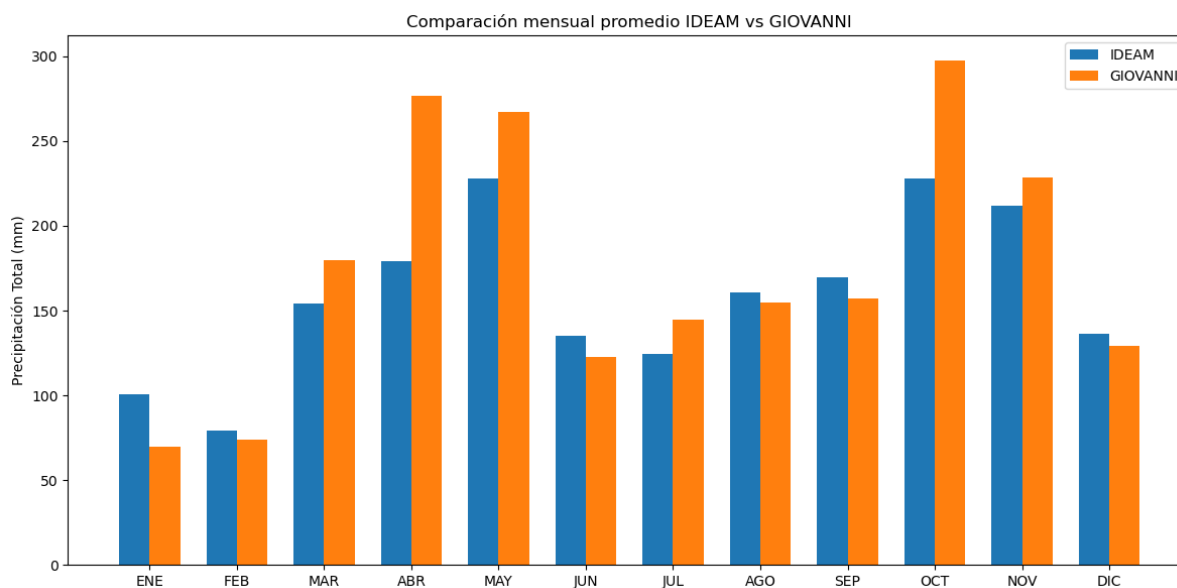
Valores Medios Mensuales de precipitación Giovanni estación Caobo.

Año	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
2004	60.96	81.66	24.51	121.71	317.37	80.25	103.86	66.84	172.47	262.62	342.72	66.18
2005	45.18	24.00	30.60	369.69	252.51	174.36	89.43	229.08	158.67	307.86	147.72	37.62
2006	100.23	21.78	251.43	326.67	199.20	139.29	202.74	200.76	95.40	314.73	300.39	68.88
2007	18.96	6.24	300.90	279.21	261.42	66.87	123.54	180.09	133.68	274.26	163.23	105.96
2008	93.00	83.28	96.72	159.54	346.71	137.40	139.41	124.74	123.54	194.31	411.54	40.53
2009	117.60	39.42	227.04	198.42	198.60	185.25	203.52	139.86	142.53	388.32	160.41	127.47
2010	20.61	78.66	241.50	295.65	246.63	79.77	160.44	98.31	266.10	286.71	191.94	311.46
2011	71.13	191.25	218.43	339.30	241.23	150.00	216.30	118.17	179.94	330.36	316.74	359.13
2012	34.11	15.27	248.22	366.54	163.02	144.75	173.01	135.42	198.36	259.41	139.38	65.01
2013	2.10	137.76	271.14	323.13	292.68	74.64	118.23	231.27	158.88	277.23	150.93	23.91
2014	33.48	94.23	136.47	186.63	300.72	72.66	106.26	212.43	191.28	423.12	230.91	154.92
2015	78.03	209.01	80.91	278.07	323.31	61.68	154.86	139.71	110.79	299.64	255.51	169.08
2016	67.35	29.16	97.53	294.09	256.20	245.76	173.88	155.10	105.48	318.45	273.24	149.43
2017	109.11	29.85	280.02	318.93	213.78	154.38	141.57	139.56	182.82	237.81	139.20	186.18
2018	227.88	7.62	179.19	323.28	361.62	65.52	77.34	179.07	139.65	305.58	216.90	40.83
2019	37.86	131.49	192.54	242.04	299.37	131.64	131.91	122.25	151.02	276.06	212.28	158.40

Nota. Fuente elaboración propia.

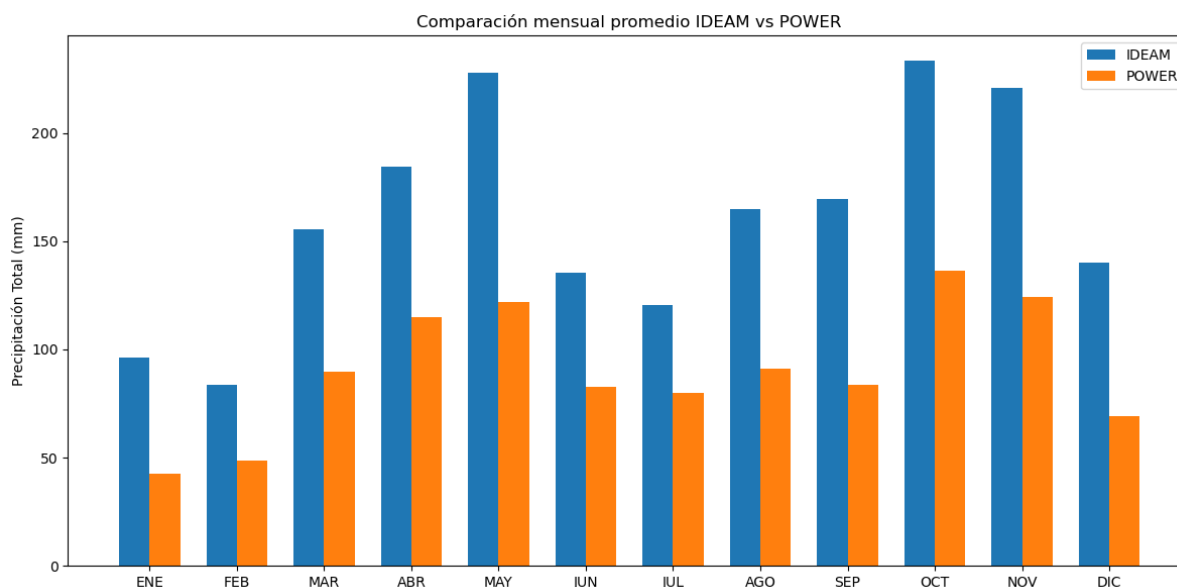
Tabla 19*Valores Medios Mensuales de precipitación Power Data estación Caobo..*

Año	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
2004	29.66	23.71	44.09	145.00	213.60	76.83	184.33	98.43	67.79	80.23	116.47	21.48
2005	82.38	50.02	18.85	137.70	123.61	85.42	49.16	29.98	69.98	137.49	133.61	94.70
2006	84.37	11.37	93.46	88.41	69.80	122.98	26.32	91.56	24.36	108.64	72.98	81.70
2007	49.08	35.69	98.76	50.26	101.65	36.76	40.67	92.88	76.62	61.69	54.86	29.00
2008	5.36	12.87	83.87	18.41	56.15	21.53	41.36	69.45	23.74	162.49	214.25	42.90
2009	31.63	95.67	80.30	114.15	113.26	90.47	81.19	74.90	29.33	105.33	92.91	56.19
2010	5.60	46.28	56.41	96.13	195.45	166.17	147.22	224.57	225.28	194.41	195.39	246.07
2011	60.28	91.07	77.11	193.23	179.71	145.28	143.19	120.25	123.13	208.42	203.67	122.49
2012	46.33	11.86	81.10	206.00	133.86	85.52	111.47	66.68	93.73	269.73	108.31	48.52
2013	29.60	136.02	112.03	111.38	204.68	69.36	41.66	152.81	99.78	98.44	162.14	46.71
2014	24.92	93.57	53.74	65.41	93.13	54.05	86.68	68.24	122.03	189.78	143.69	37.69
2015	41.32	71.76	86.00	103.72	59.84	34.40	56.34	66.75	45.62	75.73	121.74	20.17
2016	23.44	43.28	45.69	90.69	111.54	86.23	88.32	76.07	76.69	98.62	200.74	78.68
2017	44.54	31.42	128.11	139.27	118.95	83.69	69.33	123.51	109.43	164.55	121.10	64.59
2018	47.02	2.70	164.26	169.46	125.59	48.76	89.30	70.97	98.79	237.40	96.77	26.46
2019	47.42	37.36	82.76	138.50	151.85	107.08	102.87	67.85	103.62	146.17	135.86	79.98
2020	75.36	26.55	233.09	52.47	51.95	37.13	45.11	71.03	66.86	59.24	33.47	55.68
2021	52.12	31.49	12.38	0.52	0.00	0.50	0.81	0.20	0.06	7.46	42.38	86.80
2022	18.46	81.54	86.90	204.45	119.84	115.68	109.70	182.37	125.27	168.43	161.63	72.87
2023	81.02	33.65	113.29	155.12	121.54	107.77	81.33	108.23	67.62	170.60	86.71	68.09
2024	14.48	58.20	130.19	128.32	212.61	158.87	82.51	57.84	103.64	115.43	113.23	69.84

Nota. Fuente elaboración propia.**Anexo WW***Análisis comparativo datos IDEAM vs datos Giovanni Mensual estación Caobo.***Nota.** Fuente elaboración propia.

Anexo XX

Análisis comparativo datos IDEAM vs datos Power Data Mensual estación Caobo. .



Nota. Fuente elaboración propia.

Anexo 8: Análisis comparativo datos a nivel Mensual estación Escuela Agrícola Cachirá.

Tabla 20

Valores Medios Mensuales de precipitación IDEAM estación Escuela Agrícola Cachirá.

Año	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
2004	24.10	18.00	30.80	89.00	87.10	15.40	22.70	25.40	128.40	172.85	129.50	17.80
2005	33.00	21.10	8.80	53.50	151.70	99.05	72.80	48.80	96.00	106.75	147.80	9.30
2006	27.40	6.10	66.10	110.35	115.10	36.30	21.35	26.20	82.10	131.10	97.45	11.50
2007	34.20	6.70	44.50	89.40	69.40	37.80	47.60	149.20	123.70	142.70	47.30	25.00
2008	19.80	35.50	50.50	41.45	121.00	48.00	55.80	168.95	134.20	171.65	176.55	22.70
2009	24.80	39.50	140.80	37.15	111.40	89.95	90.80	50.00	51.60	89.10	100.05	27.00
2010	0.50	17.20	17.80	75.80	128.75	89.80	154.90	94.20	225.00	119.95	167.95	92.45
2011	20.50	48.40	71.80	236.00	112.35	84.40	60.30	92.20	69.00	137.50	148.00	83.65
2012	2.50	4.50	117.60	259.50	31.00	34.50	17.40	80.40	77.65	162.40	42.10	13.00
2013	8.80	64.90	15.20	90.45	55.30	7.20	18.40	125.85	78.54	3.46	0.00	0.00
2014	3.20	25.50	14.95	50.40	93.90	28.60	3.65	58.29	80.51	137.94	127.76	24.60
2015	9.70	39.40	34.40	48.63	58.98	9.05	35.80	33.80	13.20	56.44	54.91	22.20
2016	11.20	10.20	27.60	64.05	73.80	99.20	54.60	98.75	55.10	94.81	329.29	258.70
2017	175.79	40.81	138.79	242.07	84.69	40.00	86.90	98.45	119.25	41.45	94.35	9.00
2018	40.80	7.30	81.30	148.10	105.35	75.45	18.40	34.40	73.30	134.65	36.60	10.80
2019	54.25	36.90	64.60	97.25	103.70	41.00	30.20	24.20	63.35	87.90	44.50	53.30
2020	60.40	24.80	46.77	63.85	59.85	51.99	47.60	41.37	34.11	29.12	22.25	16.87
2021	10.64	4.26	33.49	120.50	88.40	94.20	59.50	147.50	80.40	92.90	58.70	6.00
2022	7.60	34.40	56.40	164.20	61.40	118.80	54.50	112.80	140.75	126.50	199.55	9.20
2023	43.00	14.00	36.90	122.30	137.00	84.10	36.35	49.05	62.70	151.40	50.40	24.50
2024	0.00	71.60	43.00	85.90	49.90	61.75	17.80	67.00	30.90	54.90	181.80	5.30

Nota. Fuente elaboración propia.

