

UNIVERSIDAD NACIONAL AUTÓNOMA DE CHOTA



FACULTAD DE CIENCIAS DE LA INGENIERÍA

Escuela Profesional de Ingeniería Civil

**“Regionalización de precipitaciones máximas diarias anuales en la
cuenca del río chamaya, aplicando el método L-momentos, 2024”**

Tesis para optar el título de Ingeniero Civil

Línea de investigación:

Diseño y gestión de sistemas hidráulicos

Autor(es):

Bach. Robbie Williams Ocsa Cubas

Bach. Jhensin Rivera Guamuro

Asesor:

Dr. Ing. Luis Alberto Orbegoso Navarro

Chota – Perú

2026



UNIVERSIDAD NACIONAL AUTÓNOMA DE CHOTA
FACULTAD DE CIENCIAS DE LA INGENIERÍA



CARTA N° 14-2026-UNACH/UI/MAST

A : **Mg. Luis Angel Mozo Cruz**
Presidente del jurado de tesis FCI-UNACH.

ASUNTO : Constancia de Originalidad de Turnitin de Robbie Williams Ocsa Cubas y Jhensin Rivera Guamuro.

FECHA : Colpa Matara, 25 de junio del 2026.

REFERENCIA: i) Carta N 22-2026-UNACH-LAMC / Envío de tesis en digital en su versión final
ii) Carta N 21-2026-UNACH-LAMC / Conformidad por parte del presidente del jurado de tesis.
iii) CARTA N° 0013 – 2026 – DGCD – DO – UNACH / Conformidad por parte del secretario del jurado de tesis.
iv) INFORME N° 019-2026-UNACH/MGHT. / Conformidad por parte del vocal del jurado de tesis

De mi mayor consideración

Es grato dirigirme a Usted para expresar mi saludo, y a la vez alcanzar la constancia de originalidad de TURNITIN de la tesis denominada: **“REGIONALIZACIÓN DE PRECIPITACIONES MÁXIMAS DIARIAS ANUALES EN LA CUENCA DEL RÍO CHAMAYA, APLICANDO EL MÉTODO L-MOMENTOS, 2024”**, elaborado por los bachilleres en ingeniería civil **Robbie Williams Ocsa Cubas y Jhensin Rivera Guamuro**, para continuar con sus trámites ante la UNACH.

Sin otro particular, es propicia la ocasión para reiterarle las muestras de mi distinguida consideración y estima.

Atentamente,



Miguel Ángel SILVA TARRILLO
INGENIERO CIVIL
N° 187752

Ing. Miguel Ángel Silva Tarrillo
Director de la unidad de
investigación FCI-UNACH

CC.

Archivo

Adjunto:

- 1) Constancia de originalidad N° 14-2026.
- 2) Reporte TURNITIN



Colpa Matara, 25 de junio del 2026.

C.O. N° 14-2026-UI-EPIC

CONSTANCIA DE ORIGINALIDAD Y USO DE IA

El que suscribe, Director de la Unidad de Investigación de la Facultad de Ciencias de la Ingeniería de la Universidad Nacional Autónoma de Chota, hace constar que el Informe Final de Tesis titulado: **“REGIONALIZACIÓN DE PRECIPITACIONES MÁXIMAS DIARIAS ANUALES EN LA CUENCA DEL RÍO CHAMAYA, APLICANDO EL MÉTODO L-MOMENTOS, 2024”**, elaborado por los bachilleres en ingeniería civil: **Robbie Williams Ocsa Cubas y Jhensin Rivera Guamuro**, para optar el Título Profesional de ingeniero civil, presenta un índice de similitud de 8% excluyendo texto citado, bibliografía y fuentes que tengan coincidencias de menos de 10 palabras, no teniendo un porcentaje de escritura con inteligencia artificial; por lo tanto, cumple con los criterios de evaluación de originalidad establecidos en el acápite g) del artículo 20 del Reglamento de Grados y Títulos UNACH, aprobado mediante la Resolución C.O. N° 120-2022-UNACH con fecha de 03 de marzo de 2022.

Se expide la presente, en conformidad a la directiva antes mencionada, para los fines que estime pertinentes.

Miguel Ángel SILVA TARRILLO
INGENIERO CIVIL

Ing. Miguel Ángel Silva Tarrillo
Director de la unidad de
investigación FCI-UNACH

Robbie Williams Ocsa Cubas y Jhensin Rivera Gua...

REGIONALIZACIÓN DE PRECIPITACIONES MÁXIMAS DIARIAS ANUALES EN LA CUENCA DEL RÍO CHAMAYA, APLIC...

 Informe Final de Tesis
 UNIDAD DE INVESTIGACION FIC 2026
 Universidad Nacional Autónoma de Chota

Detalles del documento

Identificador de la entrega

trn:oid:::1:3601507856

Fecha de entrega

25 jun 2026, 11:02 a.m. GMT-5

Fecha de descarga

25 jun 2026, 11:08 a.m. GMT-5

Nombre del archivo

INFORME_FINAL_DE_TESIS_Robbie_Jhensin_02.06.26_TNT.pdf

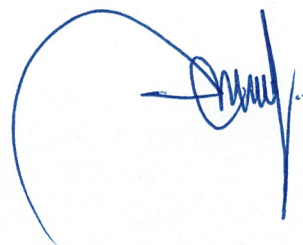
Tamaño del archivo

7.7 MB

200 páginas

49.112 palabras

250.827 caracteres



8% Similitud general

El total combinado de todas las coincidencias, incluidas las fuentes superpuestas, para ca...

Filtrado desde el informe

- Bibliografía
- Texto mencionado
- Coincidencias menores (menos de 10 palabras)

Fuentes principales

- 7%  Fuentes de Internet
- 1%  Publicaciones
- 2%  Trabajos entregados (trabajos del estudiante)

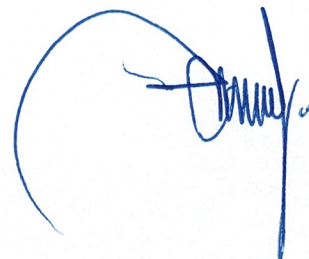
Marcas de integridad

N.º de alerta de integridad para revisión

-  **Texto oculto**
6 caracteres sospechosos en N.º de páginas
El texto es alterado para mezclarse con el fondo blanco del documento.

Los algoritmos de nuestro sistema analizan un documento en profundidad para buscar inconsistencias que permitirían distinguirlo de una entrega normal. Si advertimos algo extraño, lo marcamos como una alerta para que pueda revisarlo.

Una marca de alerta no es necesariamente un indicador de problemas. Sin embargo, recomendamos que preste atención y la revise.



Fuentes principales

- 7% Fuentes de Internet
- 1% Publicaciones
- 2% Trabajos entregados (trabajos del estudiante)

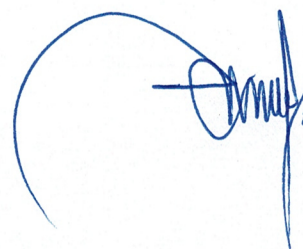
Fuentes principales

Las fuentes con el mayor número de coincidencias dentro de la entrega. Las fuentes superpuestas no se mostrarán.

1	Internet	cybertesis.uni.edu.pe	3%
2	Internet	repositorio.unach.edu.pe	2%
3	Internet	portaluni.unach.edu.pe	<1%
4	Internet	hdl.handle.net	<1%
5	Internet	repositorio.unc.edu.pe	<1%
6	Trabajos del estudiante	Universidad Alas Peruanas	<1%
7	Internet	repositorio.lamolina.edu.pe	<1%
8	Internet	uvadoc.uva.es	<1%
9	Internet	www.editorasasagai.org.ar	<1%
10	Internet	www.sri.gov.ec	<1%
11	Trabajos del estudiante	Universidad Pontificia Bolivariana	<1%

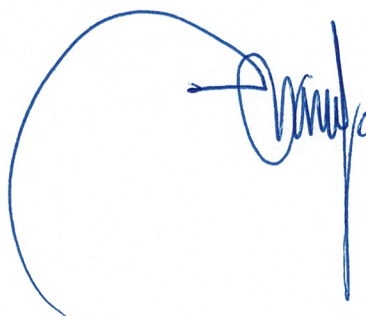
12	Internet	api-repositorio.unapikitos.edu.pe	<1%
13	Trabajos del estudiante	Universidad de Costa Rica	<1%
14	Publicación	FC INGENIERIA Y SERVICIOS AMBIENTALES SOCIEDAD ANONIMA CERRADA. "MEIA...	<1%
15	Trabajos del estudiante	Universidad TecMilenio	<1%
16	Internet	www.coursehero.com	<1%
17	Trabajos del estudiante	Account Universidad Mariana	<1%
18	Internet	senamhi.gob.pe	<1%
19	Internet	doaj.org	<1%
20	Trabajos del estudiante	Universidad Nacional Autonoma de Chota	<1%
21	Internet	dspace.ups.edu.ec	<1%
22	Trabajos del estudiante	Organismo de Evaluación y Fiscalización	<1%
23	Internet	repositorio.ucv.edu.pe	<1%
24	Trabajos del estudiante	Universidad Peruana Los Andes	<1%
25	Internet	repositorio.uancv.edu.pe	<1%

26	Trabajos del estudiante Universidad Cesar Vallejo	<1%
27	Internet cdn.www.gob.pe	<1%
28	Trabajos del estudiante Pontificia Universidad Catolica del Peru	<1%
29	Trabajos del estudiante Universidad Andina Nestor Caceres Velasquez	<1%
30	Internet repositorio.deusto.es	<1%
31	Internet www.redalyc.org	<1%
32	Trabajos del estudiante Universidad Peruana de Ciencias Aplicadas	<1%
33	Trabajos del estudiante Universiti Malaysia Sabah	<1%
34	Internet conservancy.umn.edu	<1%
35	Internet soutr.org	<1%
36	Trabajos del estudiante Universidad de Piura	<1%
37	Publicación Zhu, Laiyin, Steven M. Quiring, and Kerry A. Emanuel. "Estimating tropical cyclon..."	<1%
38	Publicación Zelio Cahuapaza, Bibyan Jesus. "El sistema de control interno y el control interno ..."	<1%
39	Internet repositorio.esan.edu.pe	<1%



40	Internet	repositorio.ugto.mx	<1%
41	Internet	repositorio.uroosevelt.edu.pe	<1%
42	Trabajos del estudiante	uncedu	<1%
43	Trabajos del estudiante	uni	<1%
44	Trabajos del estudiante	upb	<1%
45	Publicación	FC INGENIERIA Y SERVICIOS AMBIENTALES SOCIEDAD ANONIMA CERRADA. "PAM...	<1%
46	Trabajos del estudiante	Universidad Privada del Norte	<1%
47	Internet	dspace.unitru.edu.pe	<1%
48	Internet	repositorio.unsch.edu.pe	<1%
49	Publicación	Eduardo Luis Flores-Quispe, Eduardo Flores-Condori, Samuel Huaquisto-Cáceres. ...	<1%
50	Internet	alicia.concytec.gob.pe	<1%
51	Internet	ebuah.uah.es	<1%
52	Internet	repositorio.ana.gob.pe	<1%
53	Internet	repositorio.senamhi.gob.pe	<1%

54	Internet	repositorio.uas.edu.mx	<1%
55	Internet	repositorio.untumbes.edu.pe	<1%
56	Internet	upc.aws.openrepository.com	<1%
57	Internet	www.aranzadi.eus	<1%
58	Internet	www.ptolomeo.unam.mx:8080	<1%
59	Internet	www.repositorio.unach.edu.pe	<1%
60	Internet	www.researchgate.net	<1%



**REGIONALIZACIÓN DE PRECIPITACIONES MÁXIMAS
DIARIAS ANUALES EN LA CUENCA DEL RÍO CHAMAYA,
APLICANDO EL MÉTODO L-MOMENTOS, 2024.**

POR:

Bach. ROBBIE WILLIAMS OCSA CUBAS

Bach. JHENSIN RIVERA GUAMURO

**Presentada a la Facultad de Ciencias de la Ingeniería de la
Universidad Nacional Autónoma de Chota para optar el título de
INGENIERO CIVIL**

APROBADA POR EL JURADO INTEGRADO POR

Mg. Ing. Luis Angel Mozo Cruz

PRESIDENTE

Mg. Ing. Donald Gorki Collantes Delgado

SECRETARIO

Dra. Carmen Rosa Cárdenas Rosales

ACCESITARIO



UNIVERSIDAD NACIONAL AUTÓNOMA DE CHOTA

Ley de Creación N° 29531

LICENCIADA CON RESOLUCIÓN DE CONSEJO DIRECTIVO N° 160-2018-SUNEDU/CD

DECANATO DE LA FACULTAD DE CIENCIAS DE LA INGENIERÍA

ESCUELA PROFESIONAL DE INGENIERIA CIVIL



ACTA DE SUSTENTACIÓN DE TESIS

Siendo las 09:00 a.m. del día 21 de julio, reunidos en la sala de docentes de la Escuela Profesional de Ingeniería Civil, los miembros del jurado de tesis que suscriben, para escuchar y evaluar la sustentación de tesis presentado por los bachilleres: : **Robbie Williams Ocsa Cubas y Jhensin Rivera Guamuro**, denominada: **“Regionalización de precipitaciones máximas diarias anuales en la cuenca del río Chamaya, aplicando el método L-momentos, 2024”**; escuchada la sustentación, y absueltas las preguntas a las observaciones formuladas, la declaramos:

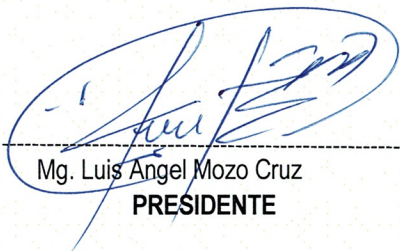
APROBADO

CON EL CALIFICATIVO (*)

17 (DIECISIETE)

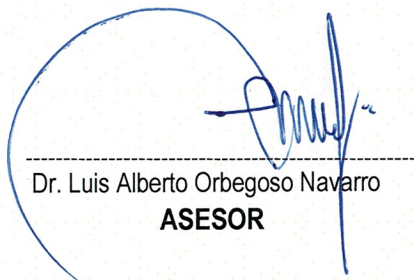
En consecuencia, se le declara **EXPEDITOS** para conferirle el Título de Ingeniero civil, elevando la presente acta al decano de la Facultad de Ciencias de la Ingeniería a fin de que se emita el acto resolutivo, en conformidad con la ley universitaria y el estatuto de la Universidad.

Chota, 21 de julio del 2026.


Mg. Luis Angel Mozo Cruz
PRESIDENTE


Mg. Donald Gorki Collantes Delgado
SECRETARIO


Dra. Carmen Rosa Cardenas Rosales
ACCESITARIO


Dr. Luis Alberto Orbegoso Navarro
ASESOR

(*) De acuerdo al reglamento específico del proyecto y tesis de investigación de la EPIC, aprobada con Resolución de coordinación N° 141-2020, Artículo 21, cuya calificación es: (20 Summa Cum Laude); (18-19: Aprobado con excelencia); (15-17: Aprobado con mención honrosa); (12-14: Aprobado); (0-11: Desaprobado).



UNIVERSIDAD NACIONAL AUTÓNOMA DE CHOTA

Ley de Creación N° 29531

LICENCIADA CON RESOLUCIÓN DE CONSEJO DIRECTIVO N° 160-2018-SUNEDU/CD
DECANATO DE LA FACULTAD DE CIENCIAS DE LA INGENIERÍA
ESCUELA PROFESIONAL DE INGENIERIA CIVIL



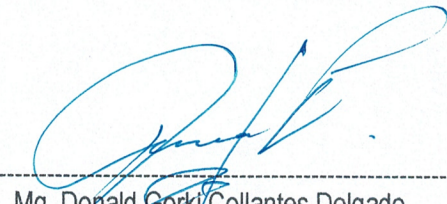
ACTA DE CONFORMIDAD DE TESIS

Los miembros del jurado, luego de evaluar la Tesis denominada: **“Regionalización de precipitaciones máximas diarias anuales en la cuenca del río Chamaya, aplicando el método L-momentos, 2024”** presentado por los bachilleres **Robbie Williams Ocsa Cubas y Jhensin Rivera Guamuro**, sustentada el día 21 de julio del 2026, según Resolución de Decanato N°001-2026-FCI/UNACH, la declaramos **CONFORME**.

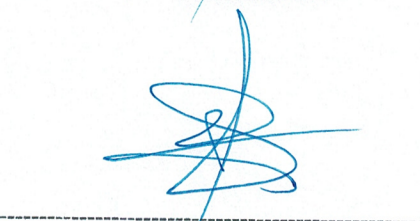
Chota, 24 de julio de 2026.



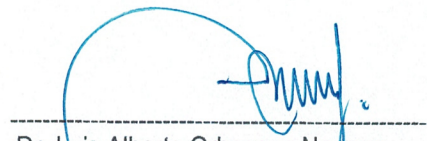
Mg. Luis Angel Mozo Cruz
PRÉSIDENTE



Mg. Donald Gorki Collantes Delgado
SECRETARIO



Dra. Carmen Rosa Cardenas Rosales
ACCESITARIO



Dr. Luis Alberto Orbegoso Navarro
ASESOR

DEDICATORIA

Dedico este trabajo, en primer lugar, a Dios, por concederme salud, fortaleza y claridad para culminar esta investigación. A mi familia, por su apoyo constante, comprensión y confianza durante cada etapa del proceso académico. Asimismo, a mis amigos, por su compañía, aliento y respaldo en los momentos de mayor exigencia, que fueron fundamentales para alcanzar este logro.

Bach. Robbie Williams Ocsa Cubas

Dedico esta investigación a Dios por guiar mi camino y brindarme perseverancia para alcanzar este objetivo académico. A mi familia, por su apoyo incondicional, paciencia y motivación permanente, que me permitieron continuar aun en las dificultades. De igual manera, a mis amigos, por su ayuda, palabras de ánimo y acompañamiento, que contribuyeron de manera significativa a la culminación de este trabajo.

Bach. Jhensin Rivera Guamuro

AGRADECIMIENTOS

Agradecemos a Dios por brindarnos salud, constancia y la oportunidad de culminar esta investigación. Expresamos nuestro reconocimiento y gratitud a nuestro asesor, Dr. Ing. Luis Alberto Orbegoso Navarro, por su orientación técnica, exigencia académica y acompañamiento durante el desarrollo del estudio. Asimismo, agradecemos al Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI) por facilitar la información pluviométrica utilizada, la cual resultó esencial para el análisis y cumplimiento de los objetivos. Finalmente, extendemos nuestro agradecimiento a todas las personas e instituciones que, de manera directa o indirecta, contribuyeron con sugerencias, apoyo y facilidades para la realización de este trabajo.

ÍNDICE DE CONTENIDOS

LISTA DE SIGLAS Y ABREVIATURAS	xiv
RESUMEN.....	xvi
ABSTRACT	xvii
CAPÍTULO I: INTRODUCCIÓN	18
1.1. Planteamiento del problema.....	18
1.2. Justificación	21
1.3. Formulación del problema.....	22
1.4. Objetivos.....	22
1.4.1. Objetivo general	22
1.4.2. Objetivos específicos	22
1.5. Limitaciones	23
CAPÍTULO II: MARCO TEÓRICO.....	24
2.1. Antecedentes	24
2.1.1. Antecedentes internacionales	24
2.1.2. Antecedentes nacionales	30
2.1.3. Antecedentes regionales.....	32
2.2. Bases teóricas	33
2.2.1. Análisis Regional de Frecuencia (ARF)	33
2.2.2. Teoría de L-Momentos.....	34
2.2.3. Delimitación e identificación de regiones homogéneas.....	41
2.2.4. Análisis de datos: Pruebas de discordancia y heterogeneidad regional (ARF).....	45
2.2.5. Selección y ajuste de distribuciones de probabilidad.....	55
2.2.6. Región de Influencia (ROI).....	57
2.3. Definiciones conceptuales	58
2.3.1. Cuenca hidrográfica y sus propiedades geomorfológicas.....	58
2.3.2. Precipitaciones	61
2.3.3. Precipitaciones máximas diarias anuales (PMDA) observadas	62

2.3.4. Regionalización.....	66
2.3.5. Regionalización PMDA por L-momentos	67
CAPÍTULO III: MARCO METODOLÓGICO	74
3.1. Tipo, diseño, nivel y enfoque de investigación	74
3.2. Ubicación, población, muestra, muestreo y unidad de análisis.....	76
3.2.1. Ubicación de la zona de estudio.....	76
3.2.2. Población.....	78
3.2.3. Muestreo.....	78
3.2.4. Muestra.....	78
3.2.5. Unidad de análisis	78
3.3. Técnicas e instrumentos de recolección de datos.....	79
3.3.1. Técnicas de recolección de datos	79
3.3.2. Instrumentos de recolección de datos	79
3.4. Hipótesis	80
3.5. Operacionalización de variables	80
3.5.1. Variable independiente: Precipitaciones máximas diarias anuales (PMDA).....	80
3.5.2. Variable dependiente: Regionalización PMDA por L-momentos	81
3.5.3. Matriz de operacionalización de variables	82
3.6. Procedimientos de recolección de datos	83
3.6.1. Delimitación del área de estudio y base cartográfica.....	83
3.6.2. Identificación, inventario y selección de estaciones meteorológicas.....	88
3.6.3. Obtención de información pluviométrica y construcción de series de PMDA.....	91
3.6.4. Control de calidad y depuración preliminar de las series (análisis exploratorio)	92
3.6.5. Preparación de atributos y matriz de trabajo para regionalización	94
3.6.6. Regionalización mediante L-momentos con Análisis Regional de Frecuencia (ARF).....	95
3.6.7. Regionalización mediante L-momentos por Región de Influencia (ROI/ RI)	103
3.6.8. Sistematización para cartografía regional de PMDA.....	104
3.7. Procedimientos de análisis de datos.....	107
3.8. Material y equipos (descripción del uso).....	108
3.9. Aspectos éticos	109

CAPÍTULO IV: RESULTADOS.....	111
4.1. Descripción de resultados	111
4.1.1. Análisis de las series de PMDA de las estaciones meteorológicas seleccionadas	111
4.1.2. Análisis regional de frecuencia y enfoque de Región de Influencia (ROI)	136
4.1.3. Mapas regionales de PMDA regionalizadas y caracterizadas espacialmente en regiones homogéneas para la cuenca del río Chamaya	174
4.1.4. Estrategias para mitigar la insuficiencia de datos hidrológicos en zonas poco instrumentadas, empleando la regionalización con L-momentos	184
4.2. Contrastación de hipótesis.....	189
4.3. Discusión.....	194
CAPÍTULO V: CONCLUSIONES Y RECOMENDACIONES	203
5.1. Conclusiones.....	203
5.2. Recomendaciones.....	205
REFERENCIAS.....	207
ANEXOS.....	216
Anexo 1. Matriz de consistencia.....	216
Anexo 2. Documentación.....	217
Anexo 3. Resultado del análisis de regionalización.....	218
Anexo 4. Mapas	219

ÍNDICE DE TABLAS

Tabla 1	Valores críticos para la discordancia Di.....	53
Tabla 2	Tamaño de cuenca según su extensión.....	58
Tabla 3	Índice de compacidad según la forma de la cuenca	59
Tabla 4	Periodo de retorno según fuente hidrográfica y proyectos u obras	61
Tabla 5	Tipo de investigación según los principales criterios	74
Tabla 6	Puntos extremos de la cuenca Chamaya	76
Tabla 7	Departamentos de la cuenca.....	76
Tabla 8	Codificación Pfafstetter unidad hidrográfica de la cuenca Chamaya	76
Tabla 9	Matriz de operacionalización de variables.....	82
Tabla 10	Parámetros geomorfológicos e identificadores de la cuenca del río Chamaya	85
Tabla 11	Características del sistema climático	86
Tabla 12	Características del sistema climático de la zona de estudio	87
Tabla 13	Materiales y equipos utilizados para el proceso de obtención de datos	108
Tabla 14	Inventario, ubicación geográfica y fuente de datos de las estaciones meteorológicas empleadas (PMDA)	113
Tabla 15	Estaciones más próximas por estación (cinco menores distancias, km)	114
Tabla 16	Matriz de distancias entre estaciones meteorológicas (km)	116
Tabla 17	Estadística descriptiva de PMDA por estación	118
Tabla 18	Estaciones que necesitan corrección por presentar tendencia en la serie de datos.....	129
Tabla 19	Resultados del análisis de tendencia/estacionariedad en series PMDA (Mann–Kendall, Spearman y tendencia en la media).....	130
Tabla 20	Estaciones que necesitan corrección por presentar saltos en la media.....	132
Tabla 21	Número de estaciones que cumplen y no cumplen las pruebas	132
Tabla 22	Resultados del análisis de homogeneidad de las series PMDA (Mann–Whitney, prueba t y prueba F)	133
Tabla 23	Selección de estaciones con datos corregidos.....	135
Tabla 24	Grupos iniciales se obtuvieron a partir de los métodos de agrupación de Ward y K-means	136
Tabla 25	Máximos valores de discordancia en las regiones preliminares formadas.....	139
Tabla 26	Valores de heterogeneidad en las regiones preliminares formadas.....	139
Tabla 27	Estaciones y sus L-momentos	140
Tabla 28	Regiones (4) y estaciones que las conforman	141
Tabla 29	Rango de altitudes de regiones hidrológicas conformadas	143
Tabla 30	Máximos valores de discordancia en las regiones formadas.....	143

Tabla 31	Valores de heterogeneidad en las regiones formadas	143
Tabla 32	Valores de Z por región según las distribuciones analizadas.....	145
Tabla 33	Distribución de frecuencia óptima para cada región hidrológica.....	146
Tabla 34	Cuantiles regionales adimensionales para la distribución con mejor ajuste de las regiones formadas.....	147
Tabla 35	Cuantiles locales de PMDA estimados mediante el ajuste probabilístico seleccionado para cada estación de la Región 1	150
Tabla 36	Cuantiles locales de PMDA estimados mediante el ajuste probabilístico seleccionado para cada estación de la Región 2	151
Tabla 37	Cuantiles locales de PMDA estimados mediante el ajuste probabilístico seleccionado para cada estación de la Región 3.....	152
Tabla 38	Cuantiles locales de PMDA estimados mediante el ajuste probabilístico seleccionado para cada estación de la Región 4.....	153
Tabla 39	Estaciones agrupadas para diferentes periodos de retorno.....	158
Tabla 40	Cuantiles registrados y estimados (mm) para diferentes periodos de retorno y métodos para la Estación Sondorillo	168
Tabla 41	Cuantiles registrados y estimados (mm) para diferentes periodos de retorno y métodos para la Estación El Limón.....	169
Tabla 42	Cuantiles registrados y estimados (mm) para diferentes periodos de retorno y métodos para la Estación Querocotillo.....	170
Tabla 43	Cuantiles registrados y estimados (mm) para diferentes periodos de retorno y métodos para la Estación Hacienda Pucará.....	171
Tabla 44	Cuantiles registrados y estimados (mm) para diferentes periodos de retorno y métodos para la Estación Hacienda Shumaya.....	172
Tabla 45	Cuantiles registrados y estimados (mm) para diferentes periodos de retorno y métodos para la Estación Huarmaca	173
Tabla 46	Resumen de cobertura espacial y número de estaciones por región homogénea (ARF) ...	174
Tabla 47	Estaciones y área de influencia por Región homogénea 1 definida mediante ARF.....	175
Tabla 48	Estaciones y área de influencia por Región homogénea 2 definida mediante ARF.....	175
Tabla 49	Estaciones y área de influencia por Región homogénea 3 definida mediante ARF.....	176
Tabla 50	Estaciones y área de influencia por Región homogénea 4 definida mediante ARF.....	176
Tabla 51	Indicadores de cobertura espacial e instrumentación por región homogénea (ARF).....	185
Tabla 52	Estrategias derivadas de la regionalización L-momentos para mitigar insuficiencia de datos (síntesis operativa)	187
Tabla 53	Priorización técnica de zonas para reforzar instrumentación (orientada a mitigación de vacíos)	188
Tabla 54	Máximos valores de discordancia y heterogeneidad en las regiones formadas	190

Tabla 55 Estadísticas descriptivas para Dmax y H1	192
Tabla 56 Prueba t-Student para Dmax	192
Tabla 57 Prueba t de Student para H ₁	192
Tabla 58 Contraste cuantitativo referencial con Cadenillas Campos (2024), considerando la diferente estructura temporal de las series	197
Tabla 59 Contraste metodológico de la sensibilidad del ARF respecto al ajuste local con Roca Calderón (2021).....	199

ÍNDICE DE FIGURAS

Figura 1	Diagrama de dendrograma para la jerarquización de estaciones.....	43
Figura 2	Esquema de nube de puntos de ratios de Momentos L.....	52
Figura 3	Ratios de Momento – L para datos observados y simulado	54
Figura 4	Ejemplo de relaciones entre L - C_S– L - C_K para distribuciones de frecuencia	55
Figura 5	Partes de una cuenca.....	58
Figura 6	Tipos de formas de las curvas hipsométricas en una cuenca.....	60
Figura 7	Esquema de investigación	75
Figura 8	Ubicación política de la cuenca Chamaya.....	77
Figura 9	Ubicación administrativa e hidrográfica de la cuenca Chamaya.....	77
Figura 10	Delimitación de la cuenca del río Chamaya	85
Figura 11	Mapa de las zonas climáticas de la cuenca Chamaya.....	87
Figura 12	Mapa de ubicación de las estaciones meteorológicas de la cuenca Chamaya	90
Figura 13	Estaciones meteorológicas de la cuenca Chamaya.....	90
Figura 14	Detalles regionales ARF – Región 1: L-momentos regionales, discordancia de estaciones y distribución espacial (longitud–latitud).....	98
Figura 15	Detalles regionales ARF – Región 2: L-momentos regionales, discordancia de estaciones y distribución espacial (longitud–latitud)	98
Figura 16	Detalles regionales ARF – Región 3: L-momentos regionales, discordancia de estaciones y distribución espacial (longitud–latitud)	99
Figura 17	Detalles regionales ARF – Región 4: L-momentos regionales	99
Figura 18	Polígono de Thiessen.....	105
Figura 19	Creación del mapa de regiones homogéneas.....	106
Figura 20	Ubicación de las estaciones meteorológicas.....	112
Figura 21	Distribución espacial de las estaciones meteorológicas utilizadas	115
Figura 22	Boxplot de PMDA por estación (comparación rápida de magnitudes y dispersión).....	119
Figura 23	Precipitaciones de la Estación 01	120
Figura 24	Precipitaciones de la Estación 02	120
Figura 25	Precipitaciones de la Estación 03	121
Figura 26	Precipitaciones de la Estación 04	121
Figura 27	Precipitaciones de la Estación 05	121
Figura 28	Precipitaciones de la Estación 06	122
Figura 29	Precipitaciones de la Estación 07	122
Figura 30	Precipitaciones de la Estación 08	122
Figura 31	Precipitaciones de la Estación 09	123

Figura 32	Precipitaciones de la Estación 10	123
Figura 33	Precipitaciones de la Estación 11	123
Figura 34	Precipitaciones de la Estación 12	124
Figura 35	Precipitaciones de la Estación 13	124
Figura 36	Precipitaciones de la Estación 14	124
Figura 37	Precipitaciones de la Estación 15	125
Figura 38	Precipitaciones de la Estación 16	125
Figura 39	Precipitaciones de la Estación 17	125
Figura 40	Precipitaciones de la Estación 18	126
Figura 41	Precipitaciones de la Estación 19	126
Figura 42	Precipitaciones de la Estación 20	126
Figura 43	Precipitaciones de la Estación 21	127
Figura 44	Precipitaciones de la Estación 22	127
Figura 45	Precipitaciones de la Estación 23	127
Figura 46	Precipitaciones de la Estación 24	128
Figura 47	Precipitaciones de la Estación 25	128
Figura 48	Dendrograma (agrupación jerárquica) de los grupos formados tentativamente	137
Figura 49	Agrupación preliminar con el método Ward	138
Figura 50	Agrupación preliminar de estaciones por el método K-means.....	138
Figura 51	Regiones hidrológicas (4) en la zona de estudio	142
Figura 52	Gráfico de dispersión de las ratios de Momentos L (relación entre el coeficiente de variación y el coeficiente de asimetría)	144
Figura 53	Gráfico de dispersión de las ratios de Momentos L (relación entre el coeficiente de curtosis y el coeficiente de asimetría)	145
Figura 54	Curvas de distribución de frecuencias de cuantiles regionales, región 1	148
Figura 55	Curvas de distribución de frecuencias de cuantiles regionales, región 2	148
Figura 56	Curvas de distribución de frecuencias de cuantiles regionales, región 3	148
Figura 57	Curvas de distribución de frecuencias de cuantiles regionales, región 4	149
Figura 58	Curvas de distribución de frecuencias locales de cada estación de la región 1	155
Figura 59	Curvas de distribución de frecuencias locales de cada estación de la región 2	156
Figura 60	Curvas de distribución de frecuencias locales de cada estación de la región 3	157
Figura 61	Curvas de distribución de frecuencias locales de cada estación de la región 4	157
Figura 62	Estaciones agrupadas para diferentes periodos de retorno	159
Figura 63	Precipitación de diseño para un periodo de retorno de 100 años y distribuciones de frecuencias para la estación Sondorillo (Grupo 1).....	160
Figura 64	Precipitación de diseño para un periodo de retorno de 100 años y distribuciones de frecuencias para la estación El Limón (Grupo 2).....	161

Figura 65	Precipitación de diseño para un periodo de retorno de 100 años y distribuciones de frecuencias para la estación Querocotillo (Grupo 3).....	162
Figura 66	Precipitación de diseño para un periodo de retorno de 100 años y distribuciones de frecuencias para la estación Hacienda Pucará (Grupo 4).....	163
Figura 67	Precipitación de diseño para un periodo de retorno de 100 años y distribuciones de frecuencias para la estación Hacienda Shumaya (Grupo 5).....	164
Figura 68	Precipitación de diseño para un periodo de retorno de 100 años y distribuciones de frecuencias para la estación Huarmaca (Grupo 6)	165
Figura 69	Comparación espacial entre regiones ARF y grupos ROI (RI)	167
Figura 70	Comparación gráfica de tendencias – Estación Sondorillo	168
Figura 71	Comparación gráfica de tendencias – Estación El Limón	169
Figura 72	Comparación gráfica de tendencias – Estación Querocotillo	170
Figura 73	Comparación gráfica de tendencias – Estación Hacienda Pucará	171
Figura 74	Comparación gráfica de tendencias – Estación Hacienda Shumaya	172
Figura 75	Comparación gráfica de tendencias – Estación Huarmaca.....	173
Figura 76	Regiones homogéneas identificadas mediante ARF en la cuenca del río Chamaya	177
Figura 77	Mapa de PMDA regionalizado para T=50 años (mm).	178
Figura 78	Mapa de PMDA regionalizada para T=100 años (mm).	179
Figura 79	Mapa de PMDA regionalizada para T=200 años (mm).	180
Figura 80	Mapa de PMDA regionalizada para T=500 años (mm).	181
Figura 81	Mapa final de regiones homogéneas por ROI	182
Figura 82	Región de influencia 1, 2, 3, 4, 5 y 6.....	183
Figura 83	Densidad de estaciones por región homogénea (ARF)	186
Figura 84	Esquema de decisión ARF vs ROI para estimación en puntos no instrumentados	189
Figura 85	Prueba de normalidad de Dmax	191
Figura 86	Prueba de normalidad de H1	191

LISTA DE SIGLAS Y ABREVIATURAS

Sigla / abrev.	Significado
AAR	<i>Annual Amount of Rainfall</i> ; indicador de precipitación anual (usado en Japón).
AED	Análisis Exploratorio de Datos.
AM/AMS	<i>Annual Maximum / Annual Maximum Series</i> ; serie de máximos anuales (máximo diario por año).
AMeDAS	<i>Automated Meteorological Data Acquisition System</i> ; sistema automatizado de adquisición de datos meteorológicos (Japón).
ANA	Autoridad Nacional del Agua.
ANORD	Indicador de orden/frecuencia anual asociado a precipitación extrema (Japón).
AR(1)	Modelo autorregresivo de primer orden (<i>Auto-Regressive</i> , orden 1).
ARF	Análisis Regional de Frecuencia.
CN	<i>Curve Number</i> ; Número de escurrimiento del Soil Conservation Service.
CV (C_v)	Coeficiente de variación.
Dmax	Máximo valor del estadístico de discordancia dentro de una región/grupo.
EEE	Error Estándar de Estimación.
ESS	<i>Error Sum of Squares</i> ; suma de cuadrados del error (función objetivo en Ward).
ET	Evapotranspiración (<i>Evapotranspiration</i>).
FDA	Función de Distribución Acumulativa.
FRA	Factor de Reducción por Área.
GEV	<i>Generalized Extreme Value</i> ; distribución de Valor Extremo Generalizado.
GHM	<i>Global Hydrological Model</i> ; modelo hidrológico global.
GLO	<i>Generalized Logistic</i> ; distribución Logística Generalizada.
GPA	<i>Generalized Pareto</i> ; distribución Pareto Generalizada.
HEC-HMS	<i>Hydrologic Engineering Center – Hydrologic Modeling System</i> ; software de modelación hidrológica.
H1	Estadístico de heterogeneidad H1 (prueba de Hosking y Wallis).
IDF	Intensidad–Duración–Frecuencia (curvas IDF).
IID	<i>Independent and Identically Distributed</i> ; independientes e idénticamente distribuidas.
I95 / I99	Índices de intensidad asociados a percentiles 95 y 99 de precipitación (Japón).
IWM	Índice de Willmott Modificado.
knn (kNN)	<i>k-nearest neighbors</i> ; k vecinos más cercanos.
L1, L2, L3, L4	L-momentos de orden 1 a 4.

Sigla / abrev.	Significado
L-Ck (t4)	Cociente L de curtosis (razón L t4).
L-Cs (t3)	Cociente L de asimetría (razón L t3).
L-CV (t)	Cociente L de variación (razón L t).
LN3	Lognormal de tres parámetros.
logMAE	<i>log Mean Absolute Error</i> ; error absoluto medio en escala logarítmica.
LNO	Lognormal (usada como familia candidata en ARF).
MPP	Momentos Ponderados Probabilísticamente.
NS	Eficiencia de Nash–Sutcliffe.
P3/PE3	Pearson tipo III (notación usada en el texto para la familia de Pearson III).
PMDA	Precipitaciones Máximas Diarias Anuales.
PT	Precipitación de diseño (cuantil de precipitación para un periodo de retorno).
QGIS	<i>Quantum Geographic Information System</i> ; software SIG (QGIS).
R	Escorrentía (<i>Runoff</i>).
R ² ajustado	Coefficiente de determinación ajustado (<i>Adjusted R-squared</i>).
RF	<i>Random Forest</i> ; algoritmo de bosque aleatorio.
RMSE	<i>Root Mean Square Error</i> ; raíz del error cuadrático medio.
ROI	Región de Influencia (<i>Region of Influence</i>).
RVM	<i>Regional Vector Method</i> ; método del vector regional.
SENAMHI	Servicio Nacional de Meteorología e Hidrología del Perú.
SI/MLR	Enfoque combinado de similitud (<i>Similarity Index</i> , SI) y regresión lineal múltiple (<i>Multiple Linear Regression</i> , MLR).
SIG	Sistema(s) de Información Geográfica.
SP	Proximidad espacial (<i>Spatial Proximity</i>).
TSA	Enfoque/método Theil–Sen (<i>Theil–Sen Approach</i>).
TFPW	<i>Trend-Free Pre-Whitening</i> ; preblanqueo sin tendencia.
TR / Tr	Tiempo o periodo de retorno.
TRMM	<i>Tropical Rainfall Measuring Mission</i> ; producto/base satelital de precipitación.
WS	Almacenamiento de agua (<i>Water Storage</i>).
Zs (Z _s)	Estadístico estandarizado de Mann–Kendall.

RESUMEN

La cuenca del río Chamaya presenta heterogeneidad en la información pluviométrica para estimar las precipitaciones máximas diarias anuales (PMDA), lo que incrementa la incertidumbre en la delimitación espacial de los extremos. En este marco, la investigación tuvo como objetivo regionalizar las PMDA en la cuenca del río Chamaya mediante el método de los L-momentos. La metodología fue de enfoque cuantitativo, tipo básico, nivel descriptivo y diseño no experimental; la muestra estuvo conformada por series de PMDA de estaciones meteorológicas seleccionadas dentro y cerca de la cuenca. Los resultados integraron 25 estaciones meteorológicas en el inventario inicial; 24 conformaron las cuatro regiones homogéneas verificadas mediante ARF y 22 aportaron áreas efectivas para la espacialización cartográfica dentro de la cuenca, con longitudes de serie de 21 a 60 años y completitud de 72.7% a 98.3%, con promedios entre 29.0 mm y 75.3 mm y máximos anuales de 145.2 mm y 130.8 mm; se corrigieron estaciones no homogéneas y se validaron 4 regiones homogéneas con 8, 8, 4 y 4 estaciones ($D_{max} \leq 1.94$; $H_1 = 0.92 - 1.74$), con distribución regional LNO consistente ($|Z| = 0.25 - 0.63$). Se elaboraron mapas regionales ARF sobre 8 137.19 km², evidenciando un mayor contraste entre extremos al aumentar el TR (de $\leq 60 - > 120$ mm en TR=50 a $\leq 75 - > 165$ mm en TR=500), y se propusieron estrategias operativas de transferencia (ARF y RI) y de priorización del refuerzo de la red (alta en R-1 y R-4). Se concluye que la regionalización mediante L-momentos permitió estimar cuantiles regionales y elaborar productos cartográficos coherentes, lo que redujo los vacíos de información y mejoró la toma de decisiones.

Palabras clave: análisis regional de frecuencias (ARF), regiones homogéneas, regionalización hidrológica.

ABSTRACT

The Chamaya River basin exhibits heterogeneous rainfall characteristics for estimating annual maximum daily precipitation (AMDP), increasing uncertainty in the spatial delineation of extreme rainfall events. Within this context, the objective of the study was to regionalize AMDP in the Chamaya River basin using the L-moments method. The research followed a quantitative approach, a basic research scope, a descriptive level, and a non-experimental design. The sample consisted of AMDP time series from selected meteorological stations located within and near the basin. The initial inventory comprised 25 meteorological stations; 24 stations were used to define four homogeneous regions verified through Regional Frequency Analysis (RFA), and 22 stations contributed effective influence areas for cartographic spatialization within the basin. The time series ranged from 21 to 60 years, with completeness levels between 72.7% and 98.3%, mean values between 29.0 mm and 75.3 mm, and notable annual maxima of 145.2 mm and 130.8 mm. Non-homogeneous stations were corrected, and four homogeneous regions containing 8, 8, 4, and 4 stations were validated, respectively ($D_{\max} \leq 1.94$; $H_1 = 0.92-1.74$), with a consistent regional LNO distribution ($|Z| = 0.25-0.63$). RFA-based regional maps were developed for an area of 8,137.19 km². These maps revealed greater differentiation among extreme precipitation values as the return period increased, with classes ranging from ≤ 60 mm to >120 mm for a 50-year return period and from ≤ 75 mm to >165 mm for a 500-year return period. Operational strategies were also proposed for transferring information through RFA and the Region of Influence (ROI) approach, as well as for prioritizing the strengthening of the monitoring network, with high priority assigned to regions R-1 and R-4. It was concluded that L-moment-based regionalization enabled the estimation of regional quantiles and the development of consistent cartographic products, thereby reducing information gaps and improving decision-making.

Keywords: regional frequency analysis, homogeneous regions, hydrological regionalization.

CAPÍTULO I: INTRODUCCIÓN

1.1. Planteamiento del problema

La caracterización y regionalización de las precipitaciones máximas diarias anuales (PMDA) constituye un insumo crítico para el diseño hidrológico y la gestión hídrica, especialmente cuando la información disponible es escasa o presenta discontinuidades. En este marco, Markovič y Faško (2024) indican que los métodos de regionalización basados en L-momentos permiten estimar cuantiles de precipitación en regiones con datos limitados, lo que mejora la confiabilidad de los parámetros de diseño. Asimismo, Prahadchai et al. (2024) advierten que el reto técnico no se reduce a atribuir los cambios a la variabilidad climática, sino que exige comprender el comportamiento extremo desde un enfoque estadístico. Por consiguiente, cuando no se dispone de registros largos y consistentes, la estimación de eventos críticos incrementa su incertidumbre y limita la toma de decisiones.

En América Latina, la coexistencia de dominios climáticos contrastantes y la cobertura desigual de las redes de monitoreo condiciona la disponibilidad y representatividad de series de precipitación extrema (Ilbay-Yupa et al., 2021). En este contexto, Cutillas-Lozano et al. (2023) señalan que la identificación de regiones homogéneas es un requisito para aplicar modelos probabilísticos regionales y obtener estimaciones confiables de eventos extremos. De igual manera, Valenzuela-Morales et al. (2023), en el caso de México, destacan el uso de los L-momentos como herramienta para regionalizar las precipitaciones máximas diarias. Sin embargo, a pesar de estos avances, aún persisten vacíos en la caracterización espacial de los extremos, lo que incrementa la incertidumbre en la estimación de cuantiles y, en consecuencia, la vulnerabilidad ante fenómenos hidrometeorológicos severos (Ilbay-Yupa et al., 2021).

Perú presenta una gran diversidad climática asociada a gradientes altitudinales, lo que complica la extrapolación espacial de las lluvias máximas y la definición de criterios de diseño representativos. Idrogo (2024) destaca la alta variabilidad espacial de las precipitaciones extremas, lo que demanda enfoques robustos para regionalizar la información y reducir sesgos derivados de la representatividad de las estaciones individuales. Mientras que los métodos tradicionales basados en el ajuste local “estación por estación” pueden ser útiles en contextos con series extensas, en contextos con registros cortos o dispersos resultan insuficientes para una planificación hídrica confiable y una gestión del riesgo con sustento probabilístico (Flores et al., 2022). De este modo, existe una brecha entre la complejidad espacial del régimen extremo y la capacidad de los métodos convencionales para capturarla de manera consistente.

Al norte del país, particularmente en Cajamarca, la variabilidad de las PMDA se intensifica debido a una topografía accidentada y a dinámicas de uso del suelo que condicionan la respuesta hidrológica ante lluvias intensas. Según el *Servicio Nacional de Meteorología e Hidrología* [SENAMHI] (2023), entre diciembre y abril se registra una mayor frecuencia de precipitaciones intensas, lo que incrementa la exigencia en la modelación de obras. En consecuencia, la necesidad de regionalizar las lluvias máximas cobra relevancia para optimizar los diseños de control y disminuir los riesgos asociados a inundaciones, erosión y movimientos en masa en cuencas expuestas a amenazas hidrometeorológicas (Llauca et al., 2021). Además, la limitación de las redes de monitoreo y la distribución desigual de las estaciones de monitoreo aumentan la incertidumbre en la estimación de eventos extremos a nivel regional.

La cuenca del río Chamaya, ubicada al norte del Perú, abarca un territorio que integra nueve provincias: Huancabamba, San Ignacio, Jaén, Cutervo, Chota, Santa Cruz, Lambayeque, Ferreñafe y Hualgayoc, con una superficie aproximada de 8 137.19

km². En esta cuenca se observa una elevada variabilidad espacial y temporal de las PMDA, lo que incrementa la exposición a eventos hidrometeorológicos extremos y dificulta la definición de parámetros de diseño representativos para obras hidráulicas. SENAMHI (2024) reporta que la precipitación anual promedio se sitúa entre 1 000 y 1 200 mm, concentrándose principalmente entre diciembre y abril; además, en eventos recientes, como en marzo de 2023, se registraron intensidades de hasta 80 mm/día en estaciones como Choropampa y Chontalí, lo que generó desbordes, inundaciones en áreas agrícolas y en vías rurales, interrupciones del tránsito y afectaciones a viviendas. En el mismo periodo, se identificaron incrementos de hasta un 40% por encima del promedio histórico en los acumulados mensuales de estaciones clave, lo que evidencia condiciones propicias para impactos recurrentes. Además, de acuerdo con la *Autoridad Nacional del Agua* [ANA] la gestión del recurso hídrico en la cuenca se desarrolla bajo presiones asociadas a la demanda agrícola y a cambios en la oferta hídrica, lo que incrementa conflictos de uso y compromete la sostenibilidad de actividades productivas, tal como fue advertido en el ámbito Chinchipe-Chamaya (ANA, 2011).

Por otro lado, en el ámbito regional existen investigaciones afines que constituyen una base técnica relevante para comprender el comportamiento hidrológico del norte de Cajamarca. Bustamante (2019) analizó la cuenca del río Chamaya desde una perspectiva morfométrica, mientras que Cadenillas (2024) regionalizó las precipitaciones máximas de 24 horas en la cuenca vecina del río Chotano, considerando estaciones próximas o compartidas con el área de influencia del río Chamaya. Estos aportes permiten reconocer avances previos; no obstante, no sustituyen un análisis específico e integral de las precipitaciones máximas diarias anuales en la cuenca del río Chamaya. La presente investigación amplía dicho marco al depurar las series pluviométricas, evaluar su estacionariedad y homogeneidad, delimitar regiones

homogéneas mediante L-momentos, estimar cuantiles de diseño para distintos periodos de retorno, comparar los enfoques de Análisis Regional de Frecuencia (ARF) y de Región de Influencia (ROI), y representar espacialmente los resultados mediante cartografía regional. El vacío de conocimiento no radica, por tanto, en la ausencia absoluta de antecedentes, sino en la falta de una evaluación articulada que reúna estos componentes en una misma cuenca y genere criterios aplicables a sectores con instrumentación meteorológica limitada. En consecuencia, la aplicación del método de L-momentos se plantea como una alternativa técnica para integrar información espacialmente, delimitar regiones homogéneas y generar una caracterización regional robusta de eventos extremos que sustente una gestión hídrica informada y sostenible.