Tabla 21*Valores Medios Mensuales de precipitación Giovanni estación Escuela Agrícola Cachirá.*

Año	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
2004	43.77	52.68	40.74	117.00	155.13	59.73	35.67	48.63	189.72	265.65	269.04	37.65
2005	79.98	60.54	19.68	234.06	237.87	124.20	113.64	118.77	90.36	264.45	144.42	28.53
2006	71.61	37.68	201.03	177.81	144.57	67.53	91.86	76.35	74.64	222.81	183.24	52.41
2007	33.21	32.40	176.37	154.08	95.82	39.66	71.73	155.25	88.26	230.85	110.58	37.38
2008	81.24	67.83	46.41	87.60	187.56	71.88	107.31	160.62	131.70	170.07	283.86	34.68
2009	63.54	53.10	183.66	91.23	127.56	158.10	85.32	127.53	69.51	302.46	148.44	70.32
2010	10.80	56.10	158.16	198.39	184.71	102.00	157.77	89.31	214.44	174.96	211.11	168.69
2011	33.00	182.13	160.29	280.65	226.14	103.89	89.97	100.86	149.79	197.97	250.20	181.77
2012	13.35	36.33	211.56	316.77	82.92	87.93	56.25	105.72	128.28	249.39	82.59	15.12
2013	15.21	82.26	86.10	105.72	142.29	36.15	76.11	151.80	92.13	201.72	129.87	22.47
2014	40.14	112.32	90.81	145.95	190.59	41.91	46.62	94.53	92.67	292.44	171.81	95.88
2015	72.45	186.90	71.97	198.84	198.66	37.56	105.30	59.85	82.53	191.55	192.12	139.71
2016	63.60	29.31	33.93	173.82	212.40	86.37	79.11	90.72	60.27	228.30	180.54	108.93
2017	73.05	35.94	204.54	246.36	135.15	85.14	73.80	95.07	106.95	127.95	94.47	95.49
2018	106.32	12.57	209.79	253.08	226.59	56.76	34.62	88.68	118.41	329.52	143.49	41.16
2019	68.37	108.87	190.77	177.06	201.15	109.23	56.58	50.97	139.98	156.66	165.63	165.54

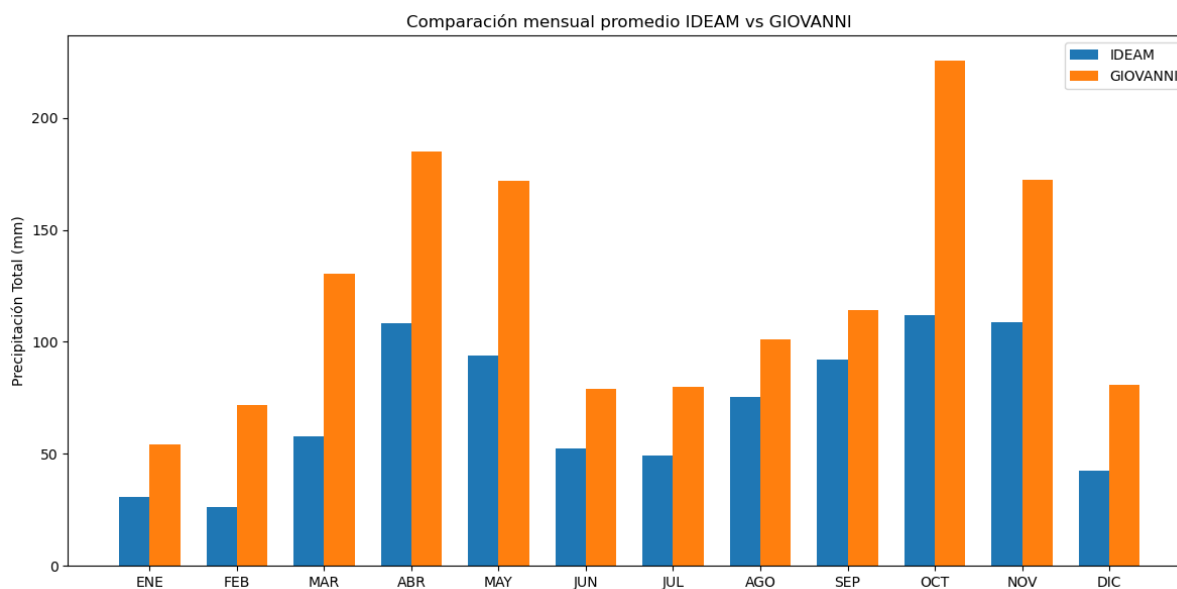
Nota. Fuente elaboración propia.**Tabla 22***Valores Medios Mensuales de precipitación Power Data estación Escuela Agrícola Cachirá.*

Año	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
2004	29.66	23.71	44.09	145.00	213.60	76.83	184.33	98.43	67.79	80.23	116.47	21.48
2005	82.38	50.02	18.85	137.70	123.61	85.42	49.16	29.98	69.98	137.49	133.61	94.70
2006	84.37	11.37	93.46	88.41	69.80	122.98	26.32	91.56	24.36	108.64	72.98	81.70
2007	49.08	35.69	98.76	50.26	101.65	36.76	40.67	92.88	76.62	61.69	54.86	29.00
2008	5.36	12.87	83.87	18.41	56.15	21.53	41.36	69.45	23.74	162.49	214.25	42.90
2009	31.63	95.67	80.30	114.15	113.26	90.47	81.19	74.90	29.33	105.33	92.91	56.19
2010	5.60	46.28	56.41	96.13	195.45	166.17	147.22	224.57	225.28	194.41	195.39	246.07
2011	60.28	91.07	77.11	193.23	179.71	145.28	143.19	120.25	123.13	208.42	203.67	122.49
2012	46.33	11.86	81.10	206.00	133.86	85.52	111.47	66.68	93.73	269.73	108.31	48.52
2013	29.60	136.02	112.03	111.38	204.68	69.36	41.66	152.81	99.78	98.44	162.14	46.71
2014	24.92	93.57	53.74	65.41	93.13	54.05	86.68	68.24	122.03	189.78	143.69	37.69
2015	41.32	71.76	86.00	103.72	59.84	34.40	56.34	66.75	45.62	75.73	121.74	20.17
2016	23.44	43.28	45.69	90.69	111.54	86.23	88.32	76.07	76.69	98.62	200.74	78.68
2017	44.54	31.42	128.11	139.27	118.95	83.69	69.33	123.51	109.43	164.55	121.10	64.59
2018	47.02	2.70	164.26	169.46	125.59	48.76	89.30	70.97	98.79	237.40	96.77	26.46
2019	47.42	37.36	82.76	138.50	151.85	107.08	102.87	67.85	103.62	146.17	135.86	79.98
2020	75.36	26.55	233.09	52.47	51.95	37.13	45.11	71.03	66.86	59.24	33.47	55.68
2021	52.12	31.49	12.38	0.52	0.00	0.50	0.81	0.20	0.06	7.46	42.38	86.80
2022	18.46	81.54	86.90	204.45	119.84	115.68	109.70	182.37	125.27	168.43	161.63	72.87
2023	81.02	33.65	113.29	155.12	121.54	107.77	81.33	108.23	67.62	170.60	86.71	68.09
2024	14.48	58.20	130.19	128.32	212.61	158.87	82.51	57.84	103.64	115.43	113.23	69.84

Nota. Fuente elaboración propia.

Anexo YY

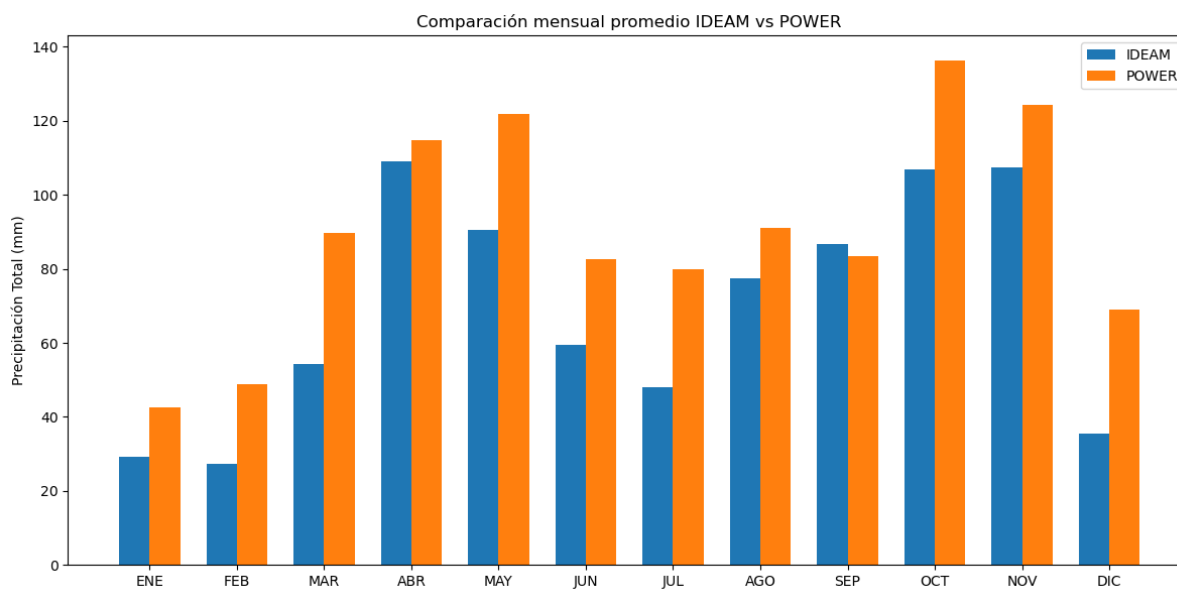
Análisis comparativo datos IDEAM vs datos Giovanni diario estación Escuela Agrícola Cachirá.



Nota. Fuente elaboración propia.

Anexo ZZ

Análisis comparativo datos IDEAM vs datos Power Data diario estación Escuela Agrícola Cachirá.



Nota. Fuente elaboración propia.

Anexo 9: Análisis comparativo datos a nivel Mensual estación Porvenir.

Tabla 23

Valores Medios Mensuales de precipitación IDEAM estación Porvenir.

Año	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
2004	0.00	0.00	20.00	136.00	93.00	100.00	194.00	254.00	279.50	262.00	0.00	149.00
2005	64.00	102.00	173.00	383.00	289.00	143.00	364.50	236.00	245.00	274.00	189.50	247.00
2006	94.00	20.00	160.00	467.00	262.00	128.00	66.00	321.00	280.00	372.00	211.00	109.00
2007	1.00	31.30	252.00	270.00	385.50	89.00	115.00	462.00	219.00	338.00	178.00	140.00
2008	96.00	87.00	238.00	246.00	426.00	433.00	267.00	136.00	304.00	415.00	510.50	51.00
2009	224.00	115.00	250.00	276.00	344.00	343.00	242.00	304.00	110.00	239.00	147.00	258.00
2010	4.00	129.00	355.00	373.00	256.00	253.00	356.00	354.50	402.50	470.00	400.50	181.50
2011	13.00	216.00	285.00	395.50	309.00	338.00	376.00	268.00	209.00	412.50	516.00	540.50
2012	38.00	43.00	134.00	644.00	440.00	299.00	247.00	209.00	217.00	41.00	227.00	227.00
2013	16.00	281.00	241.50	370.00	229.50	96.00	130.00	325.50	213.50	330.00	230.00	71.00
2014	53.00	89.00	187.00	193.00	187.00	203.00	64.00	221.00	201.00	196.00	245.00	41.00
2015	179.00	142.00	151.00	435.00	320.00	650.00	407.33	138.67	218.00	231.00	179.00	7.00
2016	71.00	16.00	19.00	136.00	235.00	94.00	0.00	67.00	384.00	366.50	98.00	91.00
2017	36.00	0.00	300.00	322.00	190.00	247.00	89.00	227.00	195.00	252.00	345.00	232.00
2018	147.00	0.00	199.00	145.00	266.00	122.00	88.00	127.00	467.50	334.00	375.00	70.00
2019	7.00	0.00	58.00	163.00	406.00	143.00	143.00	166.00	219.50	263.50	270.50	190.00
2020	98.00	13.00	40.00	199.00	225.00	336.00	400.00	369.00	478.50	779.00	516.00	166.00
2021	79.00	109.00	247.00	1104.38	381.63	201.00	213.00	302.00	393.00	356.00	373.00	41.60
2022	87.00	135.00	341.40	405.80	339.00	456.00	382.50	385.50	434.00	496.50	479.00	62.00
2023	24.00	0.30	299.00	256.50	209.00	140.00	185.00	373.00	322.00	390.00	258.20	101.00
2024	0.00	4.00	52.00	302.00	528.00	585.00	139.00	117.00	89.00	228.00	280.00	89.00

Nota. Fuente elaboración propia.

Tabla 24

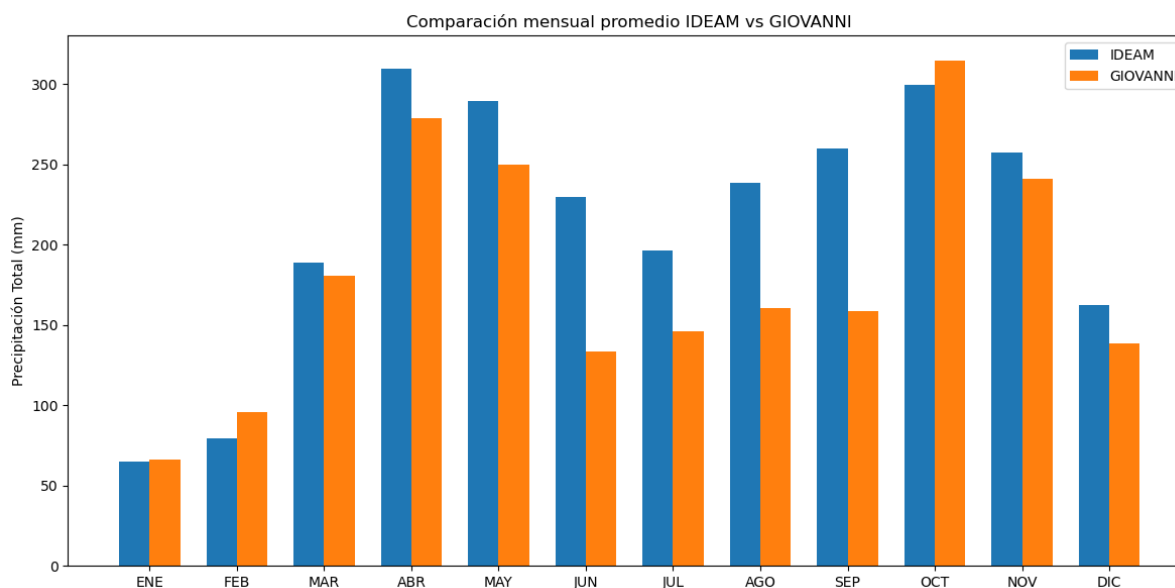
Valores Medios Mensuales de precipitación Giovanni estación Porvenir.

Año	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
2004	26.97	57.75	18.60	141.24	267.99	107.49	114.60	90.81	152.61	279.21	350.16	76.77
2005	42.84	65.61	63.03	409.98	293.16	169.62	105.15	168.60	141.93	294.75	142.83	58.56
2006	75.96	45.33	225.03	264.90	197.31	135.78	185.19	199.26	89.82	403.98	188.40	109.86
2007	12.15	36.00	262.83	226.62	188.67	73.71	97.62	141.18	128.58	243.24	138.00	115.65
2008	91.32	105.90	121.44	145.74	257.91	151.53	142.86	97.38	131.01	227.61	441.24	30.72
2009	100.74	87.09	288.60	165.00	195.78	201.60	200.76	182.55	89.61	393.72	201.27	133.02
2010	11.79	107.28	210.06	245.31	234.84	116.13	155.67	101.40	243.15	311.94	210.87	473.55
2011	68.70	269.91	177.99	366.69	292.65	242.34	234.72	131.85	143.40	421.56	377.91	329.40
2012	58.08	59.43	273.30	404.01	223.59	174.75	113.64	137.64	166.29	381.81	234.45	92.43
2013	34.53	174.39	291.93	316.08	280.38	81.00	149.31	237.09	161.04	291.60	211.56	67.50
2014	13.20	103.56	174.78	273.81	219.81	89.52	88.68	249.63	204.69	340.38	206.04	98.61
2015	100.35	248.79	63.03	258.06	329.01	54.69	158.46	175.92	261.06	296.34	218.04	170.10
2016	57.60	47.22	41.04	257.13	225.87	156.21	200.31	166.77	113.76	375.33	237.00	149.76
2017	115.17	11.19	276.81	426.33	220.20	186.84	180.15	132.30	154.08	296.16	166.86	106.59
2018	167.64	4.59	207.66	286.23	340.02	82.26	102.93	227.94	186.36	266.70	283.11	37.53
2019	85.71	113.55	200.37	272.82	229.26	109.53	106.26	133.20	168.81	208.92	247.20	169.14

Nota. Fuente elaboración propia.

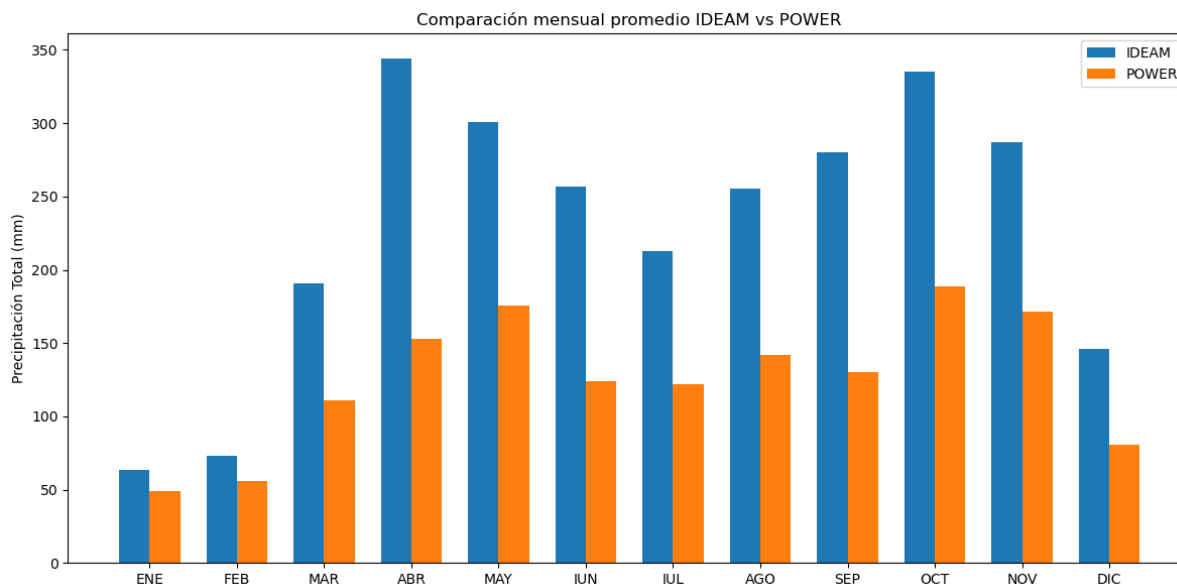
Tabla 25*Valores Medios Mensuales de precipitación Power Data estación Porvenir.*

Año	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
2004	24.53	32.91	80.36	156.63	206.69	70.17	160.74	133.33	93.53	127.08	91.24	16.66
2005	82.05	30.71	43.51	208.42	165.45	81.62	101.41	44.55	62.03	165.98	159.44	100.48
2006	62.34	17.79	99.19	106.96	63.37	124.27	28.05	101.99	42.42	138.10	144.15	75.47
2007	57.43	39.13	111.64	79.50	148.79	41.66	42.17	108.34	83.00	81.40	40.04	26.05
2008	6.53	6.44	67.92	17.42	69.38	17.40	35.35	64.81	20.45	160.26	198.56	33.06
2009	44.16	88.60	73.70	134.16	122.51	121.92	148.52	110.77	52.22	126.25	127.20	83.49
2010	11.22	71.48	99.26	145.04	297.66	277.03	249.26	425.57	288.60	289.21	283.65	260.85
2011	31.95	87.14	109.15	221.17	188.54	181.67	238.93	157.32	173.27	284.40	271.21	152.62
2012	44.91	17.87	96.69	233.08	188.36	132.14	179.54	109.31	144.96	327.35	184.62	46.77
2013	21.57	150.16	152.22	199.86	254.14	136.16	50.67	220.07	176.81	134.07	237.31	58.95
2014	39.33	146.45	82.72	74.32	111.24	61.55	118.85	91.61	214.08	253.72	171.71	60.31
2015	93.44	68.88	117.01	177.12	116.20	42.44	94.86	111.59	89.62	180.40	154.63	45.21
2016	29.39	57.49	57.86	123.87	161.28	152.51	115.95	171.16	160.07	197.09	281.00	83.95
2017	51.97	25.65	196.34	181.87	182.15	147.36	123.60	280.36	229.89	228.60	194.08	124.82
2018	106.42	2.87	211.81	250.05	237.01	60.24	141.98	81.26	167.45	258.75	197.98	14.97
2019	33.05	48.96	96.83	160.85	239.54	194.43	217.67	86.81	153.89	249.34	214.98	152.89
2020	94.18	37.57	203.52	81.73	85.93	33.47	45.32	103.15	45.89	34.68	31.62	36.71
2021	50.14	27.71	11.37	9.61	0.00	0.32	1.34	0.63	3.25	18.42	93.95	93.46
2022	27.70	114.51	151.35	310.55	264.47	264.18	215.19	298.18	244.52	278.48	202.19	45.49
2023	101.00	25.50	160.12	195.68	155.71	155.98	127.15	148.13	102.64	191.30	111.06	75.60
2024	12.93	73.29	107.96	142.69	422.55	301.79	125.18	126.42	184.75	234.93	210.55	107.13

Nota. Fuente elaboración propia.**Anexo AAA***Análisis comparativo datos IDEAM vs datos Giovanni diario estación Porvenir.***Nota.** Fuente elaboración propia.