1.2. Justificación

Esta investigación responde a la necesidad de comprender la variabilidad espacial de las precipitaciones máximas diarias anuales en la cuenca del río Chamaya, zona expuesta a eventos hidrometeorológicos extremos. La aplicación del método de L-momentos permite caracterizar las distribuciones de precipitación en regiones con datos limitados, lo que constituye un aporte significativo al conocimiento científico en hidrología regional. Según Hosking y Wallis (1997), los L-momentos resultan útiles para estimar distribuciones en contextos hidrológicos con registros incompletos, lo que favorece el desarrollo de modelos predictivos adaptados a las condiciones locales.

El estudio aporta información técnica que puede utilizarse para mejorar la gestión de los recursos hídricos y optimizar el diseño de infraestructuras hidráulicas en la región de Cajamarca. La cuenca ha sufrido históricamente daños por inundaciones y huaicos que han provocado pérdidas económicas y sociales (ANA, 2011). Al regionalizar las precipitaciones máximas, se generan herramientas que permiten a ingenieros y autoridades anticipar el comportamiento de las lluvias extremas y

planificar obras hidráulicas seguras. Esto ayuda a priorizar intervenciones en zonas vulnerables, reduciendo costos por daños y de mantenimiento a largo plazo.

La investigación incorpora el análisis de datos hidrológicos mediante los L-momentos. Este enfoque, desarrollado por Hosking (1990), ha demostrado ventajas frente a los métodos tradicionales de momentos y de máxima verosimilitud, ya que permite obtener estimaciones más estables y menos sensibles a valores atípicos, lo cual es particularmente relevante en áreas con escasez de datos confiables. El estudio no solo aplica esta metodología a la cuenca del río Chamaya, sino que además establece lineamientos replicables para futuras investigaciones en otras regiones del país, lo que contribuye a consolidar estándares metodológicos de regionalización en Perú.

Los resultados benefician a las comunidades asentadas en la cuenca del Chamaya al aportar insumos para la gestión de riesgos y la planificación territorial. La disponibilidad de información sobre lluvias extremas permite a las autoridades y a los pobladores tomar decisiones informadas para reducir su vulnerabilidad frente a inundaciones, huaicos y otros desastres relacionados.

1.3. Formulación del problema

¿Cómo se regionalizan las precipitaciones máximas diarias anuales en la cuenca del río Chamaya mediante la aplicación del método de L-momentos, 2024?

1.4. Objetivos

1.4.1. *Objetivo general*

Regionalizar las precipitaciones máximas diarias anuales (PMDA) en la cuenca del río Chamaya mediante la aplicación del método de L-momentos, 2024.

1.4.2. *Objetivos específicos*

- Analizar las series de PMDA de las estaciones meteorológicas seleccionadas en la cuenca del río Chamaya.

- Delimitar y verificar regiones homogéneas, estimar las precipitaciones de diseño mediante L-momentos y comparar los resultados obtenidos con los enfoques de Análisis Regional de Frecuencia (ARF) y Región de Influencia (ROI) en la cuenca del río Chamaya.
- Elaborar mapas de PMDA regionalizadas y caracterizar espacialmente las regiones homogéneas altoandinas de la cuenca del río Chamaya.
- Contribuir al desarrollo de estrategias para mitigar la insuficiencia de datos hidrológicos en zonas poco instrumentadas de la cuenca del río Chamaya, empleando la regionalización de PMDA mediante el método L-momentos.

1.5. Limitaciones

La estimación de cuantiles locales evidenció valores negativos para periodos de retorno extremadamente próximos a 1 año en las estaciones El Limón, Colasay, Tocmoche y Chugur. Estas estimaciones carecen de interpretación física, ya que la precipitación es una variable no negativa. El comportamiento observado no se originó en una única familia probabilística regional, sino en la extrapolación de distintos ajustes locales hacia la cola inferior de las distribuciones. Esta limitación no invalida los resultados empleados para el diseño hidrológico, ya que las aplicaciones ingenieriles se sustentaron en cuantiles con periodos de retorno iguales o superiores a 2 años. Los cálculos completos se conservaron en los archivos de respaldo de Excel entregados junto con el informe para garantizar la trazabilidad (Anexo 3).

CAPÍTULO II: MARCO TEÓRICO

2.1. Antecedentes

2.1.1. Antecedentes internacionales

Cao et al. (2025) en su investigación “*Regionalización de los cambios del ciclo hidrológico en 31 cuencas fuente de la cuenca del río Amarillo considerando múltiples variables hidrológicas*” tuvieron como objetivo caracterizar cambios integrales del ciclo hidrológico en 31 cuencas fuente de la Cuenca del Río Amarillo (2000 al 2020). La metodología fue calcular las métricas de escurrimiento (R), evapotranspiración (ET) y almacenamiento de agua (WS), además seleccionar combinaciones óptimas por error de balance hídrico y clasificar las pendientes de Sen mediante análisis de clúster para tipos de cambio. Determinaron que el error de balance hídrico de las 31 cuencas fue de -9.75% a -9.78% de precipitación media anual, en las variables hidrológicas hubo cambios significativos con pendiente media de -0.16 a -4.28 mm/a y con mayor tendencia promedio en WS (1.65 mm/a) que en ET y R (0.004 mm/a). Concluyeron que, el enfoque permite describir de forma más completa la heterogeneidad regional de los cambios hidrológicos bajo condiciones ambientales cambiantes.

Corrales y Quezada (2025) en su tesis “*Regionalización de la curva IDF para la cuenca hidrográfica del Río Puca, Provincia del Guayas*” tuvieron como objetivo desarrollar la regionalización de la curva IDF para la cuenca del río Puca. La metodología fue de enfoque cuantitativo descriptivo, teniendo como muestra la cuenca del río Puca con información máxima de 24 horas de 1990 a 2013, para ello utilizaron el programa Hydrognomon y ArcGIS. La información pluviométrica analizada proviene de 10 estaciones meteorológicas, de las cuales siete fueron seleccionadas para la regionalización: (1) Camposano, (2) Olmedo, (3) La Naranja, (4) Alajuela, (5) Pichilingue, (6) Portoviejo y (7) Río Chamotete determinando que para el tiempo de

retorno de 100 años se alcanzaba 165, 215, 97, 207.3, 210.8, 157.3 y 168.8 mm de precipitación, respectivamente. Concluyeron que, la curva IDF para el río Puca permite un entendimiento de los patrones de precipitación extrema en la región.

Pinheiro et al. (2024) en su artículo *“Regionalización del flujo utilizando datos de precipitación de diferentes bases como variable predictiva”* tuvieron como objetivo comparar diferentes bases de datos de precipitación empleadas como variables predicativas en estudios de regionalización. Aplicaron metodologías de series de caudal y seleccionaron estaciones mayores a 20 años y menor a 5% de fallas para ello emplearon 58 estaciones fluviométricas. Como resultados, las tres bases presentaron patrones espaciales similares, con precipitación media anual de 1457.5 mm (kriging), 1546.5 mm (WorldClim) y 1504.5 mm (TRMM), y amplitudes entre extremos de 584, 521 y 352 mm, respectivamente; además, los ajustes fueron satisfactorios, destacando mejor desempeño en regiones 1–3 y mayores restricciones en la región 4 para caudales más exigentes. Concluyeron que TRMM y WorldClim pueden sustituir a la precipitación observada e interpolada para regionalizar caudales en la cuenca estudiada, con resultados similares y ventajas operativas por disponibilidad y ahorro de tiempo.

Kupzig et al. (2024) en su investigación *“Regionalización en modelos hidrológicos globales y su impacto en las simulaciones de escorrentía: un estudio de caso utilizando WaterGAP3 (v 1.0.0)”* tuvieron como objetivo proveer información sobre (1) seleccionar método de regionalización adecuado para un GHM y (2) evaluar su impacto en la simulación de escorrentía. La metodología aplicada fue el modelo WaterGAP3 que simula el ciclo hidrológico terrestre, además se identificaron 4 enfoques apropiados basados en regresión, similitud física y proximidad espacial. Determinaron que la comparación de algoritmos, knn superó a K-means (logMAE 0.432 vs 0.472), por lo que K-means se consideró inadecuado; aunque RF tuvo buen

logMAE, mostró sobreajuste y se descartó por baja robustez, además en el río la mediana simulada fue de 13.7 m³/s con SP, 19.7 m³/s con knn y 3.6 m³/s con SI/MLR) y las estimaciones de esorrentía media hacia el océano variaron entre 45.622 y 47.069 km³/año (dispersión 1500 km³/año) entre los cinco métodos del conjunto. Concluyeron que se debe elegir cuidadosamente los métodos de regionalización recomendando profundizar en elección de predictores y en la robustez de estos métodos de escala.

Eraso et al. (2024) en su tesis “*Regionalización de lluvias de diseño en el departamento de Nariño*” tuvieron como objetivo construir un modelo regional de lluvias de diseño en Nariño. La metodología fue exploratoria, la muestra estuvo formada por las series de datos de lluvia máximas y se utilizaron distribuciones de probabilidad. Determinaron que el método de Gumbel predominó en 50% de las estaciones evaluadas y el método Log-Pearson Tipo III en 41.67%; además, ambos métodos estimaron valores mayores de intensidad máxima de precipitación para periodos de retorno de 2, 5 y 10 años, además identificaron que en la estación Encano, Kriging presentó el menor error (± 2.91 mm); para 51.7 mm de precipitación máxima diaria, el modelo predijo rango de 54.19 a 60.01 mm, equivalente a error aproximado de 10%. Concluyeron que la distribución Gumbel es la que más se sujeta a los datos reales de precipitación, teniendo ajustes precisos en los valores mínimos y máximos.

Yazawa y Shoji (2023) en su artículo “*Análisis espacial de las características históricas de las precipitaciones extremas mediante regionalización en la cuenca del lago Biwa y el río Yodo, Japón*” tuvieron como objetivo evaluar los cambios históricos en las características de la precipitación y su distribución espacial. Utilizaron el sistema automatizado de adquisición de datos meteorológicos (AMeDAS) recopilando datos diarios de precipitación de 61 estaciones pluviales. Lograron determinar riesgo e inundaciones en el norte, oeste, y sur, en términos espaciales la AAR aumentó en 98.2%

de las celdas y la ANORD en 53.4%; asimismo, aumentaron intensidades I95 en 84.2% e I99 en 73.8%, mientras que disminuyeron R95 en 99.99%, R99 en 86.9% y N95 en 99.3%, con N99 mostrando incremento en 50.7% de celdas y la precipitación potable el cambio relativo fue negativo en 15.8% (Tr=50 años) y 18.6% (Tr=100 años) de las celdas, y con esto los periodos de retorno disminuyeron de manera relevante a 56.1% < a 50 años y 40.4% < a 100 años. Concluyeron que la gestión local de inundaciones debe actualizarse uniendo las diferencias regionales de la precipitación extrema, y que la regionalización es especialmente útil donde la red de pluviómetros es dispersa.

Escamilla (2023) en su tesis “*Actualización de factores de regionalización de la cuenca Teapa y aplicación en MPE y HEC-HMS*” tuvo como objetivo analizar la regionalización asociados a diferentes periodos de retorno, haciendo uso de los más recientes datos climatológicos disponibles en las estaciones meteorológicas de la zona. La metodología que utilizó fue periodo de retorno, estadísticos muestrales, función Gumbel, tiempo de concentración y número de escurrimiento de Soil Conservation Service (CN). Los resultados indicaron diferencia notable en los factores de regionalización respecto al estudio del año 2017 en dicho lugar, para los periodos de retorno entre 10,000 y 1,000 se registró una variación entre 12% y 8%, en tanto para los periodos entre 1,000 y 2 años la variación osciló entre 8% y 3%. Concluyó que la actualización fue pertinente y recomendó efectuar una nueva actualización de la cuenca cada 5 años debido a que ese intervalo de datos genera diferencias significativas.

Ramírez et al. (2023) en su artículo científico “*Regionalización Hidrológica de la Cuenca del Arroyo Feliciano, Entre Ríos, Argentina*” tuvieron como objetivo analizar y determinar los parámetros morfométricos y climáticos que describen una respuesta hidrológica en la cuenca del Arroyo Feliciano. La metodología reside en la aplicación de técnicas de estadística multivariada y la delimitación SIG-QGIS de las 77

subcuencas y 22 variables con información morfología y climática. Determinaron que una cuenca de 8300 km² con precipitaciones medias anuales de 1100 a 1200 mm, la retención del 56.2% de los datos en las dos primeras componentes, la identificación de cuatro grupos/ regiones hidrológicamente homogéneas y un modelo parsimonioso con Adjusted R-squared = 0.78 y p-valor < 2.2e-16 indicando que el modelo ajusta muy bien, donde el área fue la variable de mayor influencia R² = 78.11%, seguida por CN y precipitación (precipitaciones medias de 1200 a 1300 mm anuales). Concluyeron que es posible regionalizar la cuenca en cuatro regiones y explicar el caudal medio diario.

Skidelsky et al. (2023) en su investigación “*Regionalización de caudales máximos mediante técnicas estadísticas multivariadas en cuencas de la provincia de Entre Ríos*” tuvieron como objetivo regionalizar para poder determinar ecuaciones donde permita estimar valores de caudales máximos. Utilizaron estadísticas multivariadas para determinar regiones hidrológicamente homogéneas con registros de 20 años de caudales medios. Lograron determinar ecuaciones para 5 años de recurrencia y pertinentes valores de (Adjusted R-squared) para cada grupo, analizando área y pendiente de las cuencas, en el G1: las características son áreas grandes entre 5,000 y 16,100 km², bajas pendientes entre 0.00021 y 0.00024 m/m con adjusted R-squared 98.01%, el G2: áreas pequeñas entre 160 y 170 km², altas pendientes entre 0.00092 y 0.0016 m/m su adjusted de R-squared 93.17%, G3: áreas menores a medias entre 70 y 2300 km², altas pendientes, entre 0.00042 y 0.00069, adjusted de R-squared 93.65%, G4: áreas grandes entre 740 y 5,100 km², bajas pendientes, entre 0.00036 y 0.00052 m/m, adjusted de R-squared 63.19% y G5; áreas grandes menores al grupo 1, entre 2,400 y 7,000 km² y bajas pendientes, entre 0.00022 y 0.00038 y su adjusted de R-squared de 98.88%. Concluyeron, que a regionalización admitió obtener ecuaciones prácticas para estimar caudales máximos de diseños de cuencas.

Martínez (2022) en su tesis “*Regionalización de precipitaciones para obtener lluvias de diseño en la cuenca del río Santa Catarina del estado de Nuevo León*” tuvo como objetivo tener una mejor y mayor base sobre datos de lluvia de dicha cuenca con la cual permitirá calcular tormentas de diseño para realizar obras hidráulicas civiles eficientes y seguras. En la metodología utilizó las precipitaciones máximas anuales y para la depuración el programa CLICOM para obtener valores estandarizados. Determinaron que los factores de diseño normalizados adimensionales extrapolados, para estimar tormentas de diseño de duraciones de 1, 3, 5, 10 y 15 días, en función del período de retorno ($T_r = 2$ a $10\,000$ años). En general, los factores aumentan conforme se incrementa el período de retorno, además para un mismo período de retorno, los valores tienden a ser más altos en duraciones cortas de 1 día y menores en duraciones intermedias de 3 a 10 días. Concluyó que la regionalización permitió mejorar la base de datos de lluvia y generar diseño de lluvias confiables.

Cifuentes (2020) en su tesis de maestría “*Automatización de las estimaciones de caudales de crecidas para diferentes periodos de retorno*” tuvo como objetivo proponer un método rápido y efectivo para la automatización de las estimaciones de caudales de crecida para diferentes periodos de retorno, la propuesta metodológica planteada se basó en la regionalización hidrológica, que consiste en el traspaso de información entre cuencas hidrológicamente homogéneas se aplicaron diferentes técnicas estadísticas que logró determinar un agrupamiento adecuado de 5 grupos homogéneos G1: A 799 km^2 , R_k 2.36 y Em 1931 m, G2: A 436 km^2 , R_k 2.11 y Em 355 m, G3: A 263 km^2 , R_k 1.78 y Em 1078 m, G4: A 1131 km^2 , R_k 2.45 y Em 765 m, y G5: A 258 km^2 , R_k 3.07 y Em 598 m. Concluyó que la automatización de estimación de caudales de crecida es posible, permitiendo conseguir estimaciones rápidas y confiables.

2.1.2. Antecedentes nacionales

Ticona (2024) en su tesis “*Estimación de precipitaciones máximas bajo la implementación del método de regionalización y estadístico aplicado en la cuenca Cabanillas departamento de Puno*” tuvo como objetivo evaluar los fenómenos meteorológicos extremos para mejorar la estimación de caudales máximos. Utilizó modelos probabilísticos, frecuencia de crecidas y L-momentos. Determinaron que la regionalización mediante L-momentos permitió conformar regiones homogéneas para precipitaciones (H en Cabanillas -1.68 , Juliaca -1.78 , Lampa -1.72 , Pampahuata -1.69 , Santa Lucia -1.65), temperatura (H en Cabanillas -10.78 , Juliaca -6.44 , Lampa -6.99 , Pampahuata -3.48 , Santa Lucia -4.99) y caudales (H en Rio Verde -1.527 , Puente Cabanillas -1.66 , Puente Unocolla -1.59); además, para caudales la distribución que mejor se ajusta es la GPA con $\xi = 1.66$, $\alpha = 18.33$ y $k = 0.21$, estimándose un caudal máximo para Tr 20 años en el puente Unocolla de $108.87 \text{ m}^3/\text{s}$. Concluyó que L-momentos es una alternativa firme para el análisis regional de eventos extremos.

Flores-Quispe et al. (2022) en su artículo “*Regionalización de precipitación máxima diaria en Moquegua Perú*” tuvieron como objetivo realizar regionalización de precipitación máxima de 24 horas. Utilizaron el análisis de datos de precipitación máxima en 12 estaciones meteorológicas a través de análisis clúster con registros 1965 a 2009. Obtuvieron cinco regiones homogéneas. Región 1: Ilo y Moquegua, región 2: Yacango, Cuajone, Omate, Quinistaquillas, región 3: Puquina, Ubinas, Carumas y Calacoa, región 4: Quellaveco y región 6: Ichuña, con precipitaciones de 59.80, 70.75, 83.61 y 98.75 mm, respectivamente a los Tr de 25, 50, 100 y 200 años. La relación precipitación máxima, Tr, altitud alcanzo $R^2 = 0.789$ con coeficientes significativos. Concluyeron que la regionalización estima precipitación máxima de 24 h en regiones homogéneas e incorpora control geográfico para estudios de diseño hidrológico.

Roca (2021) en su tesis *“Regionalización de precipitaciones máximas diarias anuales en la zona Sur-Occidental del Perú usando Momentos – L”* tuvo como objetivo identificar regiones hidrológicas a partir de la caracterización de precipitaciones máximas diarias con la aplicación de momentos L en la zona sur – occidental del Perú (Huancavelica, Ica, Ayacucho, Apurímac, Arequipa y Moquegua). Utilizó métodos de regionalización hidrológica, donde se redujeron las bases de 368 a 173 estaciones (47%) logrando una densidad de 1 estación/1340 km², también se halló la homogeneidad con la prueba Hosking y Wallis definiendo 8 regiones hidrológicas donde para Tr 100años la precipitación fue 80 mm. Demostró con ARF y ROI desempeño satisfactorio en 5 de 8 estaciones, con discrepancias en estaciones cercanas a la franja costera. Concluyó que los dos métodos son clave para estimar extremos diarios cuando la pluviometría es limitada al usar registros de estaciones cercanas.

Vásquez (2025) en su tesis *“Regionalización de precipitaciones promedio en la cuenca hidrográfica del Alto Marañón, mediante métodos de agrupación y el vector regional (RVM)”* tuvo como objetivo regionalizar las precipitaciones mensuales en función de la precipitación total anual y de la latitud, longitud y altitud. Utilizó series mensuales de 54 estaciones internas y 02 estaciones ficticias RAIN4PE (Rain4pe 01 para Ecuador y Rain4pe 02 para el sur del Perú), aplicando control de calidad, agrupamiento Ward y K-means en RStudio y validación con vector regional; completó datos en Hydraccess y delimitó regiones con series 1981 al 2023 e interpolación Kriging. Identificó 06 regiones homogéneas: ROI (3 estaciones; 1750.09 mm), R II (10; 1087.68 mm), R III (11; 643.67 mm, con periodos secos de 0. mm), R IV (16; 880.88 mm), RV (2; 2298.36 mm) y R VI (12; 1025.34 mm), concluyendo que la regionalización sustenta el balance hídrico y la estimación de oferta y disponibilidad hídrica en la cuenca.

2.1.3. Antecedentes regionales

Cadenillas (2024) desarrolló su tesis *“Regionalización de precipitaciones máximas de 24 horas utilizando L-Momentos en la cuenca del río Chotano, 2023”* tuvo como objetivo regionalizar estas precipitaciones en la cuenca del río Chotano mediante el método de L-Momentos, utilizando datos de cinco estaciones meteorológicas. Determinaron que los indicadores de heterogeneidad evidencian homogeneidad regional entre las estaciones Chota (−3.379), Chotano Lajas (−3.21), Cutervo (−3.052), Huambos (−2.837) y Chugur (−3.325); además, el ajuste probabilístico mostró GEV para Cutervo, Huambos y Chugur y GLO para Chota y Chotano-Lajas, estimándose precipitaciones máximas de 24 h, con rangos regionales de 29.2–31.7 mm (5 años), 35.1–36.7 mm (10 años), 38.9–41.1 mm (20 años), 41.96–48.65 mm (50 años), 41.68–52.83 mm (75 años), 41.82–55.25 mm (100 años) y 43.82–68.99 mm (500 años). Concluyó que la aplicación del método L-Momentos proporciona datos confiables.

Álvarez (2022) en su tesis *“Modelo hidrológico para pronosticar caudales recesivos y estimar el coeficiente de agotamiento: caso río Jequetepeque aguas arriba de la presa “Gallito Ciego” - Contumazá, 2020”* tuvo como objetivo proponer modelo hidrológico para pronosticar caudales recesivos del río Jequetepeque. Utilizó los métodos de SplitSample Test para la división de muestras e indicadores estadísticos, como: Nash-Sutcliffe (NS), Raíz del Error Cuadrática Medio (RMSE), Error Estándar de Estimación (EEE), Índice de Willmott Modificado (IWM), para la simulación del modelo ALVI con coeficiente de agotamiento de 0.005 día. Validó el modelo a partir de indicadores estadísticos, además, determinó: NS (0.84), RMSE (0.62), R2 (0.97), EEE (0.66) y IWM (0.84). Concluyó que el rango de valores de coeficiente de agotamiento fue de $0.0148 > \alpha > 0.0003$, también estableció modelo hidrológico ALVI para la simulación y pronósticos de caudales recesivos del río Jequetepeque.

2.2. Bases teóricas

2.2.1. *Análisis Regional de Frecuencia (ARF)*

El Análisis Regional de Frecuencia (ARF) es marco estadístico que “*sustituye espacio por tiempo*” al combinar información de varias estaciones, con el propósito de estimar cuantiles de precipitación extrema con menor incertidumbre que un análisis estrictamente local, cuando los registros son cortos o incompletos. En el caso de precipitaciones máximas diarias anuales (PMDA), el ARF se apoya en la depuración y control de calidad de series, la conformación de regiones con comportamiento similar y el ajuste de distribución regional (Estrany-Planas et al., 2025).

a) Enfoque regional y principio de homogeneidad

El enfoque regional asume que, tras normalizar las magnitudes con factor de escala, las estaciones de misma región comparten distribución similar, mediante curva de crecimiento regional. La homogeneidad se verifica cuando los cocientes de L-momentos (t , t_3 , t_4) son consistentes entre estaciones, evaluándose con medidas de discordancia, pruebas de heterogeneidad, agrupamiento y análisis espacial. Pero, deben revisarse los supuestos de estacionariedad, independencia y correlación espacial, porque su incumplimiento puede sesgar la regionalización (Estrany-Planas et al., 2025).

b) Método Index-Flood – Índice de tormenta

Los procedimientos del índice de avenida constituyen un método conciso para sintetizar estadísticos representativos de un conjunto de muestras (Hosking y Wallis, 1997). Pueden aplicarse a cualquier variable extrema; en este estudio se emplean para precipitaciones máximas, por lo que se adoptará el término índice de tormenta. Considerando N estaciones con tamaños muestrales n_i y datos Q_{ij} , el método asume homogeneidad regional: las distribuciones de frecuencia de todas las estaciones tienen la misma forma y solo difieren por factor de escala característico de la zona.

$$Q_i(F) = \mu_i q(F), \quad i = 1, \dots, N. \quad (1)$$

Siendo μ_i el índice de tormenta, que se supone ser la media de la distribución de frecuencia en la estación o el momento l_1 . El segundo parámetro, $q(F)$, es la curva de incremento regional, función cuantil adimensional común a todas las áreas de estudio de la región dada, también referida como la función cuantil de distribución de frecuencia regional, Q_{ij}/μ_i (Hosking y Wallis, 1997).

c) Ventajas y limitaciones del ARF en cuencas andinas

En cuencas andinas, el ARF mejora la estimación de la precipitación de diseño (PT) cuando existen pocas estaciones o series cortas, facilita la transferencia a zonas poco instrumentadas y respalda la regionalización. No obstante, la homogeneidad suele ser difícil por la heterogeneidad climática y topográfica; además, la no estacionariedad y la correlación espacial pueden sesgar los cuantiles si no se controlan. Por ello, se requiere delimitación regional cuidadosa, verificación estadística estricta y criterios fisiográficos consistentes para obtener cuantiles defendibles (Ishaq, 2025).

2.2.2. Teoría de L-Momentos

A inicios de la década de 1970, los primeros estudios de regionalización buscaban modelar máximos instantáneos, usualmente con Gumbel y ajuste por momentos; no obstante, Wallis et al. (1974) evidenciaron sesgos importantes del método de momentos, incrementando la incertidumbre en cuantiles (Hosking y Wallis, 1997). En el análisis regional, Wallis et al. (1997) describieron la “separación de asimetrías” en máximos anuales, interpretándola como heterogeneidad regional, mientras que Landwehr et al. (1978) señalaron que la distribución Wakeby no presenta tal separación, aunque su parametrización es compleja con métodos tradicionales. En este marco, Greenwood et al. (1979) desarrollaron los Momentos Ponderados Probabilísticamente, base para que Hosking (1990) formulara los Momentos L como

combinaciones lineales robustas para describir ubicación, escala y forma. Posteriormente, Stedinger et al. (1993) recomendaron su uso en análisis de frecuencia de avenidas, y Hosking y Wallis (1997) consolidaron el enfoque de análisis de frecuencia regional con Momentos L, destacando su resistencia a valores atípicos.

2.2.2.1. Definición de los momentos

El propósito del análisis de frecuencia es intentar ajustar los datos a una distribución cuya forma esté determinada por un número limitado de parámetros. La forma de la distribución de probabilidad está determinada tradicionalmente por los momentos de la distribución. Los momentos son la media (Hosking y Wallis, 1997).

$$\mu = E(X) \quad X: \text{Variable aleatoria} \quad (2)$$

y los momentos (centrales) de orden superior

$$\mu_r = E(X - \mu)^r \quad r = 1, 2, 3, \dots \quad (3)$$

La media está localizada en el centro de la distribución. La dispersión de la distribución alrededor de su centro se mide como la desviación estándar,

$$\sigma = \mu_2^{1/2} = \{E(X - \mu)^2\}^{1/2} \quad (4)$$

El coeficiente de variación (CV), $C_v = \sigma/\mu$ expresa la dispersión de la distribución como una proporción de la media. Mayores momentos adimensionales $\mu_r/\mu_2^{r/2}$ los cuales también son usados, particularmente el coeficiente de sesgo (medida de asimetría) y la curtosis (medida de apuntamiento).

$$\gamma_1 = \frac{\mu_3}{\mu_2^{3/2}} \quad (5)$$

Donde: γ_1 : coeficiente de asimetría de la distribución, adimensional. μ_2 : segundo momento central o varianza poblacional. μ_3 : tercer momento central de la distribución.

$$\beta_2 = \frac{\mu_4}{\mu_2^2} \quad (6)$$

Donde: β_2 : coeficiente de curtosis de la distribución, adimensional. μ_2 : segundo momento central o varianza poblacional. μ_4 : cuarto momento central de la distribución. La ecuación (6) expresa la curtosis. Cuando se desee reportar el exceso de curtosis, debe emplearse:

$$\gamma_2 = \beta_2 - 3 \quad (6.1)$$

Donde: γ_2 : exceso de curtosis, adimensional. β_2 : coeficiente de curtosis. 3: curtosis de referencia de una distribución normal.

Se pueden calcular las cantidades análogas a partir de una muestra de datos $x_1, x_2, \dots, x_n$.

$$\bar{x} = n^{-1} \sum_{i=1}^n x_i \quad (7)$$

Los mayores momentos de la muestra

$$m_r = n^{-1} \sum_{i=1}^n (x_i - \bar{x})^r \quad (8)$$

Son estimadores razonables de μ_r , pero están sesgados. Usualmente, se usan estimadores de sin sesgo. En particular, σ^2, μ_3 y el cuarto $k_4 = \mu_4 - 3\mu_2^2$ son estimados insesgablemente según:

$$s^2 = (n-1)^{-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (9)$$

$$\tilde{m} = \frac{n^2}{(n-1)(n-2)} m_3 \quad (10)$$

$$\tilde{k}_4 = \frac{n^2}{(n-1)(n-2)} \left\{ \left(\frac{n+1}{n-1} \right) m_4 - 3m_2^2 \right\} \quad (11)$$

Respectivamente. Los estimadores de la muestra del CV, coeficiente de asimetría y coeficiente de curtosis son, respectivamente (Hosking y Wallis, 1997).

$$\hat{C}_v = s/x, \quad g = \tilde{m}_3/s^3, \quad k = \tilde{k}_4/s^4 + 3 \quad (12)$$

Valores estimados como la asimetría o la curtosis tienen desventajas o inconvenientes. En el caso de la asimetría, esta se puede expresar como:

$$g = \frac{n}{(n-1)(n-2)} * \frac{\sum (x_i - \bar{x})^3}{s^3} \quad (13)$$

Se puede observar que esta contiene la suma de cubos de las desviaciones desde la media y, por lo tanto, es propenso a error (Bedient et al., 2008), lo que se entiende como su diferencia con respecto a la asimetría teórica.

Los estimadores g y k pueden presentar sesgos importantes que dependen del tamaño muestral, por lo que, en muestras pequeñas típicas de proyectos, el sesgo poblacional de la distribución puede no reflejarse adecuadamente. Por ejemplo, con 20 datos de una log-normal de dos parámetros, el sesgo muestral no debería superar 4.47 (65% del sesgo poblacional). Las inferencias mediante el método de los momentos en distribuciones sesgadas resultan poco confiables (Hosking y Wallis, 1997).

2.2.2.2. Definición de L-Momentos

a) Momentos ponderados probabilísticos (MPP)

Los momentos L Hosking (1990), basado en el método del “momento ponderado por probabilidad” (MPP) para describir la forma de las distribuciones de probabilidad introducido por Greenwood et al. (1979), dado que favorecía a aquellas distribuciones en las que la forma inversa estaba definida de manera explícita, es decir, una variable aleatoria x puede representarse mediante la función F:

$$x = x(F) \quad (14)$$

Greenwood et al. (1979) observó que, aunque los momentos de esta función podían ser representados en sus parámetros, la relación inversa no podía ser derivada directamente, es decir la estimación de parámetros por momentos no era posible.

Sea X una variable aleatoria con una función de distribución acumulativa (FDA) $F_x(x)$, luego los MPP son (Greenwood et al., 1979)

$$M_{p,r,s} = E[X^p \{F_x(x)\}^r \{1 - F_x(x)\}^s] = \int_0^1 [x(F_x(x))]^p F_x(x)^r (1 - F_x(x))^s dF_x(x) \quad (15)$$

Siendo p, r y s números reales.

Una definición alternativa de los momentos ponderados probabilísticamente es:
(Govindaraju y Das, 2007).

$$F_x(x) = u, x = Fx^{-1}(u) = x(u) \text{ con } 0 \leq u \leq 1, \quad (16)$$

$$M_{p,r,s} = \int_0^1 [x(u)]^p u^r (1-u)^s du \quad (17)$$

Si $s = r = 0$ y p es un entero no negativo, entonces $M_{p,0,0}$ denota el momento convencional de orden p con respecto al origen; por lo tanto,

$$M_{1,0,0} = \int_0^1 x(u) du = E[X] \quad (18)$$

Expresiones alternativas de los MPP son, generalmente, obtenidas a partir de hacer r o s igual a cero tal que:

$$\alpha_s = M_{1,0,s} = \int_0^1 x(u)(1-u)^s du \quad (19)$$

$$\beta_r = M_{1,r,0} = \int_0^1 x(u)u^r du \quad (20)$$

Si $r = 0$, el lado izquierdo de la distribución tiene más peso (valor pequeño) porque la distribución de frecuencia observada $x(u)$ está ponderada por el complemento $1 - u$ de la función de distribución acumulada. Por otro lado, si $s = 0$, $x(u)$ es ponderada por la función de distribución acumulativa u , en cuyo caso se le da más peso al lado derecho de la distribución (valores grandes) (Grender, 1990).

$$[X^n] = \int_{-\infty}^{\infty} x^n f_x(x) dx \quad (21)$$

Los momentos convencionales elevan x a potencias mayores, mientras que los MPP α_s y β_r mantienen x en primera potencia y elevan u o $(1-u)$. Por ello, los momentos convencionales son más sensibles al tamaño muestral y pueden mostrar tendencias marcadas. En distribuciones sesgadas, una muestra finita suele no reflejar adecuadamente esa asimetría, esto es crítico en ingeniería, donde los registros suelen ser cortos frente a los periodos de retorno analizados.

Hosking y Wallis (1997) indican que, aunque los MPP han sido utilizados como base para estimar los parámetros de las distribuciones de probabilidad, estos son difíciles de interpretar directamente como medidas de escala y forma de una distribución de probabilidad, siendo esta información proporcionada bajo ciertas combinaciones de los MPP.

De la ecuación de β_r , este (β_r) puede ser representado como una combinación lineal de α_s y viceversa. Durante el desarrollo de los L- Momentos se tiene una muestra definida como cantidad “n” de elementos ordenados de forma ascendente (Hosking y Wallis, 1997), entonces el estimador de MPP a partir de esa muestra finita es: siendo $x_{j:n}$ el j-ésimo elemento de orden ascendente.

Considérese una muestra de tamaño n, ordenada de menor a mayor:

$$x_{1:n} \leq x_{2:n} \leq \dots \leq x_{n:n}$$

$$b_r = \frac{1}{n} \sum_{j=r+1}^n \frac{\binom{j-1}{r}}{\binom{n-1}{r}} x_{j:n} \quad (22)$$

Donde: b_r : estimador muestral del momento ponderado probabilísticamente de orden r.
r: orden del momento ponderado probabilísticamente; $r=0,1,2,3,\dots$; n: número total de observaciones de la muestra. j: posición de la observación dentro de la muestra ordenada. $x_{j:n}$: valor ubicado en la posición j de una muestra de tamaño n, ordenada de menor a mayor. $\binom{j-1}{r}$: coeficiente combinatorio del numerador. $\binom{n-1}{r}$: coeficiente combinatorio del denominador.

Del desarrollo de los primeros MPP se obtiene:

$$b_0 = \frac{1}{n} \sum_{j=1}^n x_{j:n} \quad (23)$$

Donde: b_0 : estimador MPP de orden cero; equivale a la media aritmética de la muestra.
n: número total de observaciones. $x_{j:n}$: observación situada en la posición j de la muestra ordenada.

$$b_1 = \frac{1}{n} \sum_{j=2}^n \frac{(j-1)}{(n-1)} x_{j:n} \quad (24)$$

Donde: b_1 : estimador MPP de primer orden. $j-1$: ponderación asociada a la posición de la observación. $n-1$: factor de normalización. $x_{j:n}$: observación situada en la posición j de la muestra ordenada.

$$b_2 = \frac{1}{n} \sum_{j=3}^n \frac{(j-1)(j-2)}{(n-1)(n-2)} x_{j:n} \quad (25)$$

Donde: b_2 : estimador MPP de segundo orden. $(j-1)(j-2)$: ponderación del numerador. $(n-1)(n-2)$: factor de normalización. $x_{j:n}$: observación situada en la posición j de la muestra ordenada.

$$b_r = \frac{1}{n} \sum_{j=r+1}^n \frac{(j-1)(j-2)\dots(j-r)}{(n-1)(n-2)\dots(n-r)} x_{j:n} \quad (26)$$

Donde: b_3 : estimador MPP de tercer orden. $(j-1)(j-2)(j-3)$: ponderación del numerador. $(n-1)(n-2)(n-3)$: factor de normalización. $x_{j:n}$: observación situada en la posición j de la muestra ordenada.

b) Estimado de los L-Momentos

Las estimaciones de los Momentos L están linealmente relacionadas con los valores de MPP; por lo tanto, son una función lineal de la muestra original. Por lo tanto, en la práctica es posible estimar los momentos L de una muestra finita. A partir de los MPP, los primeros cuatro Momentos L (l_r), están determinados por:

$$l_1 = b_0 \quad (27)$$

$$l_2 = 2b_1 - b_0 \quad (28)$$

$$l_3 = 6b_2 - 6b_1 + b_0 \quad (29)$$

$$l_4 = 20b_3 - 30b_2 + 12b_1 - b_0 \quad (30)$$

El momento lineal de primer orden l_1 es el parámetro de localización de la muestra o el promedio (media aritmética), el de segundo orden l_2 mide la escala o la variación, representa la magnitud de dispersión de los datos, el momento de tercer orden

l_3 hace referencia a su simetría, y el de cuarto orden l_4 indica la curtosis de la muestra o el máximo de distribución.

Para calcular los L-Momentos independientes de las unidades de medición son los conocidos como cocientes o ratios de Momentos L (Hosking J. R., 1990). Estos se obtienen dividiendo aquellos de mayor orden por la medida de escala l_2 :

$$t = \frac{l_2}{l_1} = L - C_v, \quad \text{donde } 0 \leq t \leq 1 \quad (31)$$

$$t_3 = \frac{l_3}{l_2} = L - C_s, \quad \text{donde } -1 \leq t_3 \leq 1 \quad (32)$$

$$t_4 = \frac{l_4}{l_2} = L - C_k, \quad \text{donde } -1 \leq t_4 \leq 1 \quad (33)$$

Siendo, t coeficiente de variación, t_3 coeficiente de asimetría, t_4 coeficiente de curtosis, $(t, t_3 \text{ y } t_4)$ momentos lineales de muestra ponderados proporcionalmente en función de la longitud para el cálculo de las ratios medio regionales de los L-Momentos, n_i longitud del registro para el sitio i . Donde t^R, t_3^R, t_4^R se determina como:

$$t_r^R = \frac{\sum_{i=1}^N n_i t_r^i}{\sum_{i=1}^N n_i} \quad (34)$$

Estos se relacionan con los momentos lineales de la función de distribución de probabilidad para determinar los parámetros de las distribuciones. Hosking y Wallis (1997), aseguran el uso de los L-Momentos tiene varias ventajas, tales como:

- Capacidad de caracterizar un amplio rango de funciones de distribución;
- Alta estabilidad en presencia de desviaciones;
- Menor sensibilidad a la asimetría, etc.

2.2.3. Delimitación e identificación de regiones homogéneas

El procedimiento busca conformar grupos de estaciones homogéneas, condición que se cumple cuando las distribuciones de frecuencia son iguales entre sí. Para ello se emplean estadísticos derivados de las PMDA y atributos del área de estudio, especialmente la latitud, la longitud y la altitud, y se recomienda utilizar solo atributos

de estación para definir regiones (Hosking y Wallis, 1997). Asimismo, se aplica un agrupamiento preliminar con Ward y Joe (1963) y K-Means de MacQueen (1967), cuyas regiones se validan luego mediante una prueba de heterogeneidad basada en Momentos L, comparando la variación observada entre estaciones con la esperada en una región homogénea.

2.2.3.1. Selección de los atributos de las estaciones

Los factores que más influyen en la meteorología son la latitud, la altitud, la Cordillera de los Andes, la Corriente Costera Peruana (de aguas frías), el Anticiclón del Pacífico Sur y la continentalidad. Para este proyecto, se han seleccionado dos de estos factores: la latitud y la altitud, además de la longitud, ya que son características cuantificables y fáciles de obtener de las estaciones (Hosking y Wallis, 1997).

2.2.3.2. Análisis de agrupamiento según Word

Ward y Joe (1963) definieron un método de proporcionalidad entre los miembros de un grupo en una colección de variables. Su intención fue agrupar cada posible n-número de grupos, n, n-1, ..., 1, de modo que la pérdida de información en ese agrupamiento fuera mínima y pudiera medirla él mismo.

De un conjunto de 10 valores {2, 6, 5, 6, 2, 2, 2, 0, 0, 0}, una forma de representar los valores con un solo valor es su media: 2.5. La pérdida de información, dice Ward, se muestra en la suma de cuadrados del error (ESS) dada por:

$$ESS = \sum_{i=1}^n x_i^2 - \frac{1}{n} (\sum_{i=1}^n x_i)^2 \quad (35)$$

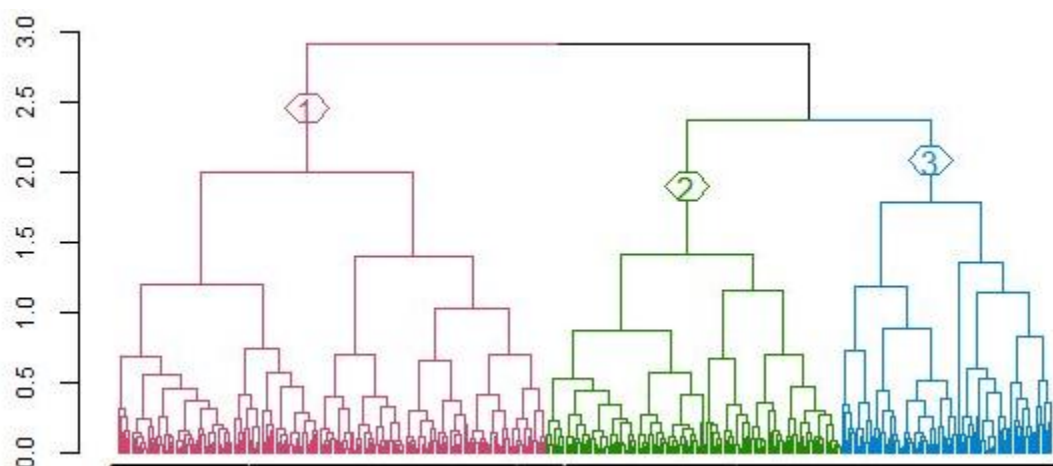
Donde x_i es el i-ésimo valor y n es el número de valores. En el ejemplo, n=10, ya que el ESS es 50.5. Ahora, si tomamos 4 grupos como {0, 0, 0}, {2, 2, 2, 2}, {5}, {6, 6}, esta partición se puede probar en términos de la suma de las cuatro sumas de cuadrados del error.

$$ESS_{(cuatro\ grupos)} = ESS_{(Grupo\ 1)} + ESS_{(Grupo\ 2)} + ESS_{(Grupo\ 3)} + ESS_{(Grupo\ 4)} \quad (36)$$

A esta relación, Ward y Joe (1963) denominan función objetivo a la medida que guía el agrupamiento, usualmente minimizando una pérdida, cuyo valor evidencia la información que se pierde al tratar los datos como un solo grupo frente a varios. En el agrupamiento jerárquico, no se fija previamente el número de grupos: se inicia con n grupos unitarios y, paso a paso, se fusionan los dos grupos cuya unión produce el menor deterioro de la función objetivo, continuando hasta formar un único grupo; los conjuntos resultantes se denominan grupos jerárquicos (Ward y Joe, 1963). La representación gráfica se realiza con dendrograma, donde el eje x muestra las estaciones y el eje y la distancia; cada fusión se identifica por una unión vertical y la altura de la barra horizontal indica el nivel de diferencia entre los grupos.

Figura 1

Diagrama de dendrograma para la jerarquización de estaciones



Nota. La función objetivo está limitada por la distancia Euclidiana (al cuadrado) entre los atributos seleccionados de las estaciones (latitud, longitud y altitud) (De-La-Hoz et al., 2019).

2.2.3.3. Análisis de agrupamiento según el método K-means

Este método es usado para particionar un conjunto de datos en k grupos de forma automatizada. Se da una colección de puntos $x_1, x_2, \dots, x_n$ el objetivo es encontrar la varianza de las agrupaciones de datos sabiendo de antemano que se tiene que formar k grupos, es decir, la función objetivo es minimizada, donde $d(x_i, m_j)$ es la

distancia euclidiana de x_i a m_j . Los puntos $\{m_j\}$ ($j= 1, 2, \dots, k$) son los centroides de los clústeres o grupos. El método K-means es una iteración de actualización de centroides para minimizar la función objetivo hasta que logre un mínimo local, el procedimiento paso a paso se muestra a continuación: (MacQueen, 1967)

$$\frac{1}{n} \sum_{i=1}^n [\min_j d^2(x_i, m_j)] \quad (37)$$

El método K-means permite formar múltiples grupos a partir de un número de clústeres previamente asumido (definido con base en el agrupamiento jerárquico de Ward). Inicialmente, asigna cada observación al clúster más cercano según la distancia euclidiana, estima los centroides y delimita las áreas de influencia mediante regiones de Voronoi. Luego, recalcula los centroides y reasigna los registros al grupo más similar de forma iterativa hasta que los centroides dejan de desplazarse (Hosking y Wallis, 1997).

En el proyecto se emplea la distancia euclidiana al cuadrado con las variables latitud, longitud y altitud. El número de grupos se definió a partir del agrupamiento jerárquico mediante el método de Ward y, una vez fijado, los centroides se calculan automáticamente. Finalmente, se evalúa si esta estrategia mejora el desempeño en comparación con el enfoque de Ward en el caso de estudio.

El resultado de un análisis de agrupamiento no es determinístico. Se pueden realizar algunas mejoras para aumentar la consistencia física de las regiones y reducir su heterogeneidad, introduciendo un valor denominado medida de heterogeneidad (H). En general, pueden realizarse algunos de los siguientes cambios (Hosking y Wallis, 1997).

- Mover una estación o varias estaciones de una región a otra;
- Excluir una estación o un conjunto de estaciones de los datos;
- Subdividir una región;

- Fragmentar una región reasignando sus estaciones a otras regiones;
- Unir una región con una o más regiones;
- Unir dos o más regiones y redefinir los grupos.
- Obtener más información o datos y volver a definir los grupos.

2.2.4. *Análisis de datos: Pruebas de discordancia y heterogeneidad regional (ARF)*

Para verificar la confiabilidad de una muestra regional, se aplicó un Análisis Exploratorio de Datos (AED) basado en la consistencia (tendencia y homogeneidad en estaciones individuales) y en la discordancia (anomalías dentro de grupos de estaciones). La tendencia se evalúa mediante la prueba no paramétrica de Yue et al. (2002). La homogeneidad se analiza con la prueba no paramétrica de Pettitt (1979) contrastando la hipótesis de ausencia de cambio frente a la alternativa de cambio. Finalmente, la discordancia identifica series no confiables por presentar regímenes de frecuencia distintos a los del grupo, evaluando las anomalías con base en los L-momentos.

a) Análisis de tendencia

Para el análisis de tendencia se aplica la prueba no paramétrica de Yue et al. (2002), adecuada para series hidrológicas con distribución no normal y registros incompletos, situación frecuente en los datos de tiempo. El método se desarrolla mediante:

Paso 1: La pendiente b de la tendencia de la serie temporal es estimada según el método de Theil – Sen (TSA):

$$b = \text{Mediana} \left(\frac{x_j - x_i}{j - i} \right), 1 \leq i < j \leq n \quad (38)$$

Donde: b : pendiente estimada mediante el método de Theil–Sen. x_i : observación correspondiente al instante i . x_j : observación correspondiente al instante j . i, j : posiciones temporales de las observaciones, con $i < j$. n : número total de observaciones

de la serie. Mediana: valor central del conjunto de pendientes obtenidas entre pares de observaciones.

Siendo X_i la observación i-ésima. Si la pendiente es cercana a cero, no es necesario continuar con el análisis de tendencias. Si es distinto de cero, se asume que hay una tendencia lineal, siendo anulada la tendencia de la muestra según: siendo t el tiempo transcurrido entre las observaciones j y i .

Serie libre de tendencia:

$$X'_t = X_t - T_t = X_t - bt \quad (39)$$

Donde: X'_t : valor de la serie sin tendencia en el instante t . X_t : valor original observado en el instante t . T_t : componente lineal de tendencia en el instante t . b : pendiente estimada mediante el método de Theil–Sen. t : índice temporal.

Paso 2: El coeficiente de correlación de la serie de lag-1, r_1 , de la serie sin tendencia X'_t se calcula utilizando la siguiente ecuación: siendo n es el tamaño de la muestra, luego el modelo de autorregresión AR(1) es anulado de X'_t según:

Coeficiente de autocorrelación serial de primer orden:

$$r_1 = \frac{\sum_{t=1}^{n-1} [X'_t - \bar{X'}][X'_{t+1} - \bar{X'}]}{\sum_{t=1}^n [X'_t - \bar{X'}]^2} \quad (40)$$

Donde: r_1 : coeficiente de autocorrelación serial de primer orden. X'_t : valor de la serie libre de tendencia en el instante t . X'_{t+1} : valor de la serie libre de tendencia en el instante posterior. $\bar{X'}$: media aritmética de la serie sin tendencia. n : número total de observaciones. t : índice temporal.