Anexo BBB

Análisis comparativo datos IDEAM vs datos Power Data diario estación Porvenir.



Nota. Fuente elaboración propia.

Anexo 10: Análisis comparativo datos a nivel Mensual estación Provincia.

Tabla 26

Valores Medios Mensuales de precipitación IDEAM estación Provincia.

Año	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
2004	5.00	14.00	51.00	148.00	177.00	135.00	186.00	326.00	296.00	312.00	235.50	115.00
2005	145.00	172.00	153.00	380.50	323.00	238.00	275.00	247.00	286.00	406.00	207.50	159.00
2006	101.00	43.00	230.00	329.00	216.00	153.50	134.00	391.00	170.00	309.00	384.00	190.00
2007	30.00	43.00	210.00	146.00	355.00	82.00	18.00	243.00	203.00	273.50	213.00	165.00
2008	44.00	156.00	106.00	208.00	376.00	241.00	242.00	224.00	225.00	264.00	397.00	195.00
2009	260.00	162.00	269.00	175.00	310.00	174.00	268.00	336.00	115.00	355.50	138.00	286.50
2010	29.00	109.00	225.00	184.00	337.50	299.00	344.00	275.00	368.00	424.00	280.00	253.50
2011	4.00	202.00	217.00	250.50	301.00	254.00	230.00	198.00	162.00	411.50	427.00	158.00
2012	47.00	58.00	151.00	341.00	126.00	219.00	98.00	171.00	126.00	123.30	230.00	84.00
2013	52.00	172.00	266.00	374.00	305.00	68.00	156.00	357.00	319.00	252.00	198.00	43.00
2014	88.00	83.00	217.00	182.00	144.00	173.00	33.00	170.00	169.00	166.00	256.00	34.00
2015	251.00	46.00	49.00	174.00	279.00	51.00	69.00	76.00	176.00	295.00	182.00	8.00
2016	18.00	22.00	17.00	195.00	140.00	92.00	187.00	136.00	238.00	299.00	212.00	85.00
2017	18.00	5.00	389.00	237.00	262.00	170.00	58.00	312.00	187.00	224.00	402.00	231.40
2018	153.00	0.00	198.00	192.00	261.00	147.00	79.00	147.00	338.00	269.00	183.00	42.00
2019	53.00	65.00	182.50	183.00	293.00	142.30	190.00	96.00	198.00	309.00	140.00	215.00
2020	43.00	8.00	81.00	308.00	99.00	269.00	330.00	116.00	235.00	402.50	200.00	39.00
2021	61.00	102.00	63.00	370.00	409.50	403.00	191.50	307.00	144.00	244.00	279.00	168.00

Año	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
2022	96.20	100.00	219.00	476.00	240.00	209.50	329.00	391.00	219.00	279.00	219.00	107.00
2023	140.00	0.30	135.00	191.00	194.00	197.00	220.00	248.00	127.00	325.00	400.00	133.00
2024	0.00	25.00	56.00	157.00	215.50	458.50	263.00	110.00	160.33	124.67	259.50	112.00

Nota. Fuente elaboración propia.

Tabla 27

Valores Medios Mensuales de precipitación Giovanni estación Provincia.

Año	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
2004	26.97	57.75	18.60	141.24	267.99	107.49	114.60	90.81	152.61	279.21	350.16	76.77
2005	42.84	65.61	63.03	409.98	293.16	169.62	105.15	168.60	141.93	294.75	142.83	58.56
2006	75.96	45.33	225.03	264.90	197.31	135.78	185.19	199.26	89.82	403.98	188.40	109.86
2007	12.15	36.00	262.83	226.62	188.67	73.71	97.62	141.18	128.58	243.24	138.00	115.65
2008	91.32	105.90	121.44	145.74	257.91	151.53	142.86	97.38	131.01	227.61	441.24	30.72
2009	100.74	87.09	288.60	165.00	195.78	201.60	200.76	182.55	89.61	393.72	201.27	133.02
2010	11.79	107.28	210.06	245.31	234.84	116.13	155.67	101.40	243.15	311.94	210.87	473.55
2011	68.70	269.91	177.99	366.69	292.65	242.34	234.72	131.85	143.40	421.56	377.91	329.40
2012	58.08	59.43	273.30	404.01	223.59	174.75	113.64	137.64	166.29	381.81	234.45	92.43
2013	34.53	174.39	291.93	316.08	280.38	81.00	149.31	237.09	161.04	291.60	211.56	67.50
2014	13.20	103.56	174.78	273.81	219.81	89.52	88.68	249.63	204.69	340.38	206.04	98.61
2015	100.35	248.79	63.03	258.06	329.01	54.69	158.46	175.92	261.06	296.34	218.04	170.10
2016	57.60	47.22	41.04	257.13	225.87	156.21	200.31	166.77	113.76	375.33	237.00	149.76
2017	115.17	11.19	276.81	426.33	220.20	186.84	180.15	132.30	154.08	296.16	166.86	106.59
2018	167.64	4.59	207.66	286.23	340.02	82.26	102.93	227.94	186.36	266.70	283.11	37.53
2019	85.71	113.55	200.37	272.82	229.26	109.53	106.26	133.20	168.81	208.92	247.20	169.14

Nota. Fuente elaboración propia.

Tabla 28

Valores Medios Mensuales de precipitación Power Data estación Provincia.

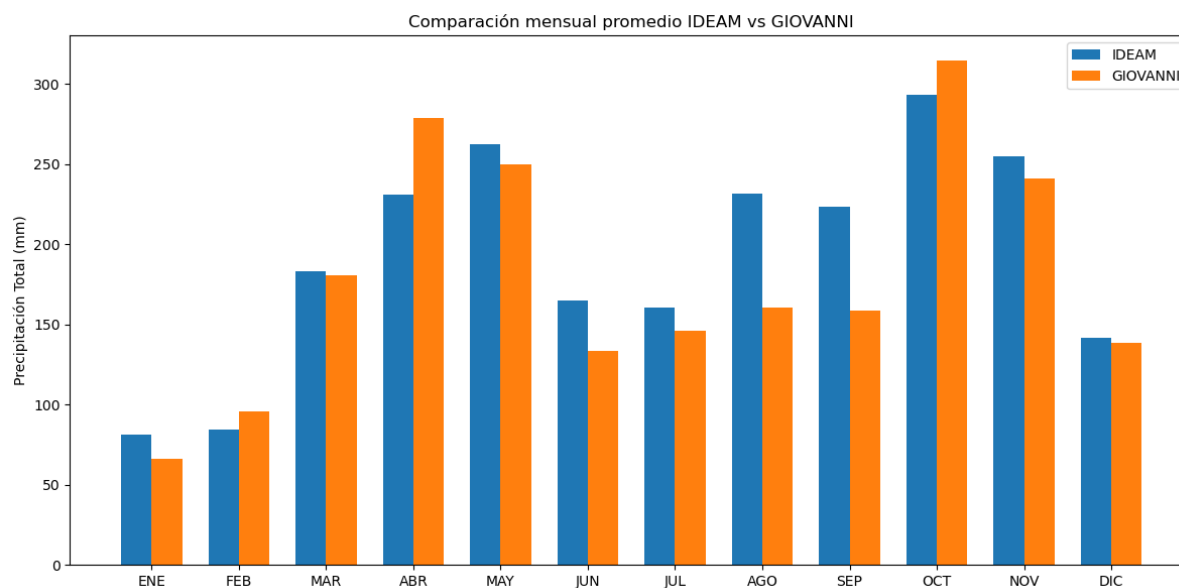
Año	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
2004	24.53	32.91	80.36	156.63	206.69	70.17	160.74	133.33	93.53	127.08	91.24	16.66
2005	82.05	30.71	43.51	208.42	165.45	81.62	101.41	44.55	62.03	165.98	159.44	100.48
2006	62.34	17.79	99.19	106.96	63.37	124.27	28.05	101.99	42.42	138.10	144.15	75.47
2007	57.43	39.13	111.64	79.50	148.79	41.66	42.17	108.34	83.00	81.40	40.04	26.05
2008	6.53	6.44	67.92	17.42	69.38	17.40	35.35	64.81	20.45	160.26	198.56	33.06
2009	44.16	88.60	73.70	134.16	122.51	121.92	148.52	110.77	52.22	126.25	127.20	83.49
2010	11.22	71.48	99.26	145.04	297.66	277.03	249.26	425.57	288.60	289.21	283.65	260.85
2011	31.95	87.14	109.15	221.17	188.54	181.67	238.93	157.32	173.27	284.40	271.21	152.62
2012	44.91	17.87	96.69	233.08	188.36	132.14	179.54	109.31	144.96	327.35	184.62	46.77
2013	21.57	150.16	152.22	199.86	254.14	136.16	50.67	220.07	176.81	134.07	237.31	58.95
2014	39.33	146.45	82.72	74.32	111.24	61.55	118.85	91.61	214.08	253.72	171.71	60.31
2015	93.44	68.88	117.01	177.12	116.20	42.44	94.86	111.59	89.62	180.40	154.63	45.21
2016	29.39	57.49	57.86	123.87	161.28	152.51	115.95	171.16	160.07	197.09	281.00	83.95
2017	51.97	25.65	196.34	181.87	182.15	147.36	123.60	280.36	229.89	228.60	194.08	124.82
2018	106.42	2.87	211.81	250.05	237.01	60.24	141.98	81.26	167.45	258.75	197.98	14.97

Año	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
2019	33.05	48.96	96.83	160.85	239.54	194.43	217.67	86.81	153.89	249.34	214.98	152.89
2020	94.18	37.57	203.52	81.73	85.93	33.47	45.32	103.15	45.89	34.68	31.62	36.71
2021	50.14	27.71	11.37	9.61	0.00	0.32	1.34	0.63	3.25	18.42	93.95	93.46
2022	27.70	114.51	151.35	310.55	264.47	264.18	215.19	298.18	244.52	278.48	202.19	45.49
2023	101.00	25.50	160.12	195.68	155.71	155.98	127.15	148.13	102.64	191.30	111.06	75.60
2024	12.93	73.29	107.96	142.69	422.55	301.79	125.18	126.42	184.75	234.93	210.55	107.13

Nota. Fuente elaboración propia.