Preblanqueo de la serie sin tendencia:

$$Y'_t = X'_t - r_1 X'_{t-1} \quad t = 2, 3, \dots, n \quad (41)$$

Donde: Y'_t : residuo obtenido después de retirar la autocorrelación serial. X'_t : valor de la serie sin tendencia en el instante t . X'_{t-1} : valor de la serie sin tendencia en el instante anterior. r_1 : coeficiente de autocorrelación serial de primer orden.

Este proceso es conocido como TFPW (pre-blanqueo sin tendencia). La serie residual, tras aplicar el procedimiento TFPW, debería ser independiente.

Paso 3: Se realiza la combinación entre la tendencia T_t y el residual Y'_t : de esta manera la serie Y_t conserva la tendencia original sin ser influenciada por efectos de la autocorrelación.

Recomposición de la serie:

$$Y_t = Y'_t - T_t = Y'_t + bt \quad (42)$$

Donde: Y_t : serie recompuesta, libre del efecto de autocorrelación serial de primer orden. Y'_t : residuo obtenido mediante el preblanqueo. T_t : componente lineal de tendencia. b : pendiente estimada mediante Theil–Sen. t : índice temporal.

Paso 4: Finalmente, se aplica la prueba de Mann – Kendall a la nueva serie para evaluar la significancia de la tendencia.

La prueba no paramétrica Mann–Kendall se usa para evaluar la significancia de tendencias en series temporales hidrometeorológicas y es adecuada cuando los datos no siguen distribución normal, como ocurre con precipitaciones extremas. Para aplicarla, la serie debe cumplir independencia serial (Yue et al., 2002).

Queremos estimar la importancia del patrón de intensas lluvias a un nivel de confianza dado. La prueba se lleva a cabo en la serie de tiempo de precipitaciones máximas diarias anuales X_i en el rango $i = 1, \dots, n - 1$, y X_j en el rango $j = i + 1, \dots, n$. Cada punto X_i se toma como referencia y se compara con el resto de los puntos X_j de la manera que El estadístico S de Kendall es calculado como:

Estadístico S

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{sgn}(x_j - x_i) \quad (43)$$

S: estadístico de Mann–Kendall. X_i : observación en el instante i. X_j : observación situada en el instante j, con $j > i$. n : número total de observaciones. $\text{sgn}(\cdot)$: función signo.

Función signo

$$sgn(x_j - x_i) = \begin{cases} 1, & \text{Si } x_j - x_i > 0 \\ 0, & \text{Si } x_j - x_i = 0 \\ -1, & \text{Si } x_j - x_i < 0 \end{cases} \quad (44)$$

Donde: $sgn(X_j - X_i)$: signo de la diferencia entre dos observaciones. $X_j - X_i$: variación entre una observación posterior y otra anterior.

Cuando los datos son IID (independientes e idénticamente distribuidos), se cumple que la esperanza del estadístico S es 0 y su varianza puede definirse como: $Var(S)$, donde t_i representa el número de valores empatados (iguales) en el i-ésimo grupo de empates. En consecuencia, el estadístico estandarizado ZS se calcula como:

Varianza del estadístico S:

$$Var(S) = \frac{n(n-1)(2n+5) - \sum_{p=1}^g t_p(t_p-1)(2t_p+5)}{18} \quad (45)$$

Donde: $Var(S)$: varianza del estadístico Mann–Kendall. n : número total de observaciones. g : número de grupos de valores empatados. t_p : número de observaciones incluidas en el grupo de empates p . p : índice de cada grupo de empates.

Estadístico estandarizado Zs:

$$Z_s = \begin{cases} \frac{S-1}{\sqrt{Var(S)}}, & \text{para } S > 0 \\ \frac{(S+1)}{\sqrt{Var(S)}}, & \text{para } S < 0 \\ 0, & \text{para } S = 0 \end{cases} \quad (46)$$

Donde: ZS: estadístico normal estandarizado de Mann–Kendall. S: estadístico de Mann–Kendall. $Var(S)$: varianza del estadístico S. 1: corrección de continuidad.

Z_s sigue la distribución normal estándar. El grado de significancia p puede ser obtenido usando la prueba estadística en la función de distribución acumulativa (F.D.A.) para una variable estándar normal F_N :

Valor p para una prueba bilateral:

$$p = 2[1 - \Phi(|Z_s|)] \quad (47)$$

Donde: p : probabilidad asociada con la prueba bilateral. $\Phi(\cdot)$: función de distribución acumulada de una variable normal estándar. $|Z_s|$: valor absoluto del estadístico normal estandarizado.

Cuando $|Z_s| > Z_{\alpha/2}$, la tendencia se considera significativa al nivel de significancia α . El valor p mide la probabilidad de rechazar la hipótesis nula de “sin tendencia”; cuanto menor es p , mayor evidencia de tendencia. Con un 95% de confianza, se asume un 5% de error (probabilidad de concluir que existe una tendencia cuando en realidad no la hay). El criterio equivalente es: $|Z_s| > 1.96$

b) Análisis de homogeneidad

La prueba no paramétrica de Pettitt identifica un posible punto de cambio abrupto en una serie temporal. Se basa en el estadístico de Mann–Whitney y permite evaluar la hipótesis nula de ausencia de cambio. La fuente original es Pettitt (1979).

Estadístico acumulado de Pettitt:

$$U_{t,T} = \sum_{i=1}^t \sum_{j=i+1}^T \text{sgn}(X_i - X_j), 1 \leq t \leq T \quad (48)$$

Donde: $U_{t,T}$: estadístico acumulado de Mann–Whitney para el posible punto de cambio t . X_i : observación perteneciente al segmento anterior o al instante t . X_j : observación correspondiente al segmento posterior al instante t . t : posición evaluada como posible punto de cambio. T : número total de observaciones. $\text{sgn}(\cdot)$: función signo.

Función signo para la prueba Pettitt

$$\text{sgn}(X_i - X_j) = \begin{cases} 1, & \text{Si } X_i - X_j > 0 \\ 0, & \text{Si } X_i - X_j = 0 \\ -1, & \text{Si } X_i - X_j < 0 \end{cases} \quad (49)$$

Donde: $\text{sgn}(X_i - X_j)$: signo de la diferencia entre dos observaciones. $X_i - X_j$: diferencia entre una observación anterior y otra posterior al punto evaluado.

Forma recursiva del estadístico:

$$U_{t,T} = U_{t-1,T} + \sum_{j=1}^t \text{sgn}(X_i - X_j), \quad t = 2, 3, \dots, T \quad (50)$$

Donde: $U_{t,T}$: estadístico acumulado en el instante t . $U_{t-1,T}$: estadístico acumulado en el instante anterior. X_t : observación correspondiente al instante evaluado. X_j : observación situada en la posición j . T : número total de observaciones.

El estadístico $U_{t,T}$ es un recuerdo del estadístico de Mann – Whitney para la prueba de igualdad de dos muestras $X_1, \dots, X_t$ y $X_{t+1}, \dots, X_T$ con respecto a la misma población. Calculamos el estadístico $U_{t,T}$ para la ventana de t -valores: $1 \leq t < T$. Para probar H sin cambio contra la alternativa A con cambio, Pettitt (1979) sugiere el uso del estadístico:

Estadístico bilateral de Pettitt:

$$K_T = \max_{1 \leq t < T} |U_{t,T}| \quad (51)$$

Donde: K_T : estadístico bilateral de Pettitt. $|U_{t,T}|$: valor absoluto del estadístico acumulado. t : posición evaluada como posible punto de cambio. T : número total de observaciones.

Para los cambios en una dirección se consideran los estadísticos:

Estadístico para un cambio ascendente:

$$K_T^+ = \max_{1 \leq t < T} U_{t,T} \quad (52)$$

Donde: K_T^+ : estadístico unilateral asociado con un posible cambio ascendente. $U_{t,T}$: estadístico acumulado evaluado en cada instante.

Estadístico para un cambio descendente:

$$K_T^- = - \min_{1 \leq t < T} U_{t,T} \quad (53)$$

Esta misma expresión puede escribirse como:

$$K_T^- = - \min_{1 \leq t < T} (-U_{t,T}) \quad (53.1)$$

Donde: K_T^- : estadístico unilateral asociado con un posible cambio descendente. $U_{t,T}$: estadístico acumulado evaluado en cada instante.

Según la teoría estadística de Mann – Whitney, el valor de K con subíndice T a la más será alto si hay un incremento abrupto en la serie, es decir, si F con subíndice 1 , se abre paréntesis x se cierra paréntesis, es menor o igual a F conann – Whitney, el valor de K_T^+ será alto si hay un incremento abrupto en la serie, es decir, si $F_1(x) \leq F_2(x)$. Del mismo modo, el valor de K_T^- será alto si hay un decaimiento brusco o $F_1(x) \geq F_2(x)$.

El nivel de confianza asociada con K_T^+ o K_T^- se determina con:

Probabilidad aproximada de la prueba bilateral

$$p \approx 2 \exp\left(-\frac{6K_T^2}{T^3+T^2}\right) \quad (54)$$

Donde: p: probabilidad aproximada asociada con la prueba bilateral de Pettitt. KT: estadístico bilateral de Pettitt. T: número total de observaciones. exp(·): función exponencial.

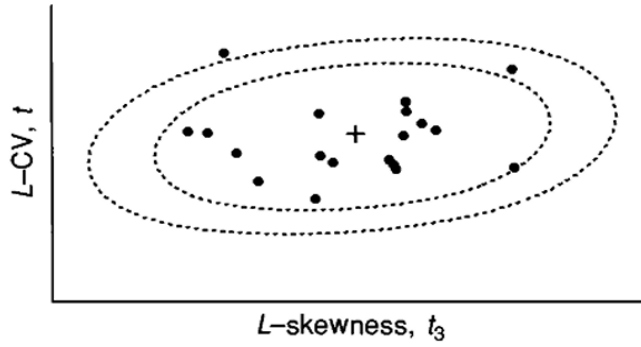
Cuando ρ es menor que el nivel de confianza dado (nivel del 95%) la hipótesis nula se rechaza. La probabilidad aproximada de significancia para el sal1 menos ro . Claramente, la serie rompe en el punto de salto en dos subseries

c) Análisis de discordancia

El análisis de discordancia busca identificar estaciones con datos poco confiables o atípicos respecto a la distribución de frecuencia, evaluando anomalías en los Momentos L de la muestra. Para ello, se calcula un estadístico de discordancia que compara las ratios de Momentos L de cada estación con el promedio regional de un grupo homogéneo, lo que permite detectar valores atípicos, tendencias o desviaciones. Gráficamente, las estaciones se representan en un diagrama de dispersión L-Cv vs L-Cs, con el promedio como centro y elipses concéntricas basadas en la covarianza; se consideran discordantes los puntos fuera de la elipse exterior (Hosking y Wallis, 1997).

Figura 2

Esquema de nube de puntos de ratios de Momentos L



Nota. (Hosking y Wallis, 1997).

Considerando que hay N estaciones en el grupo y que es un vector que contiene los valores t, t_3, t_4 correspondientes a la estación i (donde T indica la transpuesta del vector), sea el promedio grupal, se define la matriz de suma de cuadrados y productos,

$$u_i = [t^{(i)}, t_3^{(i)}, t_4^{(i)}]^T \quad (55)$$

$$\bar{u} = N^{-1} \sum_{i=1}^N u_i \quad (56)$$

$$A = \sum_{i=1}^N (u_i - \bar{u})(u_i - \bar{u})^T \quad (57)$$

Se defina la media de discordancia para la estación i de la siguiente manera (Hosking y Wallis, 1997).

$$D_i = \frac{1}{3} N (u_i - \bar{u})^T A^{-1} (u_i - \bar{u}) \quad (58)$$

Hosking y Wallis (1997) sugieren que una estación sea considerada discordante en función del número de estaciones, y si D_i supera valor crítico de la Tabla 1.

Tabla 1*Valores críticos para la discordancia D_i*

Número de estaciones en la región de análisis	Valor Crítico	Número de estaciones en la región de análisis	Valor Crítico
5	1.333	10	2.491
6	1.648	11	2.632
7	1.917	12	2.757
8	2.140	13	2.869
9	2.392	14	2.971
		≥ 15	3

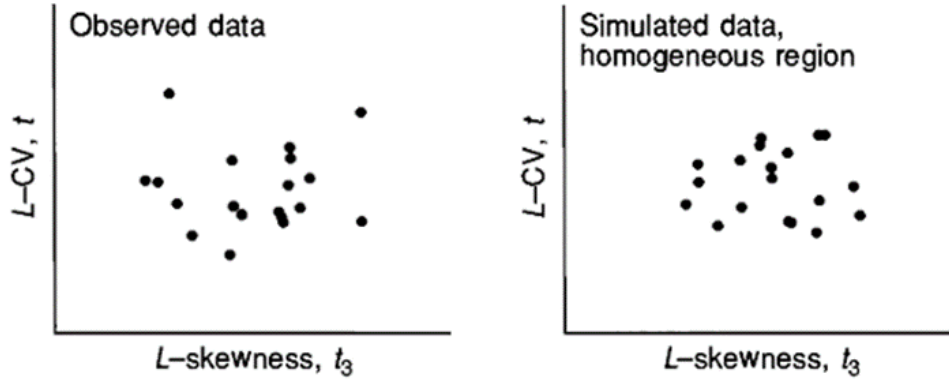
Nota. (Hosking y Wallis, 1997).**d) Medida de heterogeneidad**

La prueba de heterogeneidad busca determinar si un subconjunto de estaciones puede considerarse parte de una región homogénea, comparando la variabilidad inter-estaciones de las ratios de Momentos L con la esperada en grupos homogéneos. En este marco, una región homogénea implica distribuciones de frecuencia con forma idéntica entre estaciones; para operacionalizarlo se recurre a simulación Monte Carlo de muestras sintéticas. Para evitar elegir distribuciones de 2 o 3 parámetros, Hosking y Wallis (1997) recomiendan simular con la distribución kappa de cuatro parámetros, por su alta flexibilidad para representar diversos tipos de distribuciones.

Por ejemplo, si se tiene una región con longitudes de registro de estaciones de 20, 30, 10, 50, etc., se generará una región sintética de la misma cantidad de estaciones y de la misma longitud de registro de estaciones. Una inspección visual de la variabilidad de los cocientes de Momentos L en las estaciones se puede realizar con gráficos como $L - C_s$ vs. $L - C_v$ (ver Figura 3) y $L - C_s$ vs. $L - C_k$ (Hosking y Wallis, 1997).

Figura 3

Ratios de Momento – L para datos observados y simulado



Nota. (Hosking & Wallis, 1997)

Ahora, sea el área propuesta que contenga N estaciones, con la longitud de registro de la estación i igual a n_i y los coeficientes de L-momentos de la muestra iguales a $t^{(i)}, t_3^{(i)}$ y $t_4^{(i)}$. Los promedios regionales $L - C_v$, $L - C_s$ y $L - C_k$ ponderados por la longitud de registro de las estaciones se pueden escribir como: (Hosking y Wallis, 1997)

$$t^R = \frac{\sum_{i=1}^N n_i t^{(i)}}{\sum_{i=1}^N n_i} \quad (59)$$

La desviación ponderada en la estación de los $L - C_v$ s de la muestra es:

$$V = \left\{ \frac{\sum_{i=1}^N n_i (t^{(i)} - t^R)^2}{\sum_{i=1}^N n_i} \right\}^{1/2} \quad (60)$$

El método simula gran número N_{sim} valores de región de N estaciones con distribución de frecuencia kappa. De las simulaciones se obtiene la media (μ_v) y la desviación estándar (σ_v) de los N_{sim} valores de V . La heterogeneidad se estima como:

$$H = \frac{(V - \mu_v)}{\sigma_v} \quad (61)$$

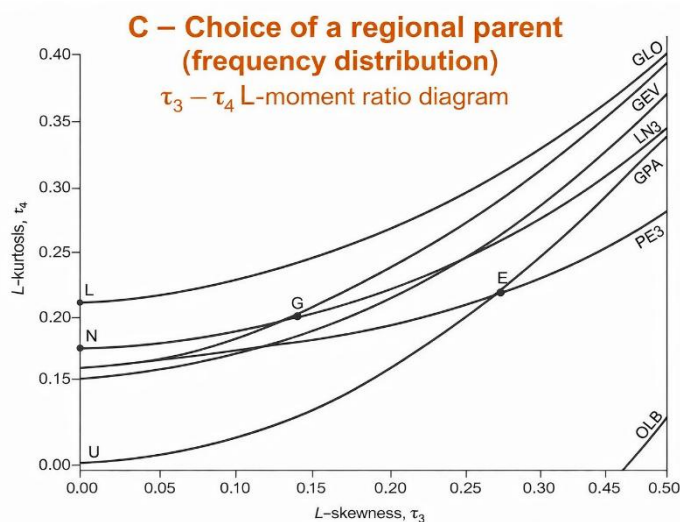
Hosking y Wallis (1997) proponen que la región sea considerada “aceptablemente homogénea” si $H < 1$, “posiblemente heterogénea” si $1 \leq H \leq 2$ y “definitivamente heterogénea” si $H \geq 2$.

2.2.5. Selección y ajuste de distribuciones de probabilidad

En análisis regional, las áreas solo pueden aproximarse mediante una distribución de frecuencia, aunque siempre existe cierta heterogeneidad; por ello, no se busca la “distribución verdadera”, sino una que produzca cuantiles adecuados para cada área (Hosking y Wallis, 1997). La selección de la distribución se apoya en gráficos de cocientes de L-momentos: para modelos de dos parámetros se usa L-Cs vs. L-Cv, pero por la variabilidad entre estaciones suele preferirse una distribución más flexible, evaluada con el diagrama t_3 vs. t_4 (L-Cs vs. L-Ck). En estos gráficos, t_3 y t_4 siguen relaciones teóricas propias de distribuciones de tres parámetros, permitiendo inferir la curtosis a partir de la asimetría. Además, se identifican casos especiales: Gumbel como caso particular de GEV (forma = 0), Normal como caso simétrico de PE3 y LN3, y Logística (2 parámetros) como caso especial de GLO.

Figura 4

Ejemplo de relaciones entre L - C_S– L - C_K para distribuciones de frecuencia



Nota. (Castellarin et al., 2007).

Hosking y Wallis (1997) proponen usar el diagrama $t_3 - t_4$ para seleccionar la distribución, comparando las relaciones teóricas con las relaciones muestrales de las estaciones del área; el centro del conjunto de puntos se adopta como promedio regional.

En GEV, el parámetro k se estima por Momentos-L igualando T_3 teórico con t_3^R regional, ubicando el punto sobre la curva; este ajuste corrige hasta orden 3 (asimetría). Para evaluar el ajuste de orden 4 (curtosis), se utiliza medida de distancia Z entre ratios regionales observadas y modeladas. Si Z supera el umbral se rechaza el modelo, y si es menor se acepta la GEV como representativa del comportamiento medio regional.

Se supone que se tiene una región con N estaciones, con la estación i teniendo longitud de registro n_i y los coeficientes de las muestras de momento L $t^{(i)}, t_3^{(i)}$ y $t_4^{(i)}$. Llámese a las medias $L - C_v, L - C_s$ y $L - C_k, t^R, t_3^R$ y t_4^R respectivamente, donde están ponderados en proporción a la longitud de registro. Algunas de las distribuciones son logística generalizada (GLO), valor extremo generalizado (GEV), Pareto generalizado (GPA), lognormal (LN3) y Pearson III (PE3) (Hosking y Wallis, 1997).

Cada distribución se ajusta a los Momentos L regionales promedio 1, t^R, t_3^R . El $L - C_k$ de la distribución ajustada se denota por T_4^{DIST} , donde DIST puede ser GLO. Luego se ajusta una distribución kappa a los coeficientes de los Momentos L regionales promedio 1, t^R, t_3^R y t_4^R . Se generan N_{sim} realizaciones de una región con N estaciones, cada una con esta distribución kappa. Las regiones simuladas son homogéneas y tienen las mismas longitudes de registro que sus contrapartes reales. Para la m -ésima región simulada, se estiman los promedios regionales $L - C_s t_3^{[m]}$ y $L - C_k t_4^{[m]}$. Ahora se estima el sesgo de t_4^R , desviación estándar de t_4^R , y, para cada distribución, la prueba de bondad: según Hosking y Wallis (1997) se satisface la prueba cuando $|Z^{DIST}| \leq 1.64$.

$$B_4 = N_{sim}^{-1} \sum_{m=1}^{N_{sim}} (t_4^{[m]} - t_4^R) \quad (62)$$

$$\sigma_4 = \left[(N_{sim} - 1)^{-1} \left\{ \sum_{m=1}^{N_{sim}} (t_4^{[m]} - t_4^R)^2 - N_{sim} B_4^2 \right\} \right]^{1/2} \quad (63)$$

$$Z^{DIST} = \frac{\tau_4^{DIST} - t_4^R + B_4}{\sigma_4} \quad (64)$$

2.2.6. Región de Influencia (ROI)

La Región de Influencia (ROI) es un enfoque de Análisis Regional de Frecuencia que construye “regiones locales” alrededor de una estación objetivo, en lugar de trabajar con regiones fijas predefinidas. Su lógica es que, en dominios hidrológicamente complejos, cada estación puede requerir un conjunto distinto de estaciones vecinas para lograr un agrupamiento estadísticamente consistente. En la práctica, ROI se implementa como un método de vecindad que selecciona estaciones similares según atributos (fisiografía, clima, etc.) y luego aplica el esquema regional (p. ej., L-momentos) sobre ese subconjunto (Ahmed et al., 2025).

a) Fundamento del enfoque ROI y diferencias con ARF “fijo”

A diferencia del ARF “fijo”, donde una región comparte un único pool de datos, en ROI la región se redefine para cada estación (en espacio geográfico o de atributos), lo que puede reducir sesgos por heterogeneidad, aunque añade decisiones (métrica, variables, tamaño del vecindario) que afectan los cuantiles (Ahmed et al., 2025).

b) Construcción de grupos ROI por estación objetivo

La construcción de grupos ROI suele seguir una secuencia operativa: (i) definir un conjunto de atributos de similitud (altitud, coordenadas, descriptores climáticos), (ii) estandarizar variables para evitar dominancias por escala, (iii) calcular distancias/similitudes entre la estación objetivo y las candidatas, (iv) conformar el vecindario eligiendo las k estaciones más cercanas, y (v) verificar la coherencia/homogeneidad del grupo antes de estimar cuantiles regionales (Zaghloul et al., 2024).

c) Sensibilidad a la distribución en ROI (curvas por estación objetivo)

Los cuantiles regionales pueden ser sensibles a las medidas de similitud usadas para agrupar estaciones, lo que obliga a justificar técnicamente el criterio de vecindad y a evaluar el ajuste de distribuciones con mayor rigor (Zaghloul et al., 2024).

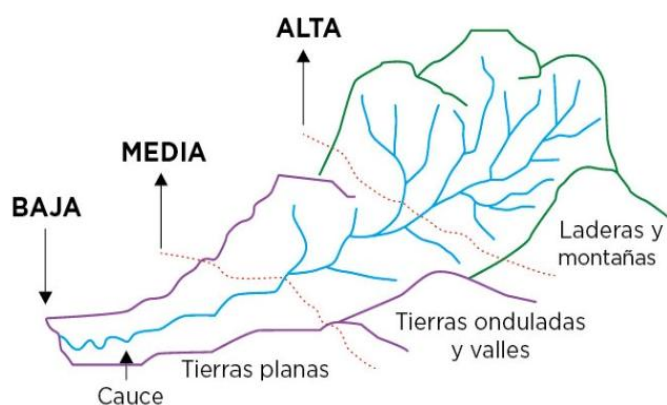
2.3. Definiciones conceptuales

2.3.1. Cuenca hidrográfica y sus propiedades geomorfológicas

Una cuenca hidrográfica es el área de terreno donde las aguas de precipitación, superficiales y subterráneas, convergen hacia un punto común, como el cauce o río principal (Puelles, 2015). Además, constituye una unidad natural en la que interactúan componentes hidrológicos, bióticos y antrópicos, funcionando como un sistema abierto e interconectado que puede verse afectado por el uso del suelo, el cambio climático y obras humanas como embalses o edificaciones ribereñas (Restrepo, 2005).

Figura 5

Partes de una cuenca



Nota. (Departamento General de Irrigación, 2016).

Tabla 2

Tamaño de cuenca según su extensión

Carácter	Área (Km ²)
Pequeña	< de 50
Mediana	de 50 a 150
grande	> de 150

Nota. Texto Básico de Hidrología (Gamez, 2010).

De acuerdo a Guevara (2015), el análisis de una cuenca resulta importante por el hecho de que plantea establecer las propiedades geomorfológicas en función al aporte que manifiestan y el comportamiento hidrológico.

Área de la cuenca (A). Es la superficie que queda delimitada dentro del parteaguas (línea imaginaria que junta los puntos más elevados del terreno).

Ancho de la cuenca (w). Es la medida perpendicular respecto al eje mayor, cuyo valor depende de las relaciones entre las longitudes de las cuencas y su área.

Perímetro de la cuenca (P). Longitud total del parteaguas, en otro término, al borde de las formas de las cuencas que se ha proyectado en función de un plano horizontal.

Índice de capacidad (k). Relación por medio del área y el perímetro de las cuencas respecto a las escorrentías, correspondientes a las particularidades del hidrograma.

$$k = 0.28 \times \frac{P}{\sqrt{A}} \quad (65)$$

Tabla 3

Índice de compacidad según la forma de la cuenca

Valores de K	Tipos de clases de forma
De 1.00 a 1.25	De casi redonda a oval redonda
De 1.26 a 1.50	De oval redonda a oval oblonga
De 1.51 a 1.75	De oval oblonga a rectangular oblonga

Nota. Texto Básico de Hidrología (Gamez, 2010)

Pendiente de la cuenca (Sm). Inclinação del terreno en la superficie de una cuenca hidrográfica, que determina la velocidad y la dirección del flujo del agua.

$$S_m = \frac{H}{L_m} \quad (66)$$

Donde, Sm relación entre el desnivel total (H), la cual se expresa en km, y el lado mayor del rectángulo equivalente (Lm), la cual se expresa en Km.

$$S_c = \frac{C}{A} \sum_{i=1}^n l_i \quad (67)$$

Donde, Sc es la pendiente de la cuenca, C es la equidistancia entre curvas de nivel, A es el área de la cuenca, li es la longitud de cada curva de nivel.

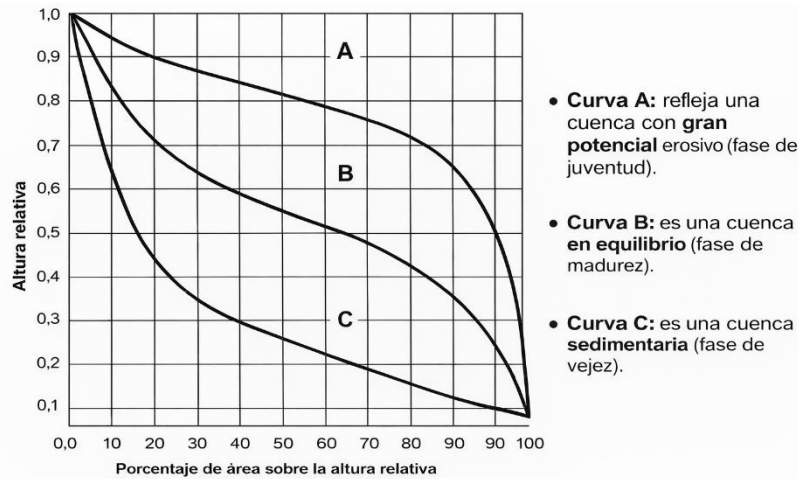
Altitud media de la cuenca. Medida que describe la altura promedio del terreno dentro de la cuenca y se utiliza para entender mejor las características topográficas de la zona.

$$H = \frac{1}{A} \sum_{i=1}^n H_i \times A_i \quad (68)$$

Donde, H es la altitud media en msnm, H_i es la altura del área acumulada A_i sobre la curva H_i , n es el número de áreas parciales, A es el área de la cuenca.

Figura 6

Tipos de formas de las curvas hipsométricas en una cuenca



Nota. (Rojas, 2015).

Número de orden de la cuenca. Índice que se utiliza para clasificar y ordenar las cuencas hidrográficas según su tamaño y jerarquía.

Longitud del máximo recorrido (L). Abarca la medición de las trayectorias de más trascendencia con respecto a cada partícula del flujo constituida por medio del emisor (punto del colector grandemente bajos) y el punto grandemente elevado.

Pendiente del cauce principal. Inclínación del lecho por donde fluye el curso de agua principal de una cuenca hidrográfica. Es la pendiente de la corriente principal que drena la cuenca y es un factor importante en la generación de velocidad y erosión del agua.

$$S_o = \left[\frac{\sum_{i=1}^{n_o} l p_i}{\sum_{i=1}^{n_o} \left(\frac{l p_i}{s_i} \right)^{1/2}} \right]^2 \quad (69)$$

Donde, S_o es la pendiente del cauce principal, $l p_i$ es la longitud de cada tramo de pendiente s_i , n_o es el número de tramos de similar pendiente.

2.3.2. Precipitaciones

Según Llamas (1993), la precipitación es la transferencia de agua entre la atmósfera y el suelo en fase líquida (lluvia) o sólida (nieve y granizo), donde una parte se evapora y el resto aporta escorrentía superficial o recarga subterránea; además, se clasifica en orográfica, convectiva y ciclónica; la orográfica ocurre por ascenso de masas de aire húmedas en zonas montañosas, mientras que la convectiva es breve pero intensa por calentamiento solar y ascenso del aire hasta condensar. Por su parte, la precipitación pluvial o lluvia resulta de la condensación del vapor de agua y presenta alta variabilidad por factores como humedad, temperatura y presión (Ricra, 2022).

a) Precipitaciones máximas

Según Villón (2004) la precipitación máxima probable es la cantidad máxima de precipitación que puede caer en un momento meteorológico posible en una cuenca, en un lugar determinado y en una época determinada del año.

b) Periodo de retorno

Intervalo promedio (en años) en el que un caudal pico de una creciente se iguala o supera al menos una vez. Asumiendo independencia entre eventos anuales, permiten estimar la probabilidad de excedencia y el riesgo de falla durante la vida útil de una obra en un horizonte de años determinado (MTC, 2018).

$$T = \frac{1}{1-(1-p)^{1/n}} \quad (70)$$

Donde, T periodo de retorno, p probabilidad de presentación del fenómeno en n años.

Tabla 4

Periodo de retorno según fuente hidrográfica y proyectos u obras

Nº	Tipo de proyecto u obra	Periodo de retorno (años)
1	Diques longitudinales [mediano riesgo]	50 a 100
2	Desarrollo de zona de inundación	100
3	Diseño de puentes (pilares)	100 a 500

Nota. (Ponce, 2008).

2.3.3. *Precipitaciones máximas diarias anuales (PMDA) observadas*

Las PMDA corresponden a la serie de máximos anuales construida a partir de registros diarios, donde para cada año hidrológico se extrae el mayor valor de precipitación diaria (24 h) por estación. Esta serie es la base para el análisis de extremos, porque conserva la información relevante de eventos críticos y permite estimar cuantiles de diseño mediante enfoques locales y regionales (Alejo-Sanchez et al., 2025).

a) Inventario y cobertura espacial de estaciones

El inventario de estaciones define la base observacional del estudio, integrando metadatos (tipo de estación, coordenadas, altitud, periodo de registro) y su relación espacial con la cuenca. La evaluación de cobertura permite identificar vacíos, redundancias y sesgos por concentración de estaciones en ciertos sectores. En este marco, la red de pluviómetros se analiza como un sistema cuya densidad y distribución influyen directamente en la estimación de extremos (Wang et al., 2022).

Ubicación de estaciones (altitud). La altitud es un descriptor en cuencas andinas porque controla procesos orográficos, gradientes térmicos y la organización espacial de la precipitación extrema. Por ello, la ubicación altitudinal de cada estación se incorpora para interpretar diferencias sistemáticas entre PMDA y para sustentar agrupamientos coherentes con condiciones fisiográficas (Fergusia et al., 2024).

Proximidad espacial. Se usa para evaluar similitud potencial de los máximos y para controlar redundancias o vacíos de información. En redes reales, estaciones muy cercanas pueden compartir señal meteorológica y reducir valor marginal de información, mientras que grandes separaciones incrementan incertidumbre de interpolación y de representación regional. Por ello, se emplean métricas de distancia (p. ej., vecino más cercano) y estructuras de dependencia espacial para sustentar criterios de agrupamiento o de influencia (Huang y Chen, 2024).

Distribución espacial. Evalúa si las estaciones están uniformemente repartidas o si presentan clustering. En términos prácticos, una red con distribución irregular puede inducir regiones “aparentes” por densidad de datos (Song y Kim, 2025).

b) Completitud y consistencia del registro

Requisitos previos para que las PMDA representen extremos reales y no artefactos de faltantes, cambios operativos o errores de medición. En series de precipitación, los faltantes y anomalías pueden distorsionar estadísticos, alterar pruebas de tendencia y afectar cuantiles de diseño, por lo que la gestión de calidad incluye verificación, depuración y trazabilidad de correcciones (Mohammed y Scholz, 2023).

Longitud del registro por estación (n años). Condiciona la estabilidad de la estimación de extremos, porque en series anuales los cuantiles dependen fuertemente del tamaño muestral. Registros cortos suelen exhibir mayor incertidumbre muestral y sensibilidad a valores atípicos, mientras que registros más largos estabilizan estadísticos y mejoran la representatividad climática (Alejo-Sanchez et al., 2025).

Años faltantes, % faltantes, % completos. Los faltantes se cuantifican mediante años sin PMDA, porcentaje de datos ausentes y proporción de años completos, porque estas métricas definen el grado de confiabilidad del registro para análisis de extremos (Fitriana et al., 2025).

c) Caracterización estadística de PMDA

La caracterización estadística resume el comportamiento de las PMDA por estación y permite comparar magnitud y variabilidad antes de regionalizar. En análisis de extremos, estos estadísticos no sustituyen el modelamiento probabilístico, pero sí permiten detectar inconsistencias, outliers plausibles/no plausibles y diferencias sistemáticas entre zonas. Además, constituyen la base para interpretar por qué ciertas estaciones se agrupan o se separan en regiones homogéneas (Achite et al., 2024).

Media / Mediana de PMDA. La media y la mediana describen el nivel típico de los máximos anuales y permiten comparar estaciones en términos de “escala” del extremo. En contextos con asimetría o presencia de valores altos, la mediana puede aportar robustez frente a extremos muy influentes, mientras que la media es útil para lecturas integradas del comportamiento general (Achite et al., 2024).

Desviación estándar. Cuantifica la dispersión absoluta de las PMDA alrededor de su valor típico y refleja cuán variable es el extremo anual en cada estación. Valores altos sugieren mayor fluctuación interanual y potencial sensibilidad a condiciones climáticas de gran escala o a mecanismos locales (Achite et al., 2024).

Coefficiente de variación (CV). Expresa la dispersión relativa respecto a la media y permite comparar variabilidad entre estaciones (Achite et al., 2024).

Máximo, mínimo. Los valores máximo y mínimo de la serie PMDA delimitan el rango observado y sirven para identificar años críticos y posibles inconsistencias por errores de medición o transcripción. En análisis de extremos, estos valores se revisan con criterio hidrometeorológico, ya que un máximo muy alto puede ser físicamente plausible (evento extremo real) o un outlier por falla de registro (Achite et al., 2024).

d) Estacionariedad / tendencia

La estacionariedad se refiere a la estabilidad estadística del proceso generador de extremos en el tiempo, supuesto operativo frecuente en análisis de frecuencia. Sin embargo, en precipitación extrema pueden presentarse tendencias o cambios graduales que alteren la media o la distribución, afectando cuantiles estimados bajo supuestos estacionarios (Silva et al., 2022).

Estaciones con tendencia y requerimiento de corrección. Cuando una estación presenta tendencia estadísticamente significativa, se asume que la serie puede violar el supuesto de estabilidad temporal requerido para comparaciones regionales. En

práctica, ello obliga a justificar una corrección, un tratamiento de la serie o una decisión de exclusión según el impacto esperado sobre cuantiles de diseño (Silva et al., 2022).

Resultados Mann–Kendall, Spearman y tendencia en la media. Se emplean para detectar tendencias monotónicas sin exigir normalidad, lo que las hace adecuadas para series hidrometeorológicas con asimetría. Mann–Kendall evalúa concordancia temporal de rangos, mientras que Spearman mide asociación monotónica entre tiempo y variable mediante correlación por rangos (Kavwenje et al., 2024).

e) Homogeneidad (cambios en media/varianza)

La homogeneidad se refiere a la ausencia de cambios artificiales en la serie por factores no climáticos, como reubicación de estación, cambios de instrumento u operación. La evaluación de homogeneidad se considera una etapa formal de control de calidad y consistencia previa a análisis regional (Mohammed y Scholz, 2023).

Estaciones no homogéneas y corrección. Si una estación resulta no homogénea, se asume que existe al menos un quiebre que altera la comparabilidad interna del registro y, por extensión, la comparabilidad regional. La corrección busca restituir coherencia estadística mediante ajustes (Mohammed y Scholz, 2023).

Resumen cumplen/no cumplen por prueba. Consolida el diagnóstico de calidad y permite justificar, de forma verificable, qué estaciones ingresan a la fase regional y bajo qué condiciones (Mohammed y Scholz, 2023).

Resultados Mann-Whitney, t y F. La prueba Mann–Whitney contrasta si dos segmentos provienen de distribuciones con distinta localización, siendo útil ante datos no normales; enfoques como Pettitt se fundamentan en Mann–Whitney para detectar quiebres. La prueba t evalúa diferencias de medias bajo supuestos paramétricos, mientras que la prueba F contrasta igualdad de varianzas, aportando diagnóstico específico sobre cambios en dispersión (Mohammed y Scholz, 2023).

2.3.4. Regionalización

La regionalización en hidrología consiste en transferir información entre sitios de una región con comportamiento hidrológico similar, especialmente para estimar variables en lugares con pocos o ningún dato (Zucarelli, 2017). En el enfoque de Hosking y Wallis (1997), este proceso agrupa estaciones en regiones homogéneas mediante técnicas estadísticas para modelar la distribución de eventos extremos, como las precipitaciones máximas, asumiendo similitud en sus distribuciones de frecuencia.

Para Zamanillo et al. (2009), la regionalización de precipitaciones máximas incluye seleccionar estaciones dentro y fuera de la cuenca, realizar análisis estadístico, elegir una distribución teórica y aplicar una metodología heurística, comparando resultados con la propuesta de Hosking y Wallis (1997). Finalmente, la regionalización se materializa en mapas que representan la variabilidad espacial de las PMDA y sus valores para distintos periodos de retorno a lo largo de la cuenca.

a) Tipos de regionalización

Idrogo (2024) señala que la regionalización hidrológica puede realizarse mediante: (1) variables, estimando precipitación, caudales o tiempos característicos a partir de relaciones obtenidas con datos puntuales; (2) funciones, ajustando distribuciones estadísticas a registros históricos y transfiriendo cuantiles (p. ej., picos para distintos periodos de retorno) apoyándose en curvas regionales adimensionales Q/Q_m vs. periodo de retorno, útil con series cortas o escasez de datos; (3) parámetros, vinculando parámetros de modelos con características físicas de cuenca (p. ej., CN del SCS según suelo, cobertura y uso); y (4) indicadores regionales, usando valores medios representativos (como caudal específico medio o razón caudal mínimo/medio) para estimaciones rápidas del comportamiento hidrológico.

b) Proceso de regionalización

Parizaca (2012) menciona que el procedimiento implementado para llevar a cabo de regionalización de tormentas convectivas implica relacionar, en forma integral, varios conceptos asociados con las hipótesis de partida, la distribución regional de las lluvias máximas, los factores de ajuste asociados con cortas y largas duraciones, el factor de reducción por periodo de retorno, el factor de reducción por área (FRA) y la distribución temporal de la lluvia.

c) Regionalización utilizando L-Momentos

Permite ver el comportamiento hidrológico para el diseño de proyectos; este método está basado en la estadística obteniendo datos regionalizados así parámetros que permitan calcular eventos máximos en sitios donde no se cuente con este tipo de información. El método L-momentos se han utilizado para ajustar las distribuciones de frecuencia e identificar regiones homogéneas para las precipitaciones con mayor cantidad de registro observada en 24 horas durante un año (Cadenillas, 2024).

2.3.5. Regionalización PMDA por L-momentos

La regionalización de PMDA mediante L-momentos se sustenta en el Análisis Regional de Frecuencia (ARF), el cual combina información de varias estaciones para estimar cuantiles de diseño en sitios con registros limitados o con alta incertidumbre. El enfoque utiliza L-momentos y razones-L por su robustez frente a valores extremos y por su estabilidad en muestras moderadas, permitiendo comparar estaciones y definir regiones estadísticamente consistentes (Ahmed et al., 2025).

a) Agrupación preliminar

La agrupación preliminar constituye el primer filtrado para proponer regiones candidatas, antes de aplicar pruebas formales de discordancia y heterogeneidad. En estudios de RFA con L-momentos se recomienda que esta fase combine criterios

estadísticos con criterios físico-geográficos, con el fin de evitar regiones “artificiales” que luego fallan en homogeneidad. La estandarización de variables y evaluación del número de grupos son pasos clave para no inducir sesgos (Ahmed et al., 2025).

Grupos preliminares Ward + K-means. El procedimiento se interpreta como estrategia de dos etapas: Ward identifica estructuras jerárquicas (dendrograma) minimizando la varianza intragrupo, mientras que K-means consolida la partición final optimizando la asignación por centroides (Ahmed et al., 2025).

Dmax por región preliminar. Resume el mayor valor del estadístico de discordancia dentro del grupo, utilizado para identificar estaciones que “no se comportan” como el resto según sus razones-L. La discordancia evalúa si una estación presenta una combinación de L-momentos atípica respecto a la nube regional, lo que puede reflejar errores de datos o diferencias físicas no capturadas (Ahmed et al., 2025).

H1 por región preliminar. El estadístico H1 (heterogeneidad) cuantifica, a nivel regional, si la dispersión observada de las razones-L es compatible con una región homogénea. La lógica es comparar la variabilidad interna del grupo contra una referencia generada por simulación, de modo que H1 sintetiza si el grupo es aceptable o requiere reconfiguración (Ahmed et al., 2025).

b) Cálculo de L-momentos y razones-L

El cálculo de L-momentos ($L_1, L_2, \dots$) se fundamenta en combinaciones lineales de estadísticas de orden, lo que reduce la sensibilidad a valores extremos aislados frente a momentos convencionales. Para regionalización, el interés principal se centra en las razones-L: t ($L-CV$), t_3 ($L-Cs$) y t_4 ($L-Ck$), porque permiten comparar forma y variabilidad entre estaciones sin depender de la escala. Estas métricas constituyen el insumo para pruebas de discordancia, heterogeneidad y para el uso de diagramas razón-L en la selección de distribuciones (Ahmed et al., 2025).

l1, t (L-CV), t3 (L-Cs), t4 (L-Ck), n. L1 representa el componente de escala, mientras que t, t3 y t4 caracterizan variación, asimetría y curtosis en forma adimensional, facilitando comparaciones entre estaciones. El tamaño muestral (n) controla la estabilidad de estimación: en series cortas se incrementa la incertidumbre de t3 y t4, por lo que su interpretación debe apoyarse en consistencia regional y no en valores aislados (Ahmed et al., 2025).

c) Ajuste iterativo y regiones definitivas (ARF)

El ajuste iterativo es el proceso de refinamiento que transforma regiones preliminares en regiones definitivas, asegurando simultáneamente ausencia de estaciones discordantes y homogeneidad aceptable. Metodológicamente, el ajuste no es un “resultado” sino una rutina de control de calidad regional: se reagrupan estaciones, se evalúa el efecto sobre D y H, y se reitera hasta cumplir criterios (Ahmed et al., 2025).

Composición final de regiones homogéneas. Se define cuando cada región tiene coherencia estadística y espacial minimizando “islas” o estaciones desconectadas. Una región homogénea no implica igualdad de magnitudes, sino similitud en la forma de la distribución una vez normalizada por el índice local (Ahmed et al., 2025).

Rango altitudinal por región. Se usa como descriptor fisiográfico de soporte para interpretar la consistencia del agrupamiento, especialmente en cuencas montañosas donde la lluvia presenta gradientes orográficos (Ahmed et al., 2025).

Dmax final por región. Confirma la etapa de control de discordancia tras los ajustes, verificando que no permanezcan estaciones atípicas que puedan sesgar la distribución regional. Su interpretación es preventiva: aun si una región pasa heterogeneidad, una estación discordante puede distorsionar parámetros regionales. Por ello, el Dmax final se entiende como un criterio de “limpieza regional” que debe cumplirse antes de estimar cuantiles por periodos de retorno (Ahmed et al., 2025).

H1 final por región. Representa el veredicto estadístico de homogeneidad, en el sentido de que la variabilidad interna del grupo es compatible con una región común bajo el marco de RFA (Ahmed et al., 2025).

Verificación gráfica ratios-L. Complementa los estadísticos, porque permite identificar visualmente dispersión, agrupamientos extraños o puntos atípicos que podrían pasar inadvertidos en métricas resumen. En particular, los diagramas $t-t_3$ y t_3-t_4 se usan para contrastar la nube regional con el comportamiento esperado de familias de distribuciones candidatas (Ahmed et al., 2025).

d) Selección de distribución regional (ARF)

La selección de distribución regional corresponde al tramo inferencial del ARF: una vez validada la homogeneidad, se elige la familia probabilística que mejor representa la forma regional. El principio es evitar sobre-ajuste por estación y, en cambio, seleccionar una distribución que explique el comportamiento conjunto del grupo, respaldada por métricas de bondad de ajuste y por compatibilidad con los diagramas razón-L. En la práctica, se evalúan varias distribuciones (LNO, GEV, P3, GLO) y se prioriza aquella con desempeño consistente y estable (Ahmed et al., 2025).

Z por distribución y región (criterio $|Z| \leq 1.64$). El estadístico Z por distribución resume la discrepancia entre la curtosis-L regional observada y la curtosis-L que implicaría la distribución candidata bajo el ajuste regional. El criterio $|Z| \leq 1.64$ se interpreta como una regla de aceptación práctica para no rechazar la distribución como representación regional (Ahmed et al., 2025).

Distribución óptima por región. Se define como la alternativa aceptada por el criterio Z y que, adicionalmente, presenta coherencia con el diagrama razón-L y estabilidad en cuantiles para periodos de retorno altos (Ahmed et al., 2025).

Dimensión: Cuantiles regionales y locales (ARF). La dimensión de cuantiles regionales y locales operacionaliza el propósito de la regionalización: estimar lluvias de diseño (PT) de forma consistente dentro de cada región. Teóricamente, el ARF separa forma y escala: la forma se captura en la curva regional de crecimiento y la escala en el índice local (p. ej., L1 o media-L) (Ahmed et al., 2025).

Cuantiles regionales adimensionales q_R por TR. Corresponden a la curva de crecimiento regional y se interpretan como factores multiplicativos que crecen con el periodo de retorno. La obtención de $q_R(T)$ depende de la distribución seleccionada y de parámetros estimados por L-momentos regionales (Ahmed et al., 2025).

Cuantiles locales por estación y región. Los cuantiles locales se derivan al escalar $q_R(T)$ por el índice local de la estación (índice de tormenta), produciendo PT en unidades de precipitación (mm) (Ahmed et al., 2025).

e) Región de Influencia (ROI)

El enfoque ROI redefine la regionalización como un esquema “localmente adaptativo”, donde cada estación objetivo construye su propio conjunto de estaciones influyentes en función de una medida de similitud (hidroclimática y/o fisiográfica). En bases teóricas, la diferencia clave con el ARF fijo es que ROI no impone fronteras regionales rígidas: permite que dos estaciones cercanas compartan parte de su grupo, pero no necesariamente pertenezcan a la misma región cerrada (Ahmed et al., 2025).

Grupos ROI por estación objetivo. La construcción del grupo ROI por estación objetivo se basa en ordenar estaciones potenciales por cercanía y seleccionar un subconjunto que maximice consistencia y estabilidad de estimación. ROI se apoya en validaciones internas (discordancia/heterogeneidad local) para mantener un equilibrio entre robustez y similitud (Ahmed et al., 2025).

Sensibilidad a distribución en ROI (curvas por estación objetivo). En ROI, la sensibilidad a la distribución suele aumentar porque la composición del grupo cambia por estación, afectando parámetros regionales y, por ende, la forma de la curva de crecimiento. Esta sensibilidad no invalida el enfoque; al contrario, obliga a documentar la estabilidad de PT como criterio de confiabilidad, especialmente en retornos altos (Ahmed et al., 2025).

f) Cartografía y caracterización espacial

La cartografía en regionalización se sustenta en trasladar resultados puntuales a superficies continuas, con el fin de apoyar la planificación y el diseño en zonas no instrumentadas. Teóricamente, esto exige dos decisiones: (i) cómo representar regiones homogéneas y (ii) cómo interpolar cuantiles de diseño por TR (Ferreira et al., 2024).

Cobertura espacial por región ARF (área, % área, rangos lon/lat). Cuantifica la representatividad territorial del agrupamiento y permite evaluar si una región domina el área de estudio o si existen regiones pequeñas con alta incertidumbre por baja densidad de estaciones (Ferreira et al., 2024).

Área de influencia por estación y región. El área de influencia se entiende como la zona para la cual una estación actúa como referencia dominante en la interpolación, y su delimitación depende del método cartográfico. Su objetivo es transparentar cómo se “propaga” la información observada hacia zonas sin medición, evitando interpretaciones erróneas de continuidad espacial (Ferreira et al., 2024).

Mapa regiones homogéneas ARF. Representa la partición espacial del territorio en zonas con comportamiento estadístico comparable de extremos, y funciona como soporte de decisión para seleccionar la curva regional apropiada. Este mapa no es un “producto GIS” aislado, sino la traducción espacial de pruebas estadísticas y de criterios físicos de coherencia (Ferreira et al., 2024).