Anexo CCC

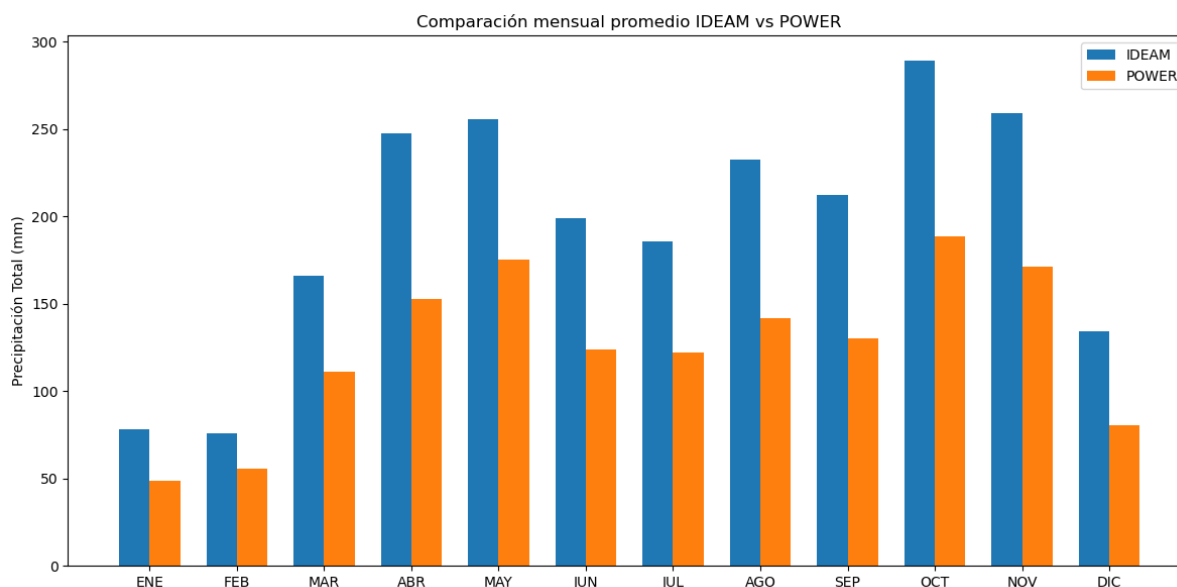
Análisis comparativo datos IDEAM vs datos Giovanni diario estación Provincia.



Nota. Fuente elaboración propia.

Anexo DDD

Análisis comparativo datos IDEAM vs datos Power Data diario estación Provincia.



Nota. Fuente elaboración propia.

Anexo 11: Análisis comparativo datos a nivel Mensual estación San Rafael.

Tabla 29

Valores Medios Mensuales de precipitación IDEAM estación San Rafael.

Año	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
2004	44.00	21.00	72.00	175.00	303.50	82.00	265.00	257.00	284.00	266.00	316.00	35.00
2005	30.00	22.00	132.00	360.00	226.50	236.00	155.00	314.00	258.00	311.00	219.50	27.00
2006	100.00	2.00	197.50	197.00	246.00	261.00	127.00	153.00	208.00	257.00	290.00	48.00
2007	1.00	27.00	196.50	168.00	439.00	121.00	180.00	260.50	169.00	208.00	224.20	164.10
2008	40.20	18.40	108.90	149.20	288.60	279.10	94.85	180.50	176.60	264.90	402.20	101.90
2009	26.10	29.90	110.90	162.95	151.40	290.60	98.70	191.60	180.60	384.40	221.20	59.90
2010	0.00	87.70	355.20	163.10	410.00	189.90	163.30	481.20	286.95	343.50	305.40	159.90
2011	11.10	93.80	228.10	372.50	249.90	248.00	295.90	312.30	148.40	378.50	405.25	242.90
2012	0.00	4.00	90.10	351.80	203.20	170.00	168.80	156.80	182.30	214.80	251.30	28.00
2013	14.00	87.15	298.60	258.80	475.20	146.70	77.80	277.10	239.00	232.20	198.00	33.90
2014	24.50	53.30	146.30	154.70	251.10	71.30	43.10	222.70	205.80	453.30	190.80	39.80
2015	93.20	69.00	57.90	187.60	217.20	36.10	195.70	72.40	231.40	283.20	245.80	26.20
2016	10.00	5.00	43.30	238.50	151.30	83.40	187.40	78.60	243.50	338.20	169.20	197.60
2017	9.50	0.00	142.40	133.80	350.95	278.50	184.60	340.20	85.30	128.60	204.05	39.80
2018	94.50	0.00	224.00	147.95	248.10	105.60	175.30	44.10	279.95	363.30	0.00	0.00
2019	0.00	0.00	183.70	138.75	222.95	144.50	87.90	136.70	201.85	462.00	119.80	74.10
2020	30.20	0.00	22.00	110.40	219.80	132.30	260.20	226.50	235.80	318.20	347.20	63.10
2021	29.50	9.80	125.40	172.10	238.70	180.15	147.40	303.80	234.50	349.80	220.60	22.20
2022	49.80	118.10	223.20	378.20	189.40	218.75	339.50	319.05	310.00	187.30	210.60	3.10
2023	132.80	58.80	120.20	149.60	260.80	117.05	246.10	112.30	167.60	371.40	137.90	143.80
2024	0.00	7.00	44.80	138.80	215.20	327.70	108.65	120.30	126.30	136.60	210.80	128.30

Nota. Fuente elaboración propia.

Tabla 30*Valores Medios Mensuales de precipitación Giovanni estación San Rafael.*

Año	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
2004	44.00	21.00	72.00	175.00	303.50	82.00	265.00	257.00	284.00	266.00	316.00	35.00
2005	30.00	22.00	132.00	360.00	226.50	236.00	155.00	314.00	258.00	311.00	219.50	27.00
2006	100.00	2.00	197.50	197.00	246.00	261.00	127.00	153.00	208.00	257.00	290.00	48.00
2007	1.00	27.00	196.50	168.00	439.00	121.00	180.00	260.50	169.00	208.00	224.20	164.10
2008	40.20	18.40	108.90	149.20	288.60	279.10	94.85	180.50	176.60	264.90	402.20	101.90
2009	26.10	29.90	110.90	162.95	151.40	290.60	98.70	191.60	180.60	384.40	221.20	59.90
2010	0.00	87.70	355.20	163.10	410.00	189.90	163.30	481.20	286.95	343.50	305.40	159.90
2011	11.10	93.80	228.10	372.50	249.90	248.00	295.90	312.30	148.40	378.50	405.25	242.90
2012	0.00	4.00	90.10	351.80	203.20	170.00	168.80	156.80	182.30	214.80	251.30	28.00
2013	14.00	87.15	298.60	258.80	475.20	146.70	77.80	277.10	239.00	232.20	198.00	33.90
2014	24.50	53.30	146.30	154.70	251.10	71.30	43.10	222.70	205.80	453.30	190.80	39.80
2015	93.20	69.00	57.90	187.60	217.20	36.10	195.70	72.40	231.40	283.20	245.80	26.20
2016	10.00	5.00	43.30	238.50	151.30	83.40	187.40	78.60	243.50	338.20	169.20	197.60
2017	9.50	0.00	142.40	133.80	350.95	278.50	184.60	340.20	85.30	128.60	204.05	39.80
2018	94.50	0.00	224.00	147.95	248.10	105.60	175.30	44.10	279.95	363.30	0.00	0.00
2019	0.00	0.00	183.70	138.75	222.95	144.50	87.90	136.70	201.85	462.00	119.80	74.10

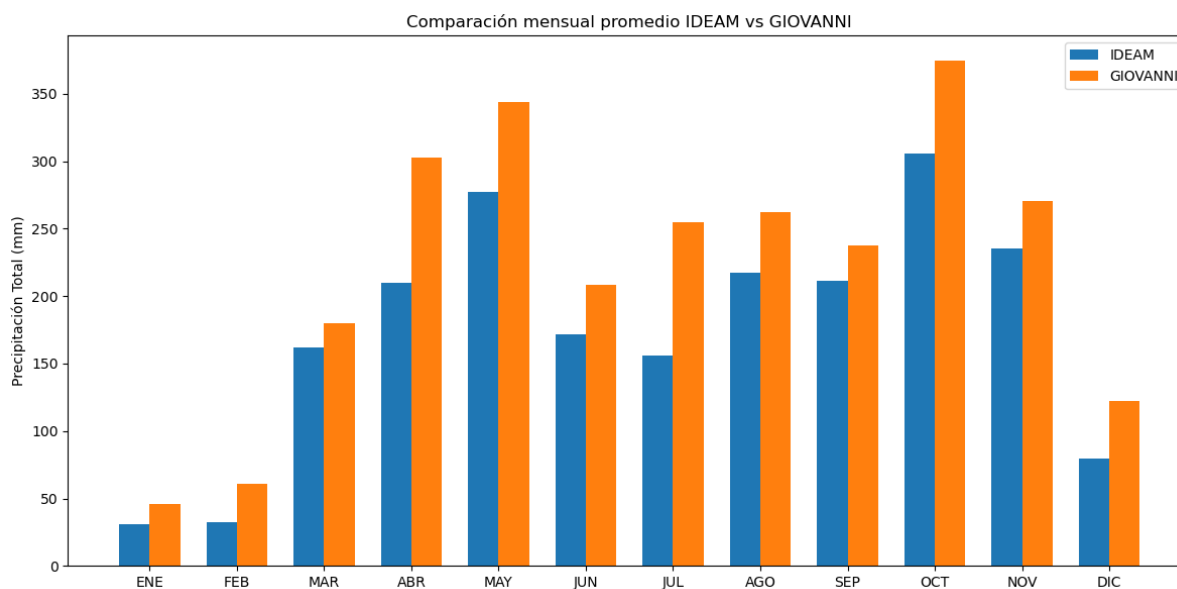
Nota. Fuente elaboración propia.**Tabla 31***Valores Medios Mensuales de precipitación Power Data estación San Rafael.*

Año	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
2004	24.53	32.91	80.36	156.63	206.69	70.17	160.74	133.33	93.53	127.08	91.24	16.66
2005	82.05	30.71	43.51	208.42	165.45	81.62	101.41	44.55	62.03	165.98	159.44	100.48
2006	62.34	17.79	99.19	106.96	63.37	124.27	28.05	101.99	42.42	138.10	144.15	75.47
2007	57.43	39.13	111.64	79.50	148.79	41.66	42.17	108.34	83.00	81.40	40.04	26.05
2008	6.53	6.44	67.92	17.42	69.38	17.40	35.35	64.81	20.45	160.26	198.56	33.06
2009	44.16	88.60	73.70	134.16	122.51	121.92	148.52	110.77	52.22	126.25	127.20	83.49
2010	11.22	71.48	99.26	145.04	297.66	277.03	249.26	425.57	288.60	289.21	283.65	260.85
2011	31.95	87.14	109.15	221.17	188.54	181.67	238.93	157.32	173.27	284.40	271.21	152.62
2012	44.91	17.87	96.69	233.08	188.36	132.14	179.54	109.31	144.96	327.35	184.62	46.77
2013	21.57	150.16	152.22	199.86	254.14	136.16	50.67	220.07	176.81	134.07	237.31	58.95
2014	39.33	146.45	82.72	74.32	111.24	61.55	118.85	91.61	214.08	253.72	171.71	60.31
2015	93.44	68.88	117.01	177.12	116.20	42.44	94.86	111.59	89.62	180.40	154.63	45.21
2016	29.39	57.49	57.86	123.87	161.28	152.51	115.95	171.16	160.07	197.09	281.00	83.95
2017	51.97	25.65	196.34	181.87	182.15	147.36	123.60	280.36	229.89	228.60	194.08	124.82
2018	106.42	2.87	211.81	250.05	237.01	60.24	141.98	81.26	167.45	258.75	197.98	14.97
2019	33.05	48.96	96.83	160.85	239.54	194.43	217.67	86.81	153.89	249.34	214.98	152.89
2020	94.18	37.57	203.52	81.73	85.93	33.47	45.32	103.15	45.89	34.68	31.62	36.71
2021	50.14	27.71	11.37	9.61	0.00	0.32	1.34	0.63	3.25	18.42	93.95	93.46
2022	27.70	114.51	151.35	310.55	264.47	264.18	215.19	298.18	244.52	278.48	202.19	45.49
2023	101.00	25.50	160.12	195.68	155.71	155.98	127.15	148.13	102.64	191.30	111.06	75.60
2024	12.93	73.29	107.96	142.69	422.55	301.79	125.18	126.42	184.75	234.93	210.55	107.13

Nota. Fuente elaboración propia.