Mapas PMDA regionalizada por TR (50, 100, 200, 500). Los mapas por TR materializan la variable de diseño (PT) como superficie continua, permitiendo identificar zonas críticas por retorno y comparar gradientes espaciales del peligro hidrometeorológico. La coherencia metodológica exige que estos mapas usen los cuantiles locales derivados del ARF/ROI y un método de interpolación documentado y consistente (Ferreira et al., 2024).

Mapas finales ROI (grupos). Pueden incluir, además de superficies de PT, la visualización de la composición espacial de los grupos influyentes (“vecindarios hidrológicos” por estación), lo que mejora la interpretabilidad del enfoque adaptativo. Estos mapas sirven para evidenciar transiciones climáticas y para justificar por qué dos puntos cercanos pueden presentar grupos ROI distintos (Ferreira et al., 2024).

CAPÍTULO III: MARCO METODOLÓGICO

3.1. Tipo, diseño, nivel y enfoque de investigación

El enfoque de la investigación fue cuantitativo, debido a que se trabajó con series numéricas de precipitaciones máximas diarias anuales por estación y año. Los registros fueron organizados, depurados y analizados mediante estadística descriptiva, L-momentos y procedimientos de regionalización hidrológica (Hernández-Sampieri y Mendoza, 2020).

El estudio fue de tipo básico, debido a que generó conocimiento metodológico y estadístico sobre la regionalización de PMDA sin intervenir físicamente en el sistema hidrológico analizado. Sus resultados ofrecen aplicaciones técnicas para el diseño hidráulico y la gestión del riesgo (Esteban, 2018).

El nivel fue descriptivo, ya que se caracterizaron las series de PMDA, su variabilidad espacial y temporal, la conformación de regiones homogéneas y los cuantiles estimados para diferentes periodos de retorno (Morales, 2012).

El diseño fue no experimental y longitudinal, de corte retrospectivo. No se manipuló ninguna variable; se analizaron registros pluviométricos históricos provenientes de estaciones meteorológicas y se evaluó su comportamiento durante los periodos disponibles (Hernández-Sampieri y Mendoza, 2020).

Tabla 5

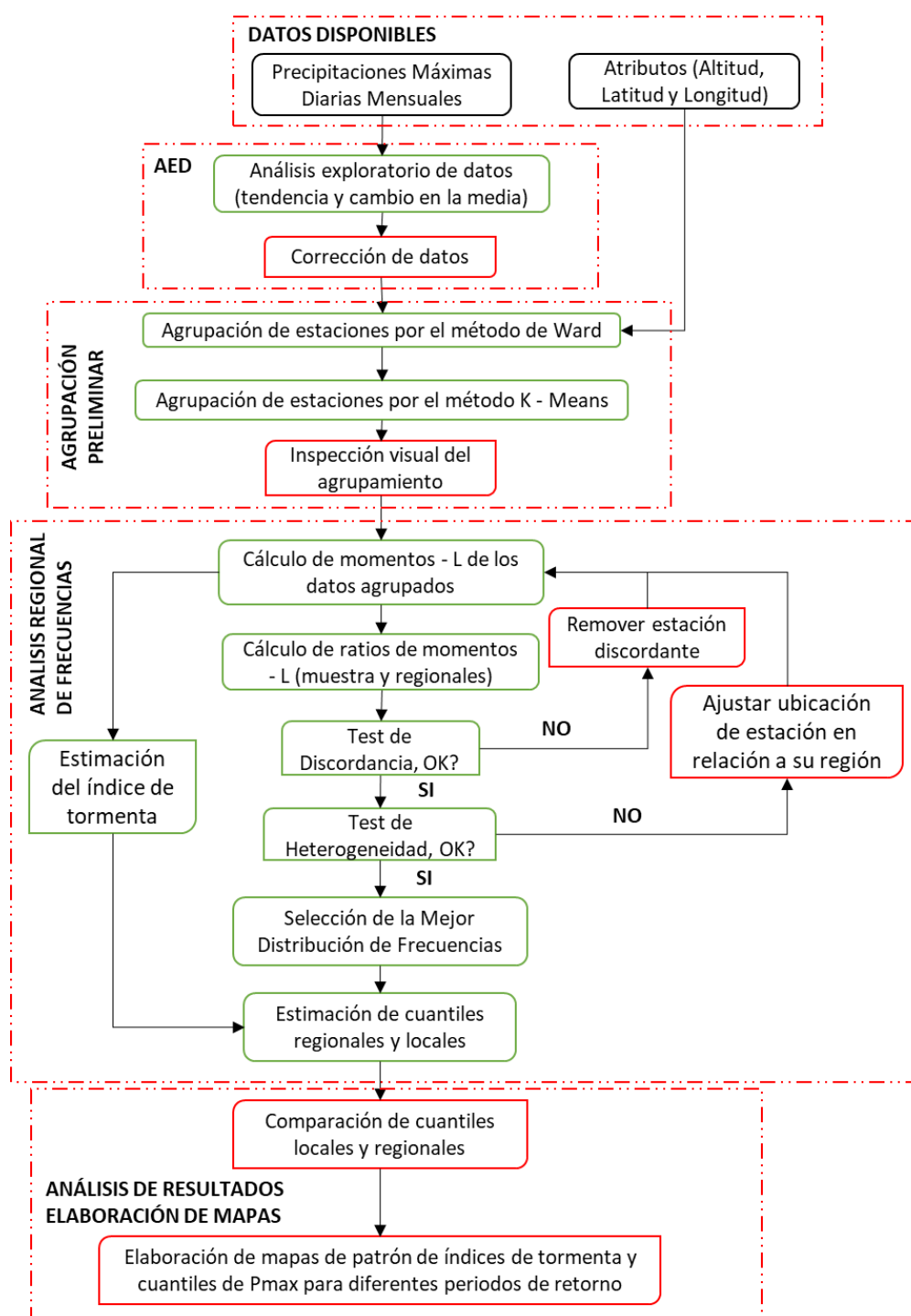
Tipo de investigación según los principales criterios

Criterio	Finalidad	Enfoque metodológico	Objetivos	Fuente de datos	Control de diseño	Temporalidad
Investigación	Básica	Cuantitativa	Descriptiva	Documental: datos secundarios	No experimental	Longitudinal retrospectiva

Nota. (Grajales, 2000).

Figura 7

Esquema de investigación



Nota. El esquema resume la secuencia metodológica aplicada para regionalizar las PMDA en la cuenca del río Chamaya mediante L-momentos, desde el análisis exploratorio y corrección de series, la agrupación preliminar (Ward y K-means) y la verificación de homogeneidad (discordancia y heterogeneidad), hasta la comparación de cuantiles regionales y locales, y la elaboración de mapas.

3.2. Ubicación, población, muestra, muestreo y unidad de análisis

3.2.1. Ubicación de la zona de estudio

Geográficamente, la cuenca Chamaya está ubicada entre los 9250000.00 y 9450000.00 en la coordenada norte y entre los 650000.00 y 800000.00 en la coordenada este. Políticamente, la cuenca Chamaya se sitúa al norte del país y pertenece a la vertiente del Atlántico con una superficie de 8 137.19 km² aproximadamente, el área de estudio abarca 9 provincias de la zona norte del Perú: Huancabamba, San Ignacio, Jaén, Cutervo, Chota, Santa Cruz, Lambayeque, Ferreñafe y Hualgayoc.

Tabla 6

Puntos extremos de la cuenca Chamaya

Punto	Este (m)	Norte (m)	Ubicación
1	759775	9260786	Por el Sur
2	773003	9268930	Por el Sureste
3	673538	9321065	Por el Suroeste
4	754646	9360806	Por el Este
5	653123	9367219	Por el Oeste
6	670764	9448571	Por el Norte

Tabla 7

Departamentos de la cuenca

Departamento	Superficie (km ²)	Porcentaje de cuenca hidrográfica
Cajamarca	6001.4	74%
Piura	1703.1	21%
Lambayeque	324.4	4%
Amazonas	81.1	1%

Tabla 8

Codificación Pfafstetter unidad hidrográfica de la cuenca Chamaya

Nivel	Nombre	Código
N1	Región Hidrográfica del Amazonas	4
N2	Alto Amazonas	49
N3	Marañón	498
N4	Alto Marañón	4889
N5	Chamaya	49896

Nota. (Bustamante, 2019).

Ubicación política de la cuenca Chamaya



MAPA DE PERÚ

UNIVERSIDAD NACIONAL AUTÓNOMA DE CHOTA

FACULTAD DE CIENCIAS DE LA INGENIERÍA

ESCUELA PROFESIONAL DE INGENIERÍA CIVIL

Presentado por: Bach. Robba Williams Oca Cubas
Bach. Jherson Rivera Guamuro

Asesor: Dr. Ing. Luis Alberto Orbeago Navarro

Regionalización de precipitaciones máximas diarias anuales en la cuenca del río Chamaya, aplicando el método L-Momentos, 2024

Plano: MAPA HIDROLÓGICO DE LA CUENCA CHAMAYA

Fecha: Enero, 2025
Escala: 1:800,000

Mapa: U-1

3.2.2. Población

La población de la investigación estuvo constituida por el conjunto total de registros diarios de precipitación disponibles en las estaciones meteorológicas con influencia hidrometeorológica sobre la cuenca del río Chamaya, ubicadas dentro y en el entorno inmediato de la cuenca. Comprende las series históricas diarias provistas por el SENAMHI, a partir de las cuales se derivaron las precipitaciones máximas diarias anuales (PMDA) por estación para el análisis regional.

3.2.3. Muestreo

El muestreo fue no probabilístico, de tipo intencional (por criterios técnicos), debido a que la selección de estaciones no se realizó al azar, sino conforme a la disponibilidad y pertinencia espacial respecto a la cuenca. Se seleccionaron estaciones con información suficiente para construir series PMDA y ejecutar el ARF y el ROI con L-momentos.

3.2.4. Muestra

La muestra estuvo conformada por las series de PMDA (máximos diarios en 24 h por año) obtenidas de las estaciones meteorológicas seleccionadas para el estudio, tanto dentro como cercanas a la cuenca del río Chamaya. Estas series correspondieron a las estaciones que resultaron aptas para el análisis regional y para la conformación y verificación de regiones homogéneas.

3.2.5. Unidad de análisis

La estación meteorológica (pluviométrica/convencional) considerada como unidad espacial de observación, junto con su serie anual de PMDA asociada. Es decir, el análisis se ejecutó a nivel de cada estación mediante su conjunto de máximos anuales y, posteriormente, a nivel regional con el agrupamiento de estaciones en regiones homogéneas para estimar cuantiles de diseño.

3.3. Técnicas e instrumentos de recolección de datos

3.3.1. Técnicas de recolección de datos

Revisión documental. La revisión documental consiste en identificar, seleccionar, recopilar y analizar información proveniente de documentos, registros, bases de datos, informes técnicos y literatura científica relevante para el objeto de estudio (Hernández-Sampieri y Mendoza, 2020). En la investigación, la revisión se centró en la recopilación de datos históricos de precipitaciones máximas diarias anuales registrados en las estaciones meteorológicas de la cuenca del río Chamaya y de zonas cercanas. Se usaron fuentes como la web del SENAMHI y ANA, además de datos solicitados al SENAMHI.

Análisis estadístico de datos: Aplicación de Método L- Momentos. El análisis estadístico de datos secundarios consiste en aplicar métodos y procedimientos matemáticos para examinar y resumir información cuantitativa previamente recolectada, lo que permite identificar patrones, relaciones y tendencias significativas (Wilks, 2006). En este estudio, el análisis estadístico permitió calcular los L-momentos a partir de datos de precipitación, ajustar distribuciones de probabilidad y regionalizar las precipitaciones máximas para diferentes periodos de retorno. Además, se generaron mapas que muestran la variabilidad espacial de las lluvias extremas.

3.3.2. Instrumentos de recolección de datos

Ficha de registro de datos hidrometeorológicos (Formatos de hojas de Excel). Se empleó un formato estructurado en hojas de cálculo para consolidar los registros históricos de precipitación diaria y las series de PMDA de cada estación. La ficha incluyó el código y nombre de la estación, coordenadas geográficas, altitud, periodo disponible, años faltantes, valores máximos anuales y observaciones de control de calidad.

Software hidrológico y estadístico (R, Excel, Qgis). El análisis de los datos se realizó utilizando programas estadísticos especializados, que permiten calcular las L-momentos, ajustar distribuciones de probabilidad y generar mapas de regionalización. En particular, se utilizó R (versión 4.2.2) con los paquetes *lmom* y *lmomco*, Microsoft Excel 365 para organización preliminar de datos, y HydroOffice 2023 para análisis hidrológico aplicado. Estos programas permitieron calcular medidas robustas de forma, escala y localización; y evaluar la homogeneidad regional según los criterios de Hosking y Wallis (1997). Los cálculos incluyeron cuantiles para periodos de retorno seleccionados (50, 100, 200, 500 años) y la generación de mapas temáticos.

3.4. Hipótesis

La aplicación del método de L-momentos permite regionalizar las precipitaciones máximas diarias anuales en la cuenca del río Chamaya mediante la delimitación de regiones homogéneas estadísticamente consistentes y la estimación de cuantiles de diseño para diferentes periodos de retorno.

3.5. Operacionalización de variables

3.5.1. Variable independiente: Precipitaciones máximas diarias anuales (PMDA)

Las precipitaciones máximas diarias anuales (PMDA) observadas corresponden al mayor valor de precipitación diaria (24 h) registrado en cada año hidrológico en una estación meteorológica. Esta serie de máximos anuales representa el comportamiento extremo de la precipitación y es la base para el análisis de frecuencia y la estimación de cuantiles de diseño en estudios hidrológicos. La PMDA observada se operacionalizó a partir de los registros diarios de precipitación por estación, extrayendo el máximo diario por año para conformar la serie anual (AM/AMS). Posteriormente, estas series se depuraron y evaluaron según disponibilidad y consistencia, caracterización

estadística, análisis de estacionariedad y homogeneidad para la regionalización (Alejo-Sanchez et al., 2025).

3.5.2. *Variable dependiente: Regionalización PMDA por L-momentos*

La regionalización de PMDA por L-momentos es un proceso de Análisis Regional de Frecuencia que agrupa estaciones en regiones homogéneas y estima cuantiles de precipitación extrema mediante L-momentos y razones-L, reduciendo la incertidumbre de estimación en sitios con registros cortos o limitada instrumentación. Permite obtener cuantiles regionalizados por periodos de retorno y su representación espacial, sustentados en criterios de homogeneidad regional (Zucarelli, 2017).

La regionalización se operacionalizó calculando, por estación, los L-momentos y razones-L ($L1$, t , $t3$, $t4$) y aplicando criterios de control regional para conformar regiones homogéneas (discordancia D y heterogeneidad $H1$). Luego se seleccionó la distribución regional mediante el estadístico Z , se estimaron cuantiles regionales y locales PT por periodos de retorno definidos y se compararon los enfoques ARF y ROI. Finalmente, los resultados se representaron mediante cartografía regional.

3.5.3. Matriz de operacionalización de variables

Tabla 9

Matriz de operacionalización de variables

Variable	Definición conceptual	Definición operacional	Dimensión	Indicadores	Unidad de medida	Técnica	Instrumento
Precipitaciones máximas diarias anuales (PMDA) observadas	Máximo valor de precipitación diaria registrado en cada año hidrológico/ calendario en una estación meteorológica.	A partir de registros diarios por estación, se extrae el máximo diario por año (serie AMS/PMDA), se depura y se resume estadísticamente para su uso en regionalización (Alejo-Sanchez et al., 2025).	Inventario y cobertura espacial de estaciones	Ubicación de estaciones (altitud)	msnm	Recolección y sistematización	Base SENAMHI
				Proximidad espacial	km	Análisis espacial (distancias)	Matriz de distancias
				Distribución espacial		Cartografía descriptiva	QGis
			Compleitud y consistencia del registro	Longitud del registro por estación (n años)	Años	Depuración/ control de calidad	Excel
				Años faltantes, % faltantes, % completos	%	Depuración/ control de calidad	Excel
			Caracterización estadística de PMDA	Media / Mediana de PMDA	mm	Análisis estadístico descriptivo	Excel
				Desviación estándar	mm		Excel
				Coeficiente de variación (CV)	%		Excel
				Máximo, mínimo	mm		Excel
			Estacionariedad/ tendencia	Estaciones con tendencia y requerimiento de corrección	n (conteo)	Pruebas de tendencia	Excel
				Resultados Mann–Kendall, Spearman y tendencia en la media	Adimensional	Pruebas de tendencia	Excel
			Homogeneidad (cambios en media/varianza)	Estaciones no homogéneas y corrección	n (conteo)	Pruebas de homogeneidad	Excel
				Resumen cumplen/no cumplen por prueba	n (conteo)	Síntesis de contrastes	Excel
				Resultados Mann-Whitney, t y F	Adimensional	Pruebas de homogeneidad	Excel
Regionalización PMDA por L-momentos	Procedimiento de Análisis Regional de Frecuencia que agrupa estaciones en regiones homogéneas y estima cuantiles regionalizados con estadísticos robustos (L-momentos) y razones-L.	Se calculan L-momentos/razones-L, se evalúan discordancia y heterogeneidad, se delimitan regiones homogéneas (ARF), se ajustan distribuciones, se estiman cuantiles (ARF y ROI) y se espacializan resultados (Zucarelli, 2017).	Agrupación preliminar	Grupos preliminares Ward + K-means	categoría	Clasificación/ agrupamiento	Software estadístico + SIG
				Dmax por región preliminar	adimensional	Discordancia	Excel
				H1 por región preliminar	adimensional	Heterogeneidad	Excel
			Cálculo de L-momentos y razones-L	l1, t (L-CV), t3 (L-Cs), t4 (L-Ck), n	mm; adimensional	Análisis regional	Plantilla L-momentos (Excel)
			Ajuste iterativo y regiones definitivas (ARF)	Composición final de regiones homogéneas	categoría	Ajuste iterativo	Excel / SIG
				Rango altitudinal por región	msnm	Caracterización física	SIG / Excel
				Dmax final por región	adimensional	Discordancia	Excel
				H1 final por región	adimensional	Heterogeneidad	Excel
				Verificación gráfica ratios-L	—	Diagnóstico gráfico	Excel
			Selección de distribución regional (ARF)	Z por distribución y región (criterio $ Z \leq 1.64$)	adimensional	Bondad de ajuste	Excel
				Distribución óptima por región	categoría	Selección de modelo	Excel
			Cuantiles regionales y locales (ARF)	Cuantiles regionales adimensionales qR por TR	adimensional	Curva regional de crecimiento	Excel
				Cuantiles locales por estación y región	mm	Estimación de cuantiles	Excel
			Región de Influencia (ROI)	Grupos ROI por estación objetivo	categoría	Agrupamiento adaptativo	Excel / SIG
				Sensibilidad a distribución en ROI (curvas por estación objetivo)	mm	Comparación de distribuciones	Excel
			Comparación ARF vs ROI	Comparación espacial ARF vs ROI	—	Análisis comparativo espacial	SIG
				Cuantiles comparados (Local vs ARF vs ROI) por estación	mm	Comparación de enfoques	Excel
			Cartografía y caracterización espacial	Cobertura espacial por región ARF (área, % área, rangos lon/lat)	km², %	Análisis espacial	SIG
				Área de influencia por estación y región	km²	Delineación espacial	SIG
				Mapa regiones homogéneas ARF	—	Cartografía	SIG
				Mapas PMDA regionalizada por TR (50, 100, 200, 500)	mm	Interpolación/representación	SIG
				Mapas finales ROI (grupos)	—	Cartografía	SIG

3.6. Procedimientos de recolección de datos

3.6.1. Delimitación del área de estudio y base cartográfica

Se delimitó la cuenca del río Chamaya como unidad hidrológica de análisis para definir el ámbito espacial de la regionalización. Con dicha delimitación se establecieron los límites, el área aproximada y el marco de referencia espacial para la ubicación de estaciones meteorológicas. Enseguida, se consolidó una base cartográfica de trabajo para superponer la cuenca con la red de estaciones y facilitar la verificación de las estaciones ubicadas dentro de la cuenca y en su entorno inmediato.

a) Insumos principales.

- Polígono de cuenca (límite oficial o delimitación hidrológica validada).
- Cartografía base (hidrografía, límites administrativos si aplica).
- Modelo de elevación (DEM) para contextualización altitudinal.
- Coordenadas geográficas (longitud/latitud) de estaciones.

b) Software utilizado.

- QGIS (3.40): gestión espacial, Voronoi/Thiessen, intersecciones, disoluciones, cálculo de áreas.
- ArcGIS (10.8): herramientas equivalentes a Thiessen y de geoprocésamiento.
- Excel: depuración de tablas de coordenadas y atributos antes de SIG.

c) Procedimiento en QGIS.

- Se creó un proyecto SIG y se definió su sistema de referencia. Para la revisión geométrica de coordenadas se trabajó en WGS84 (EPSG:4326), mientras que para cálculos métricos (áreas y distancias) se reproyectaron las capas a una proyección en metros, utilizándose UTM WGS84 zona 17S (EPSG:32717).

- Se incorporó el polígono de la cuenca Chamaya (vector) y se revisó su geometría mediante la verificación/corrección de geometrías, a fin de evitar errores en intersecciones y cálculos de área.
- Se cargaron capas base (hidrografía y límites administrativos cuando aplicó) para contextualizar el área de estudio y mejorar la lectura cartográfica.
- Se cargó el DEM y, para su representación, se generaron productos auxiliares como sombreado de relieve y rangos altitudinales (según disponibilidad del raster), con el propósito de describir el gradiente altitudinal del ámbito.
- Se determinó las propiedades geomorfológicas de la cuenca Chamaya.
- Se delimitó el mapa climático de la cuenca Chamaya para caracterizarla de forma local a fin de definir la estacionalidad climática en contraste con las estaciones ubicadas dentro y en áreas cercanas.

d) Mapa de delimitación de cuenca y sus parámetros geomorfológicos.

La cuenca del río Chamaya es extensa (8 137.19 km²) y con red de drenaje desarrollada (orden 7). Sus parámetros de forma indican una cuenca alargada y poco compacta ($F = 0.142$; $K_c = 2.272$), por lo que la respuesta hidrológica tiende a distribuirse en el tiempo y no concentrarse en picos muy abruptos. El relieve es marcadamente montañoso (altitud media 2 216.841 m s.n.m.; pendiente de cuenca 47.26%), lo que favorece alta variabilidad espacial de la lluvia y escorrentías rápidas en laderas, mientras que el cauce principal presenta pendiente relativamente baja (0.78%). Finalmente, la densidad de drenaje es baja a moderada (0.743 km/km²) y el tiempo de concentración es alto (1 742.3 min = 29 h), consistente con su gran tamaño y longitud de recorrido (238.792 km).

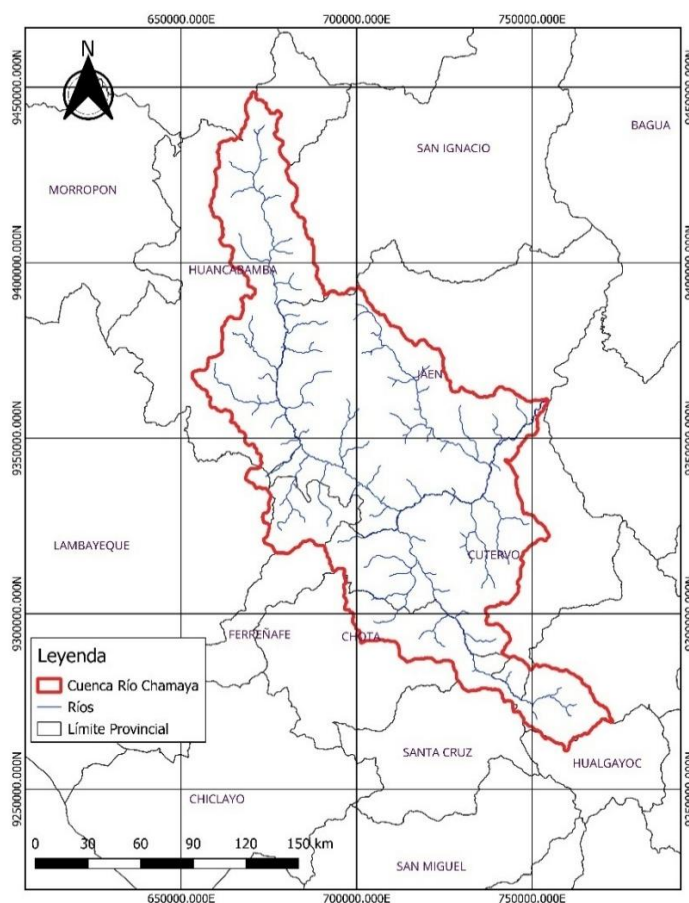
Tabla 10

Parámetros geomorfológicos e identificadores de la cuenca del río Chamaya

Categoría	Parámetro (símbolo)	Chamaya	Unidad
Identificación	IDUH	183	—
	Orden	118	—
	Orden de cursos de agua	7	—
	Emisor (E)	754 512 m E, 9 360 032 m S	m
Forma	Área (A)	8 137.19	km ²
	Perímetro (P)	729.189	km
	Factor de forma (F)	0.142	—
	Coef. de compacidad (Kc)	2.272	—
Relieve	Altitud media (H)	2 216.841	m s.n.m.
	Pendiente de cuenca (Sc)	47.26	%
	Pendiente cauce principal (Sp)	0.78	%
	Longitud máx. recorrido (Lmr)	238.792	km
	Coef. orográfico (Co)	0.001	—
Drenaje	Relación confluencias (Rc)	1.818	—
	Densidad de drenaje (Dd)	0.743	km/km ²
	Frecuencia de ríos (Fr)	0.701	N°/km ²
	Tiempo de concentración (Tc)	1 742.3	min

Figura 10

Delimitación de la cuenca del río Chamaya



e) Mapa de delimitación climática de cuenca.