Anexo EEE

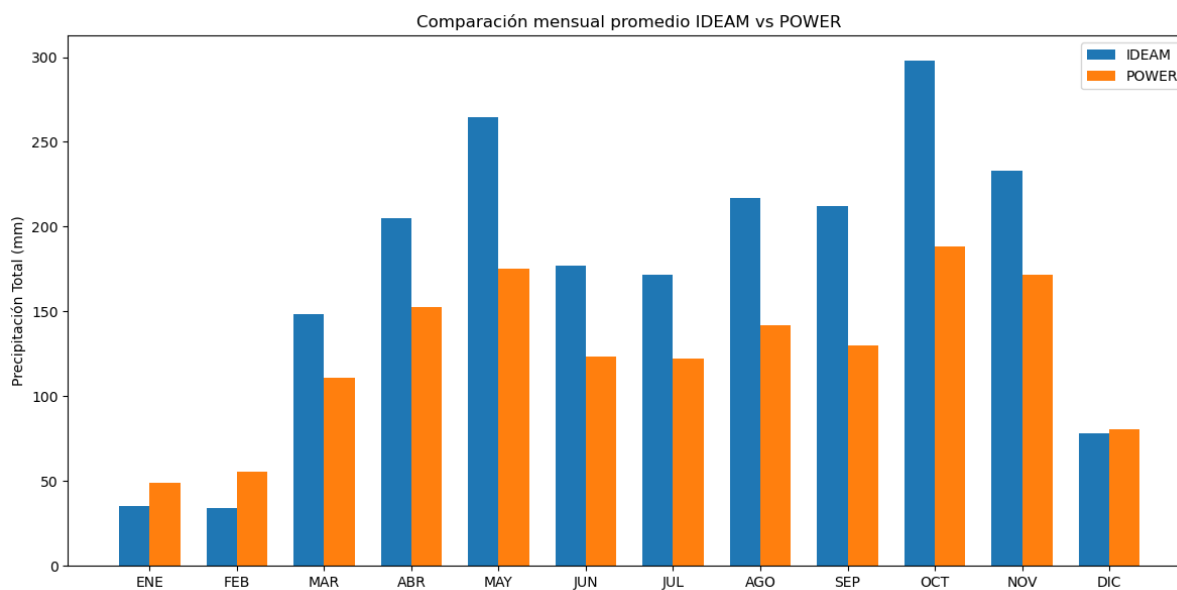
Análisis comparativo datos IDEAM vs datos Giovanni diario estación San Rafael.



Nota. Fuente elaboración propia.

Anexo FFF

Análisis comparativo datos IDEAM vs datos Power Data diario estación San Rafael.



Nota. Fuente elaboración propia.

Anexo 12: Análisis comparativo datos a nivel Mensual estación La Vega

Tabla 32

Valores Medios Mensuales de precipitación IDEAM estación La Vega

Año	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
2004	13.00	71.00	61.50	224.00	277.00	103.00	97.00	32.00	222.00	279.00	376.50	99.00
2005	53.00	71.00	8.00	112.00	244.00	258.00	101.00	110.00	243.00	200.50	217.50	163.00
2006	79.50	7.50	176.00	177.50	112.20	56.00	123.20	0.00	76.35	106.90	219.20	19.10
2007	52.70	0.00	139.25	109.00	321.50	104.00	298.20	474.70	314.10	365.80	114.60	28.40
2008	25.40	37.80	76.30	78.70	138.40	2.30	0.00	275.70	119.20	147.20	368.30	20.00
2009	76.20	58.70	149.00	55.80	31.00	82.00	58.40	61.80	220.50	249.10	171.00	34.50
2010	0.00	11.40	106.70	116.10	150.20	71.90	64.50	179.20	233.40	138.80	169.80	195.00
2011	67.20	184.50	72.60	341.95	215.55	107.40	51.40	73.30	56.50	184.90	152.45	35.30
2012	13.40	3.30	45.45	205.15	102.45	66.30	62.40	103.30	91.30	194.60	15.70	6.10
2013	0.00	83.00	80.70	117.00	96.05	0.00	48.35	130.90	65.10	117.10	49.00	17.30
2014	119.70	159.50	56.50	137.40	232.50	71.60	16.30	55.10	122.20	161.90	131.45	10.10
2015	9.40	89.90	70.00	176.00	145.20	20.70	61.75	46.80	15.45	72.55	114.25	0.00
2016	12.60	14.30	73.10	213.80	99.80	83.00	0.00	30.90	60.05	257.75	290.20	0.00
2017	40.20	55.20	368.95	207.10	0.00	158.80	84.10	89.00	206.90	125.05	107.60	0.00
2018	71.30	9.90	164.30	232.80	187.80	51.90	43.40	16.20	55.70	238.20	225.95	39.90
2019	57.30	98.00	100.35	107.20	205.40	50.85	13.20	68.30	118.00	82.90	43.80	106.95
2020	25.10	59.20	4.80	52.70	30.50	24.95	100.40	119.45	102.05	50.00	206.30	10.40
2021	0.00	35.80	121.25	179.60	96.00	92.00	53.30	208.30	77.95	128.10	89.45	8.40
2022	35.40	108.40	97.60	69.80	164.20	152.90	50.15	106.50	143.00	258.65	180.15	39.10
2023	111.10	9.70	210.10	201.05	133.40	30.85	24.40	42.40	51.00	207.30	133.70	81.50
2024	10.50	69.00	57.00	205.00	275.00	201.00	128.50	18.00	117.00	134.00	205.00	199.00

Nota. Fuente elaboración propia.

Tabla 33

Valores Medios Mensuales de precipitación Giovanni estación Provincia.

Año	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
2004	43.77	52.68	40.74	117.00	155.13	59.73	35.67	48.63	189.72	265.65	269.04	37.65
2005	79.98	60.54	19.68	234.06	237.87	124.20	113.64	118.77	90.36	264.45	144.42	28.53
2006	71.61	37.68	201.03	177.81	144.57	67.53	91.86	76.35	74.64	222.81	183.24	52.41
2007	33.21	32.40	176.37	154.08	95.82	39.66	71.73	155.25	88.26	230.85	110.58	37.38
2008	81.24	67.83	46.41	87.60	187.56	71.88	107.31	160.62	131.70	170.07	283.86	34.68
2009	63.54	53.10	183.66	91.23	127.56	158.10	85.32	127.53	69.51	302.46	148.44	70.32
2010	10.80	56.10	158.16	198.39	184.71	102.00	157.77	89.31	214.44	174.96	211.11	168.69
2011	33.00	182.13	160.29	280.65	226.14	103.89	89.97	100.86	149.79	197.97	250.20	181.77
2012	13.35	36.33	211.56	316.77	82.92	87.93	56.25	105.72	128.28	249.39	82.59	15.12
2013	15.21	82.26	86.10	105.72	142.29	36.15	76.11	151.80	92.13	201.72	129.87	22.47
2014	40.14	112.32	90.81	145.95	190.59	41.91	46.62	94.53	92.67	292.44	171.81	95.88
2015	72.45	186.90	71.97	198.84	198.66	37.56	105.30	59.85	82.53	191.55	192.12	139.71
2016	63.60	29.31	33.93	173.82	212.40	86.37	79.11	90.72	60.27	228.30	180.54	108.93
2017	73.05	35.94	204.54	246.36	135.15	85.14	73.80	95.07	106.95	127.95	94.47	95.49

Año	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
2018	106.32	12.57	209.79	253.08	226.59	56.76	34.62	88.68	118.41	329.52	143.49	41.16
2019	68.37	108.87	190.77	177.06	201.15	109.23	56.58	50.97	139.98	156.66	165.63	165.54

Nota. Fuente elaboración propia.

Tabla 34

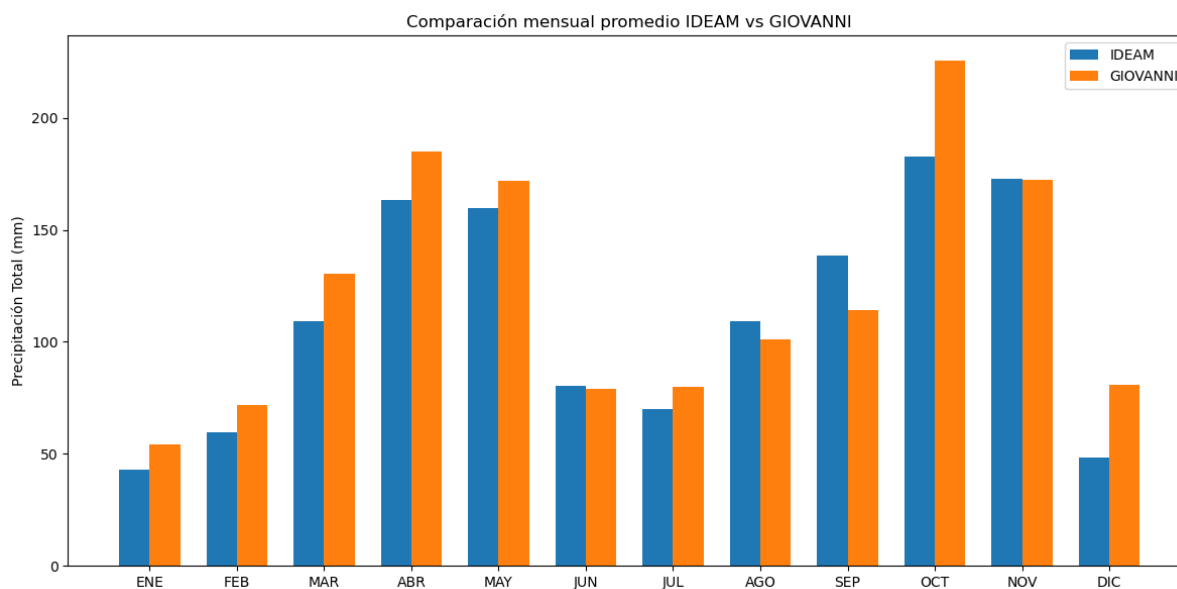
Valores Medios Mensuales de precipitación Power Data estación Provincia.

Año	ENE	FEB	MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT	NOV	DIC
2004	29.66	23.71	44.09	145.00	213.60	76.83	184.33	98.43	67.79	80.23	116.47	21.48
2005	82.38	50.02	18.85	137.70	123.61	85.42	49.16	29.98	69.98	137.49	133.61	94.70
2006	84.37	11.37	93.46	88.41	69.80	122.98	26.32	91.56	24.36	108.64	72.98	81.70
2007	49.08	35.69	98.76	50.26	101.65	36.76	40.67	92.88	76.62	61.69	54.86	29.00
2008	5.36	12.87	83.87	18.41	56.15	21.53	41.36	69.45	23.74	162.49	214.25	42.90
2009	31.63	95.67	80.30	114.15	113.26	90.47	81.19	74.90	29.33	105.33	92.91	56.19
2010	5.60	46.28	56.41	96.13	195.45	166.17	147.22	224.57	225.28	194.41	195.39	246.07
2011	60.28	91.07	77.11	193.23	179.71	145.28	143.19	120.25	123.13	208.42	203.67	122.49
2012	46.33	11.86	81.10	206.00	133.86	85.52	111.47	66.68	93.73	269.73	108.31	48.52
2013	29.60	136.02	112.03	111.38	204.68	69.36	41.66	152.81	99.78	98.44	162.14	46.71
2014	24.92	93.57	53.74	65.41	93.13	54.05	86.68	68.24	122.03	189.78	143.69	37.69
2015	41.32	71.76	86.00	103.72	59.84	34.40	56.34	66.75	45.62	75.73	121.74	20.17
2016	23.44	43.28	45.69	90.69	111.54	86.23	88.32	76.07	76.69	98.62	200.74	78.68
2017	44.54	31.42	128.11	139.27	118.95	83.69	69.33	123.51	109.43	164.55	121.10	64.59
2018	47.02	2.70	164.26	169.46	125.59	48.76	89.30	70.97	98.79	237.40	96.77	26.46
2019	47.42	37.36	82.76	138.50	151.85	107.08	102.87	67.85	103.62	146.17	135.86	79.98
2020	75.36	26.55	233.09	52.47	51.95	37.13	45.11	71.03	66.86	59.24	33.47	55.68
2021	52.12	31.49	12.38	0.52	0.00	0.50	0.81	0.20	0.06	7.46	42.38	86.80
2022	18.46	81.54	86.90	204.45	119.84	115.68	109.70	182.37	125.27	168.43	161.63	72.87
2023	81.02	33.65	113.29	155.12	121.54	107.77	81.33	108.23	67.62	170.60	86.71	68.09
2024	14.48	58.20	130.19	128.32	212.61	158.87	82.51	57.84	103.64	115.43	113.23	69.84

Nota. Fuente elaboración propia.

Anexo GGG

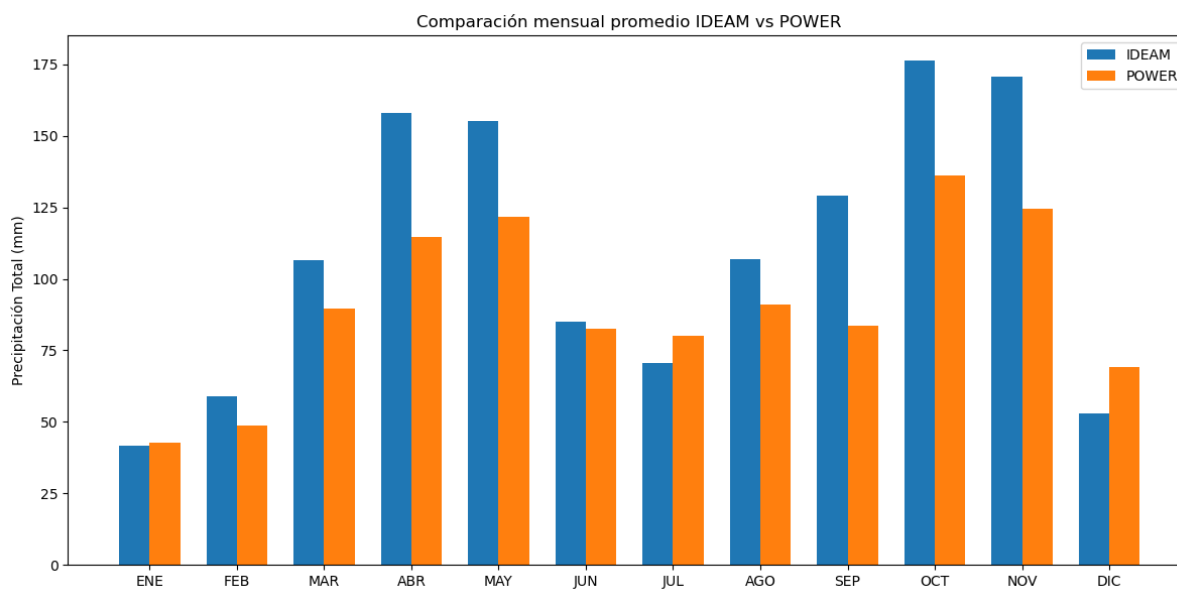
Análisis comparativo datos IDEAM vs datos Giovanni diario estación Provincia.



Nota. Fuente elaboración propia.

Anexo HHH

Análisis comparativo datos IDEAM vs datos Power Data diario estación Provincia.



Nota. Fuente elaboración propia.

Anexo 13: Resultado análisis SARIMAX

Anexo III

Análisis de modelo estadístico de SAIMAX Giovanni vs IDEAM estación Caobo..

SARIMAX Results						
=====						
Dep. Variable:	Precipitation	GIOVANNI	No. Observations:	5844		
Model:	SARIMAX(1, 1, 1)x(1, 1, 1, 12)		Log Likelihood	-22002.204		
Date:	Tue, 20 May 2025		AIC	44014.408		
Time:	12:29:32		BIC	44047.763		
Sample:	01-01-2004		HQIC	44026.008		
	- 12-31-2019					
Covariance Type:	opg					
=====						
	coef	std err	z	P> z	[0.025	0.975]

ar.L1	0.0067	0.013	0.512	0.609	-0.019	0.032
ma.L1	-0.9399	0.005	-200.513	0.000	-0.949	-0.931
ar.S.L12	0.0141	0.012	1.151	0.250	-0.010	0.038
ma.S.L12	-0.9999	0.079	-12.595	0.000	-1.155	-0.844
sigma2	109.4634	8.583	12.754	0.000	92.642	126.285
=====						
Ljung-Box (L1) (Q):	0.00	Jarque-Bera (JB):	72243.71			
Prob(Q):	0.96	Prob(JB):	0.00			
Heteroskedasticity (H):	1.18	Skew:	3.18			
Prob(H) (two-sided):	0.00	Kurtosis:	19.03			
=====						
Warnings:						
[1] Covariance matrix calculated using the outer product of gradients (complex-step).						

Nota. Fuente elaboración propia.

Anexo JJJ

Análisis de modelo estadístico de SAIMAX PoweData vs IDEAM estación Caobo..

SARIMAX Results						
Dep. Variable:	Precipitation	POWER Data	No. Observations:	7671		
Model:	SARIMAX(1, 1, 1)x(1, 1, 1, 12)		Log Likelihood	-23329.593		
Date:	Tue, 20 May 2025		AIC	46669.186		
Time:	12:28:50		BIC	46703.903		
Sample:	01-01-2004		HQIC	46681.095		
	- 12-31-2024					
Covariance Type:	opg					
	coef	std err	z	P> z	[0.025	0.975]
ar.L1	0.2847	0.006	45.075	0.000	0.272	0.297
ma.L1	-0.9824	0.002	-467.883	0.000	-0.986	-0.978
ar.S.L12	0.0012	0.012	0.099	0.921	-0.022	0.024
ma.S.L12	-0.9981	0.003	-305.898	0.000	-1.005	-0.992
sigma2	25.6762	0.109	236.288	0.000	25.463	25.889
Ljung-Box (L1) (Q):	0.14	Jarque-Bera (JB):	722360.91			
Prob(Q):	0.70	Prob(JB):	0.00			
Heteroskedasticity (H):	0.80	Skew:	4.61			
Prob(H) (two-sided):	0.00	Kurtosis:	49.68			
Warnings:						
[1] Covariance matrix calculated using the outer product of gradients (complex-step).						

Nota. Fuente elaboración propia.

Anexo KKK

Análisis de modelo estadístico de SAIMAX Giovanni vs IDEAM estación Escuela Agrícola Cachirá..

SARIMAX Results						
Dep. Variable:	Precipitacion GIOVANNI	No. Observations:	5844			
Model:	SARIMAX(1, 1, 1)x(1, 1, 1, 12)	Log Likelihood	-19415.012			
Date:	Tue, 20 May 2025	AIC	38840.024			
Time:	12:27:13	BIC	38873.378			
Sample:	01-01-2004	HQIC	38851.623			
	- 12-31-2019					
Covariance Type:	opg					
	coef	std err	z	P> z	[0.025	0.975]
ar.L1	0.1391	0.011	12.714	0.000	0.118	0.160
ma.L1	-0.9348	0.005	-193.554	0.000	-0.944	-0.925
ar.S.L12	-0.0160	0.013	-1.274	0.202	-0.041	0.009
ma.S.L12	-0.9990	0.009	-112.702	0.000	-1.016	-0.982
sigma2	45.1063	0.402	112.253	0.000	44.319	45.894
Ljung-Box (L1) (Q):	0.26	Jarque-Bera (JB):	61540.21			
Prob(Q):	0.61	Prob(JB):	0.00			
Heteroskedasticity (H):	1.41	Skew:	2.95			
Prob(H) (two-sided):	0.00	Kurtosis:	17.78			
Warnings:						
[1] Covariance matrix calculated using the outer product of gradients (complex-step).						

Nota. Fuente elaboración propia.

Anexo LLL

Análisis de modelo estadístico de SAIMAX PoweData vs IDEAM estación Escuela Agrícola Cachirá..

SARIMAX Results						
Dep. Variable:	Precipitation	POWER Data	No. Observations:	7671		
Model:	SARIMAX(1, 1, 1)x(1, 1, 1, 12)		Log Likelihood	-23329.593		
Date:	Tue, 20 May 2025		AIC	46669.186		
Time:	12:26:44		BIC	46703.903		
Sample:	01-01-2004		HQIC	46681.095		
	- 12-31-2024					
Covariance Type:	opg					
	coef	std err	z	P> z	[0.025	0.975]
ar.L1	0.2847	0.006	45.075	0.000	0.272	0.297
ma.L1	-0.9824	0.002	-467.883	0.000	-0.986	-0.978
ar.S.L12	0.0012	0.012	0.099	0.921	-0.022	0.024
ma.S.L12	-0.9981	0.003	-305.898	0.000	-1.005	-0.992
sigma2	25.6762	0.109	236.288	0.000	25.463	25.889
Ljung-Box (L1) (Q):	0.14	Jarque-Bera (JB):	722360.91			
Prob(Q):	0.70	Prob(JB):	0.00			
Heteroskedasticity (H):	0.80	Skew:	4.61			
Prob(H) (two-sided):	0.00	Kurtosis:	49.68			
Warnings:						
[1] Covariance matrix calculated using the outer product of gradients (complex-step).						

Nota. Fuente elaboración propia.

Anexo MMM

Análisis de modelo estadístico de SAIMAX Giovanni vs IDEAM estación Porvenir..

SARIMAX Results						
Dep. Variable:	Precipitacion GIOVANNI	No. Observations:	5844			
Model:	SARIMAX(1, 1, 1)x(1, 1, 1, 12)	Log Likelihood	-22343.771			
Date:	Tue, 20 May 2025	AIC	44697.542			
Time:	12:23:30	BIC	44730.897			
Sample:	01-01-2004	HQIC	44709.142			
	- 12-31-2019					
Covariance Type:	opg					
	coef	std err	z	P> z	[0.025	0.975]
ar.L1	-0.0499	0.013	-3.746	0.000	-0.076	-0.024
ma.L1	-0.9373	0.004	-212.938	0.000	-0.946	-0.929
ar.S.L12	-0.0122	0.013	-0.971	0.331	-0.037	0.012
ma.S.L12	-0.9998	0.051	-19.550	0.000	-1.100	-0.900
sigma2	123.0724	6.115	20.126	0.000	111.087	135.058
Ljung-Box (L1) (Q):	0.01	Jarque-Bera (JB):	66792.44			
Prob(Q):	0.91	Prob(JB):	0.00			
Heteroskedasticity (H):	1.39	Skew:	3.19			
Prob(H) (two-sided):	0.00	Kurtosis:	18.30			
Warnings:						
[1] Covariance matrix calculated using the outer product of gradients (complex-step).						

Nota. Fuente elaboración propia.

Anexo NNN

Análisis de modelo estadístico de SAIMAX PoweData vs IDEAM estación Porvenir..

SARIMAX Results						
Dep. Variable:	Precipitacion POWER Data	No. Observations:	7671			
Model:	SARIMAX(1, 1, 1)x(1, 1, 1, 12)	Log Likelihood	-25472.623			
Date:	Tue, 20 May 2025	AIC	50955.245			
Time:	12:22:51	BIC	50989.963			
Sample:	01-01-2004	HQIC	50967.155			
	- 12-31-2024					
Covariance Type:	opg					
	coef	std err	z	P> z	[0.025	0.975]
ar.L1	0.2995	0.007	43.117	0.000	0.286	0.313
ma.L1	-0.9753	0.003	-389.262	0.000	-0.980	-0.970
ar.S.L12	0.0035	0.011	0.328	0.743	-0.018	0.025
ma.S.L12	-0.9997	0.015	-64.750	0.000	-1.030	-0.969
sigma2	44.8901	0.599	74.966	0.000	43.716	46.064
Ljung-Box (L1) (Q):	2.54	Jarque-Bera (JB):	600823.80			
Prob(Q):	0.11	Prob(JB):	0.00			
Heteroskedasticity (H):	1.39	Skew:	4.26			
Prob(H) (two-sided):	0.00	Kurtosis:	45.55			
Warnings:						
[1] Covariance matrix calculated using the outer product of gradients (complex-step).						

Nota. Fuente elaboración propia.

Anexo OOO

Análisis de modelo estadístico de SAIMAX Giovanni vs IDEAM estación Provincia..

SARIMAX Results						
Dep. Variable:	Precipitation GIOVANNI	No. Observations:	5844			
Model:	SARIMAX(1, 1, 1)x(1, 1, 1, 12)	Log Likelihood	-22343.771			
Date:	Tue, 20 May 2025	AIC	44697.542			
Time:	12:24:54	BIC	44730.897			
Sample:	01-01-2004	HQIC	44709.142			
	- 12-31-2019					
Covariance Type:	opg					
	coef	std err	z	P> z	[0.025	0.975]
ar.L1	-0.0499	0.013	-3.746	0.000	-0.076	-0.024
ma.L1	-0.9373	0.004	-212.938	0.000	-0.946	-0.929
ar.S.L12	-0.0122	0.013	-0.971	0.331	-0.037	0.012
ma.S.L12	-0.9998	0.051	-19.550	0.000	-1.100	-0.900
sigma2	123.0724	6.115	20.126	0.000	111.087	135.058
Ljung-Box (L1) (Q):	0.01	Jarque-Bera (JB):	66792.44			
Prob(Q):	0.91	Prob(JB):	0.00			
Heteroskedasticity (H):	1.39	Skew:	3.19			
Prob(H) (two-sided):	0.00	Kurtosis:	18.30			
Warnings:						
[1] Covariance matrix calculated using the outer product of gradients (complex-step).						

Nota. Fuente elaboración propia.

Anexo PPP

Análisis de modelo estadístico de SAIMAX PoweData vs IDEAM estación Provincia..

SARIMAX Results						
Dep. Variable:	Precipitation POWER Data	No. Observations:	7671			
Model:	SARIMAX(1, 1, 1)x(1, 1, 1, 12)	Log Likelihood	-25472.623			
Date:	Tue, 20 May 2025	AIC	50955.245			
Time:	12:24:15	BIC	50989.963			
Sample:	01-01-2004	HQIC	50967.155			
	- 12-31-2024					
Covariance Type:	opg					
	coef	std err	z	P> z	[0.025	0.975]
ar.L1	0.2995	0.007	43.117	0.000	0.286	0.313
ma.L1	-0.9753	0.003	-389.262	0.000	-0.980	-0.970
ar.S.L12	0.0035	0.011	0.328	0.743	-0.018	0.025
ma.S.L12	-0.9997	0.015	-64.750	0.000	-1.030	-0.969
sigma2	44.8901	0.599	74.966	0.000	43.716	46.064
Ljung-Box (L1) (Q):	2.54	Jarque-Bera (JB):	600823.80			
Prob(Q):	0.11	Prob(JB):	0.00			
Heteroskedasticity (H):	1.39	Skew:	4.26			
Prob(H) (two-sided):	0.00	Kurtosis:	45.55			
Warnings:						
[1] Covariance matrix calculated using the outer product of gradients (complex-step).						

Nota. Fuente elaboración propia.

Anexo QQQ

Análisis de modelo estadístico de SAIMAX Giovanni vs IDEAM estación San Rafael..

SARIMAX Results						
Dep. Variable:	Precipitacion GIOVANNI	No. Observations:	5844			
Model:	SARIMAX(1, 1, 1)x(1, 1, 1, 12)	Log Likelihood	-23489.095			
Date:	Tue, 20 May 2025	AIC	46988.189			
Time:	12:26:09	BIC	47021.544			
Sample:	01-01-2004	HQIC	46999.789			
	- 12-31-2019					
Covariance Type:	opg					
	coef	std err	z	P> z	[0.025	0.975]
ar.L1	-0.0404	0.013	-3.100	0.002	-0.066	-0.015
ma.L1	-0.9453	0.004	-214.720	0.000	-0.954	-0.937
ar.S.L12	0.0015	0.013	0.118	0.906	-0.024	0.027
ma.S.L12	-0.9959	0.003	-385.950	0.000	-1.001	-0.991
sigma2	182.7741	1.338	136.644	0.000	180.152	185.396
Ljung-Box (L1) (Q):	0.00	Jarque-Bera (JB):	40488.84			
Prob(Q):	0.95	Prob(JB):	0.00			
Heteroskedasticity (H):	1.22	Skew:	2.82			
Prob(H) (two-sided):	0.00	Kurtosis:	14.61			
Warnings:						
[1] Covariance matrix calculated using the outer product of gradients (complex-step).						

Nota. Fuente elaboración propia.

Anexo RRR

Análisis de modelo estadístico de SAIMAX PoweData vs IDEAM estación San Rafael..

SARIMAX Results						
Dep. Variable:	Precipitacion POWER Data	No. Observations:	7671			
Model:	SARIMAX(1, 1, 1)x(1, 1, 12)	Log Likelihood	-25472.623			
Date:	Tue, 20 May 2025	AIC	50955.245			
Time:	12:25:41	BIC	50989.963			
Sample:	01-01-2004	HQIC	50967.155			
	- 12-31-2024					
Covariance Type:	opg					
	coef	std err	z	P> z	[0.025	0.975]
ar.L1	0.2995	0.007	43.117	0.000	0.286	0.313
ma.L1	-0.9753	0.003	-389.262	0.000	-0.980	-0.970
ar.S.L12	0.0035	0.011	0.328	0.743	-0.018	0.025
ma.S.L12	-0.9997	0.015	-64.750	0.000	-1.030	-0.969
sigma2	44.8901	0.599	74.966	0.000	43.716	46.064
Ljung-Box (L1) (Q):	2.54	Jarque-Bera (JB):	600823.80			
Prob(Q):	0.11	Prob(JB):	0.00			
Heteroskedasticity (H):	1.39	Skew:	4.26			
Prob(H) (two-sided):	0.00	Kurtosis:	45.55			
Warnings:						
[1] Covariance matrix calculated using the outer product of gradients (complex-step).						

Nota. Fuente elaboración propia.

Anexo SSS

Análisis de modelo estadístico de SAIMAX Giovanni vs IDEAM estación Vega.

```

=====
SARIMAX Results
=====
Dep. Variable:      Precipitation GIOVANNI    No. Observations:      5844
Model:             SARIMAX(1, 1, 1)x(1, 1, 1, 12)  Log Likelihood         -19415.012
Date:              Tue, 20 May 2025            AIC                   38840.024
Time:              12:28:16                   BIC                   38873.378
Sample:            01-01-2004                 HQIC                  38851.623
                  - 12-31-2019
Covariance Type:    opg
=====
              coef      std err          z      P>|z|      [0.025      0.975]
-----
ar.L1          0.1391      0.011      12.714      0.000      0.118      0.160
ma.L1         -0.9348      0.005     -193.554      0.000     -0.944     -0.925
ar.S.L12       -0.0160      0.013      -1.274      0.202     -0.041      0.009
ma.S.L12       -0.9990      0.009     -112.702      0.000     -1.016     -0.982
sigma2         45.1063      0.402     112.253      0.000     44.319     45.894
=====
Ljung-Box (L1) (Q):          0.26  Jarque-Bera (JB):          61540.21
Prob(Q):                    0.61  Prob(JB):              0.00
Heteroskedasticity (H):      1.41  Skew:              2.95
Prob(H) (two-sided):         0.00  Kurtosis:          17.78
=====

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).|

```

Nota. Fuente elaboración propia.

Anexo TTT

Análisis de modelo estadístico de SAIMAX PoweData vs IDEAM estación Vega.

```

=====
SARIMAX Results
=====
Dep. Variable:      Precipitation POWER Data    No. Observations:      7671
Model:             SARIMAX(1, 1, 1)x(1, 1, 1, 12)  Log Likelihood         -23329.593
Date:              Tue, 20 May 2025            AIC                   46669.186
Time:              12:27:49                   BIC                   46703.903
Sample:            01-01-2004                 HQIC                  46681.095
                  - 12-31-2024
Covariance Type:    opg
=====
              coef      std err          z      P>|z|      [0.025      0.975]
-----
ar.L1          0.2847      0.006      45.075      0.000      0.272      0.297
ma.L1         -0.9824      0.002     -467.883      0.000     -0.986     -0.978
ar.S.L12       0.0012      0.012      0.099      0.921     -0.022      0.024
ma.S.L12       -0.9981      0.003     -305.898      0.000     -1.005     -0.992
sigma2         25.6762      0.109     236.288      0.000     25.463     25.889
=====
Ljung-Box (L1) (Q):          0.14  Jarque-Bera (JB):          722360.91
Prob(Q):          0.70  Prob(JB):              0.00
Heteroskedasticity (H):      0.80  Skew:              4.61
Prob(H) (two-sided):         0.00  Kurtosis:          49.68
=====

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).|

```

Nota. Fuente elaboración propia.