En la Figura 11 se detalla la zonificación climática de la cuenca del río Chamaya según la clasificación de *Werren Thornthwaite*, en la que cada color corresponde a un código climático compuesto que integra: (i) la precipitación efectiva (A–E), (ii) la distribución estacional de la lluvia (r, i, p, v, o), (iii) el carácter térmico (A', B'1, B'2, B'3, C', etc.) y (iv) la humedad relativa (H1–H4), conforme a la Tabla 11. En la zona de estudio predomina el régimen C (semiseco), con distribución o,i,p (otoño, invierno y primavera secos) y condición de humedad H3 (húmedo); no obstante, se distinguen diferencias térmicas entre sectores: al norte prevalece B'3 (semifrío), mientras que al sur se registra A' (cálido), tal como resume la Tabla 12. Esta configuración espacial refleja el control del relieve y de los pisos altitudinales sobre el balance hídrico y la estacionalidad de la precipitación dentro de la cuenca, lo cual es relevante para interpretar la variabilidad de las PMDA y la necesidad de regionalización.

Tabla 11

Características del sistema climático

Componente	Código	Descripción	Componente	Código	Descripción
Precipitación efectiva	A	Muy lluvioso	Carácter del clima	A'	Cálido
	B	Lluvioso		B'1	Semicálido
	C	Semiseco		B'2	Templado
	D	Semiárido		B'3	Semifrío
	E	Árido		C'	Frío
Distribución anual de la precipitación	r	Precipitación abundante en todas las estaciones	Humedad relativa	H1	Muy seco
	i	Con invierno seco		H2	Seco
	p	Con primavera seca		H3	Húmedo
	v	Con verano seco		H4	Muy húmedo
	o	Con otoño seco			

Nota. (SENAMHI, 2020).

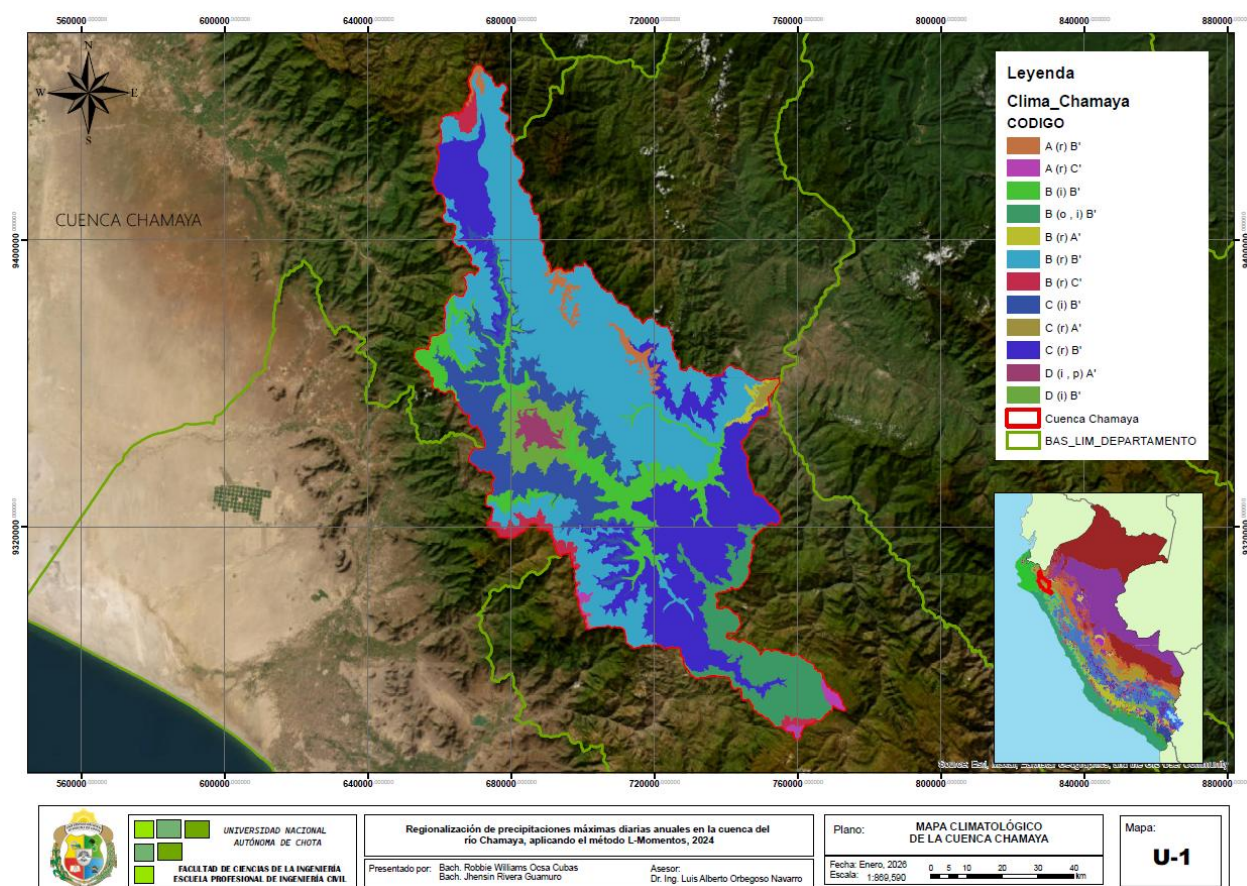
Tabla 12

Características del sistema climático de la zona de estudio

Precipitación Efectiva			Distribución de la Precipitación a Través del Año		Carácter del clima		Humedad relativa	
Norte	C	Semiseco	o,i,p	Con otoño seco, Con	B'3	Semifrío	H3	Húmedo
				invierno seco, Con primavera seca				
Sur	C	Semiseco	o,i,p	Con otoño seco, Con	A'	Cálido	H3	Húmedo
				invierno seco, Con primavera seca				

Figura 11

Mapa de las zonas climáticas de la cuenca Chamaya



Nota. Adaptado a partir de información del (SENAMHI, 2020).

3.6.2. *Identificación, inventario y selección de estaciones meteorológicas*

Se identificaron las estaciones meteorológicas con registro de precipitación diaria disponibles para el área de influencia de la cuenca. Se elaboró un inventario técnico por estación que incluyó código, tipo de estación, coordenadas geográficas, altitud y periodo de registro, a fin de documentar la trazabilidad de la información y sustentar la selección posterior. Con ese inventario se verificó la relación espacial estación–cuenca (dentro, borde o fuera), manteniendo estaciones externas solo cuando aportaron continuidad espacial o cobertura altitudinal relevante para la regionalización.

a) Criterios aplicados en la práctica.

- Tipo de estación: Convencional (CM).
- Variable: precipitación diaria (para derivar PMDA).
- Longitud mínima razonable de serie: Se trabajó con series cuyas longitudes variaron entre 21 y 60 años.
- Completitud aceptable: Se aceptaron series con niveles de completitud entre 72.7% y 98.3%, previa evaluación de su consistencia temporal.
- Inclusión espacial: 19 estaciones dentro de Chamaya y 6 estaciones de cuencas colindantes por proximidad y soporte regional (Chinchi, Chancay Lambayeque, Motupe, IAM III, IAM IV).

b) Fuentes de datos.

- SENAMHI Data (descarga directa).
- Solicitud formal al SENAMHI (con periodos explícitos; predominó 1995–2024 en varias estaciones y 2008–2024 para Porculla).

c) Procedimiento de inventario de estaciones.

- Consolidación del inventario maestro con atributos: código/ID, nombre, región/provincia/distrito, longitud, latitud, altitud, cuenca de pertenencia y condición (dentro/borde).
- Revisión rápida de disponibilidad temporal por estación (años totales y años faltantes).
- Selección operativa manteniendo todas las 25 estaciones y priorizando corrección/ajuste (en lugar de descarte), para no degradar la cobertura espacial.

d) Incorporación y validación espacial de estaciones meteorológicas en QGIS.

- Se depuró en Excel la tabla de estaciones (campos, formato numérico y vacíos), y se importó a QGIS como capa de puntos a partir de las coordenadas longitud/latitud, asignando el CRS correspondiente (WGS84).
- Se completaron/confirmaron los atributos mínimos por estación: ID, nombre, longitud, latitud y altitud.
- Se verificó la extensión espacial del conjunto de estaciones mediante la inspección de la envolvente espacial y la revisión de valores extremos. En el estudio, las estaciones cubrieron aproximadamente longitudes -79.53 a -78.46 , latitudes -6.69 a -5.10 y altitudes 1012–3300 msnm.
- Se ejecutó una comprobación visual y atributiva para detectar puntos duplicados, coordenadas invertidas (lat/long) y estaciones fuera del ámbito esperado. Cuando se detectaron inconsistencias, se corrigieron antes del análisis espacial.
- Se generó la visualización de estaciones sobre la cuenca y cartografía base, asegurando simbología legible (por categoría, región preliminar o región final, según etapa del análisis).

Figura 12

Mapa de ubicación de las estaciones meteorológicas de la cuenca Chamaya

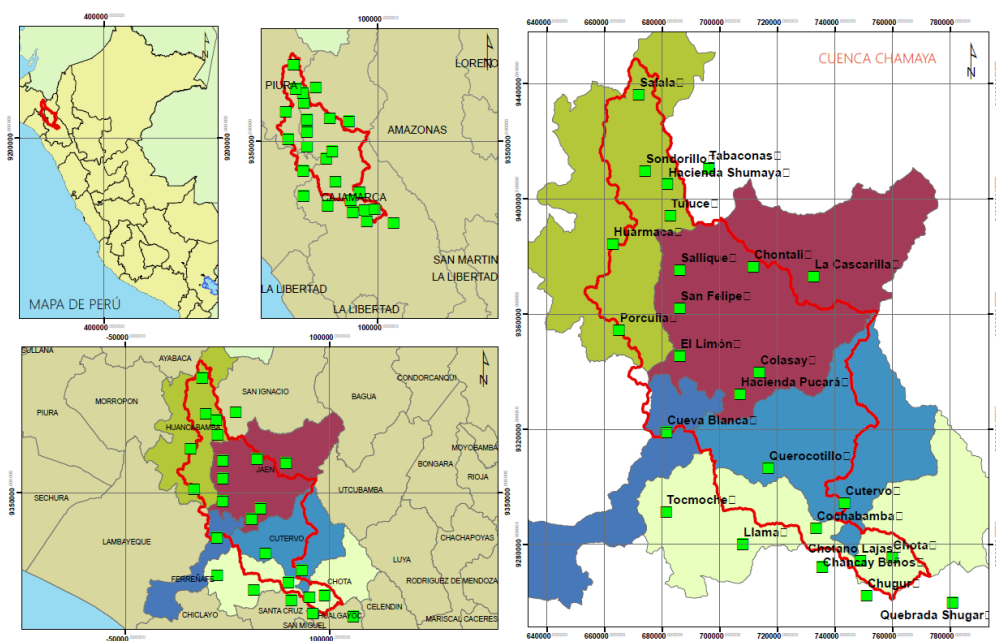
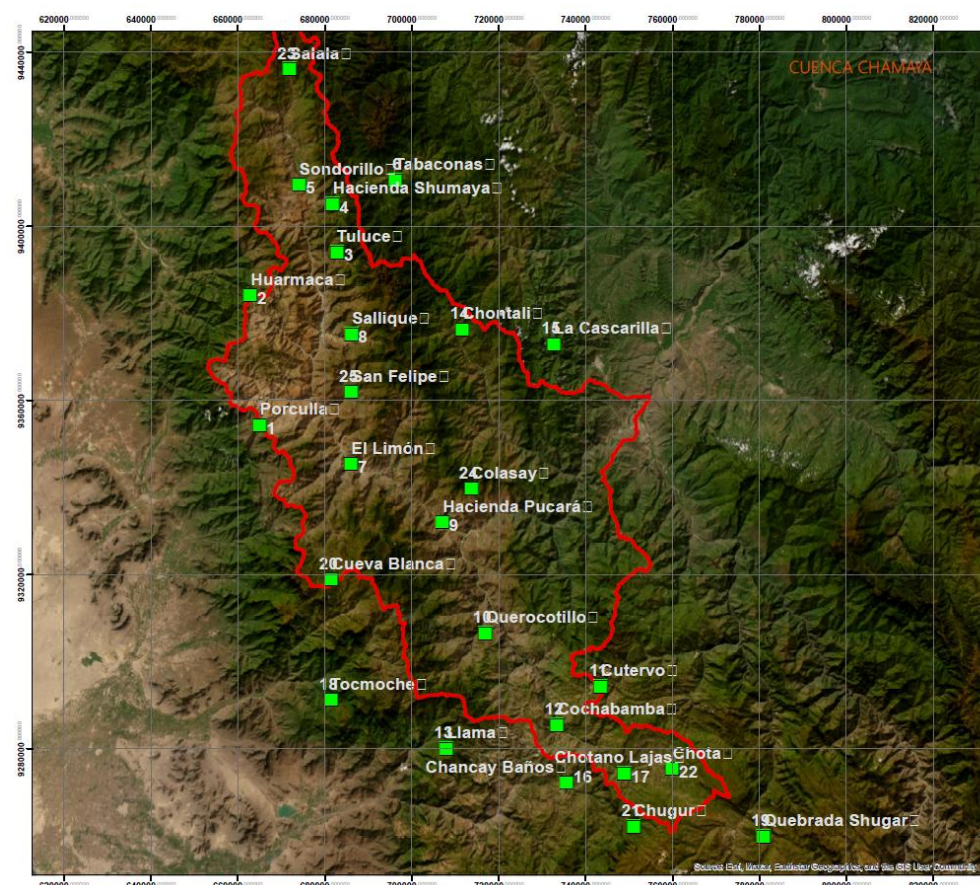


Figura 13

Estaciones meteorológicas de la cuenca Chamaya



3.6.3. Obtención de información pluviométrica y construcción de series de PMDA

Se recopilaron los datos de precipitación diaria de las estaciones seleccionadas desde las fuentes oficiales disponibles para consulta y/o solicitud. Posteriormente, para cada estación se construyó la serie anual de precipitaciones máximas diarias anuales (PMDA) mediante la extracción del máximo diario de cada año hidrológico/calendario, según la estructura de la base de datos obtenida. Esta etapa permitió estandarizar el insumo, asegurando que todas las estaciones contaran con series comparables de máximos anuales antes de aplicar los procedimientos regionales con L-momentos.

a) Definiciones y ecuaciones de trabajo.

Sea $P_{i,d,y}$ la precipitación diaria en la estación i , día d , año y .

La PMDA anual se define como:

$$PMDA_{i,y} = \max_{d \in y}(P_{i,d,y}) \quad (71)$$

La serie PMDA para cada estación es:

$$\{PMDA_{i,y}\}_{y=1}^{n_i} \quad (72)$$

Donde n_i es el número de años con PMDA válida para la estación i .

b) Regla operativa frente a faltantes.

- Si un año presenta ausencia severa de datos diarios (p. ej., meses críticos sin registro), se marca $PMDA_{i,y}$ como faltante.
- Caso contrario, se calcula el máximo con lo disponible, registrando trazabilidad del año.

c) Herramientas.

- Microsoft Excel 365: limpieza inicial, unificación de formatos de fecha y valores.
- R / RStudio: cálculo masivo de PMDA desde series diarias (si se trabajó con scripts) o validación estadística.

3.6.4. Control de calidad y depuración preliminar de las series (análisis exploratorio)

Con las series de PMDA consolidadas, se efectuó un análisis exploratorio orientado a identificar inconsistencias temporales que pudieran sesgar los estadísticos regionales. En esa fase se realizaron dos verificaciones principales: (i) evaluación de tendencia para detectar cambios sistemáticos en el tiempo y (ii) evaluación de homogeneidad para identificar saltos o cambios en la media. En paralelo, se ejecutó una inspección visual de series temporales para reconocer periodos con datos faltantes o comportamientos atípicos. Cuando se detectaron inconsistencias, se aplicaron procedimientos de corrección/ajuste definidos en la investigación y se documentaron los cambios en anexos técnicos para asegurar transparencia en la depuración.

3.6.4.1. Análisis exploratorio descriptivo.

Se calcularon: media, desviación estándar, coeficiente de variación, cuartiles, IQR y detección básica de atípicos (boxplot).

3.6.4.2. Pruebas de tendencia / estacionariedad.

Se aplicaron tres verificaciones por estación con $\alpha=0.05$, contrastando el estadístico con el valor crítico (se usó umbral aproximado ± 1.96 para decisión).

(a) Mann–Kendall (MK)

Para una serie $x_1, \dots, x_n$, el estadístico:

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{sgn}(x_j - x_i) \quad (73)$$

Y el estadístico normalizado:

$$Z = \begin{cases} \frac{S-1}{\sqrt{\text{Var}(S)}}, & S > 0 \\ 0, & S = 0 \\ \frac{S+1}{\sqrt{\text{Var}(S)}}, & S < 0 \end{cases} \quad (74)$$

Decisión: hay tendencia si $Z > Z_{1-\alpha/2}$.

(b) Spearman (correlación rango-tiempo)

$$\rho_s = 1 - \frac{6 \sum d_k^2}{n(n^2-1)} \quad (75)$$

Donde d_k es diferencia entre rangos. Se contrasta significancia de ps.

(c) Tendencia en la media (enfoque por contraste)

Se evaluó cambio sistemático, reportando aceptación/rechazo.

Procedimiento de corrección recomendado.

Para remover tendencia lineal sin alterar el nivel medio, puede usarse detrend:

Ajuste lineal $x_t = a + bt + \epsilon_t$

Serie corregida:

$$x_t^* = x_t - (a + bt) + \bar{x} \quad (76)$$

3.6.4.3. Pruebas de homogeneidad (saltos en media/varianza).

Se aplicaron tres contrastes:

- Mann–Whitney (diferencia en distribución/mediana entre segmentos).
- t de Student (diferencia de medias).
- F de Fisher (diferencia de varianzas).

La regla de decisión se basó en aceptación/rechazo por test y una decisión final integrada.

a) Procedimiento de corrección recomendado (homogeneización por cambio en media).

Si se identifica un quiebre en el año y_b (por inspección/contraste), se ajusta un segmento para igualar el nivel:

Factor multiplicativo:

$$k = \frac{\bar{x}_{ref}}{\bar{x}_{seg}} \rightarrow x_t^* = kx_t \quad (77)$$

o corrección aditiva si corresponde (según naturaleza del sesgo).

3.6.5. Preparación de atributos y matriz de trabajo para regionalización

Luego de depurar las series, se construyó una matriz integrada por estación, que incluyó: PMDA anual, número de años válidos, y atributos geográficos (latitud, longitud y altitud), además de variables de ubicación relativa respecto a la cuenca. Esta matriz se preparó específicamente para sustentar la regionalización, ya que los atributos geográficos se utilizaron en la conformación preliminar de grupos y en la comparación del desempeño entre el enfoque ARF y el enfoque ROI.

a) Variables/atributos usados.

- Localización: longitud, latitud.
- Fisiografía: altitud (msnm).
- Estadística de PMDA: n, x, CV.
- L-momentos y ratios: l1, t, t3, t4.

b) Software.

- R / RStudio: matrices, estandarización, clustering (Ward, k-means).
- Excel: control tabular y verificación.

c) Distancias espaciales entre estaciones

Para análisis de proximidad y vecinos más cercanos, se construyó la matriz de distancias. Una formulación estándar es Haversine:

$$d = 2R \arcsin \left(\sqrt{\sin^2 \left(\frac{\Delta\varphi}{2} \right) + \cos \varphi \cos \varphi_2 \sin^2 \left(\frac{\Delta\lambda}{2} \right)} \right) \quad (78)$$

Donde R es el radio terrestre, φ latitud, y λ longitud en radianes.

d) Estandarización para clustering

Cuando se mezclan variables con distintas unidades, se recomienda estandarizar:

$$z = \frac{x - \mu}{\sigma} \quad (79)$$

3.6.6. Regionalización mediante L-momentos con Análisis Regional de Frecuencia (ARF)

El Análisis Regional de Frecuencia (ARF) se refiere al conjunto de procedimientos mediante los cuales se examina la distribución de frecuencias de muestras de datos procedentes de diferentes ubicaciones dentro de una zona de estudio denominada región (Hosking y Wallis, 1997).

Con la base de datos lista, se inició la regionalización aplicando el método de L-momentos en el marco del esquema de Análisis Regional de Frecuencia (ARF). Primero, se calcularon los L-momentos y las ratios de L-momentos para cada estación a partir de su serie de PMDA, generando los insumos para las pruebas regionales. Luego, se realizó una agrupación preliminar de estaciones mediante técnicas de clasificación (aglomeración jerárquica y partición), con el propósito de delinear regiones tentativas con similitud espacial e hidrometeorológica. Posteriormente, para cada región tentativa se evaluaron la discordancia (consistencia de estaciones dentro del grupo) y la heterogeneidad (grado de homogeneidad regional), y se realizaron ajustes iterativos de composición regional (traslado o exclusión de estaciones) hasta alcanzar regiones con desempeño aceptable en ambos criterios. Este proceso se desarrolló como una rutina sistemática de refinamiento, evitando incorporar resultados numéricos en esta etapa metodológica y reservando su presentación para el capítulo de resultados.

Software utilizado.

- R / RStudio: se utilizó para el agrupamiento Ward y K-means, así como para el cálculo y la verificación de L-momentos.
- HydroOffice 2023: se utilizó para sistematizar el flujo del ARF y revisar los indicadores regionales de discordancia, heterogeneidad y ajuste.
- Excel: tablas finales de cuantiles y reportes.

- QGIS/ArcGIS: verificación espacial de regiones.

3.6.6.1.Cálculo de L-momentos y ratios.

Para cada estación, con la muestra ordenada $x_{1:n} \leq \dots \leq x_{n:n}$, los L-momentos λ_r pueden obtenerse a partir de PWM y combinaciones lineales. Para fines operativos (tesis), lo esencial es declarar:

$$l_1 = \lambda_1(\text{L-media}),$$

$$l_2 = \lambda_2(\text{L-escala}),$$

y los ratios:

$$t = \frac{l_2}{l_1}, t_3 = \frac{l_3}{l_2}, t_4 = \frac{l_4}{l_2} \quad (80)$$

3.6.6.2.Agrupación preliminar (Ward y K-means)

- Construcción de matriz de atributos (estandarizada).
- Clustering jerárquico método Ward (ward.D2) para sugerir estructura natural (dendrograma).
- Consolidación con k-means con $k=4$ para asignación preliminar de regiones.
- Esto produjo 4 regiones iniciales desbalanceadas (incluida una de tamaño muy pequeño), que luego se ajustaron.

3.6.6.3.Prueba de discordancia (D).

Para cada estación i en una región, se usa el vector:

$$u_i = (t_i, t_3i, t_4i) \quad (81)$$

y la medida de discordancia:

$$D_i = \frac{1}{3}(u_i - \bar{u})^T S^{-1}(u_i - \bar{u}) \quad (82)$$

Donde $\bar{u}$ es el promedio regional y S es la matriz de covarianza regional de u .

Criterio aplicado: no discordante si $D_i < 3$ (umbral típico usado en ARF).

3.6.6.4.Prueba de heterogeneidad (H).

Se evalúa dispersión regional de t (L-CV) y se compara con simulaciones; el indicador más usado:

$$H1 = \frac{V - \mu v}{\sigma v} \quad (83)$$

donde V es la variabilidad observada regional y μV , σV provienen de simulación bajo región homogénea.

Criterio aplicado: región aceptable si $H1 < 2$.

3.6.6.5.Ajuste iterativo y regiones definitivas.

Dado que las regiones preliminares fueron heterogéneas (valores H1 altos), se aplicó una rutina iterativa:

- mover estaciones entre regiones,
- redistribuir regiones pequeñas,
- re-evaluar D y H tras cada cambio,

Hasta obtener 4 regiones finales con tamaños 8, 8, 4 y 4 estaciones, cumpliendo simultáneamente $D_{max} < 3$ y $H1 < 2$.

Figura 14

Detalles regionales ARF – Región 1: L-momentos regionales, discordancia de estaciones y distribución espacial (longitud–latitud).

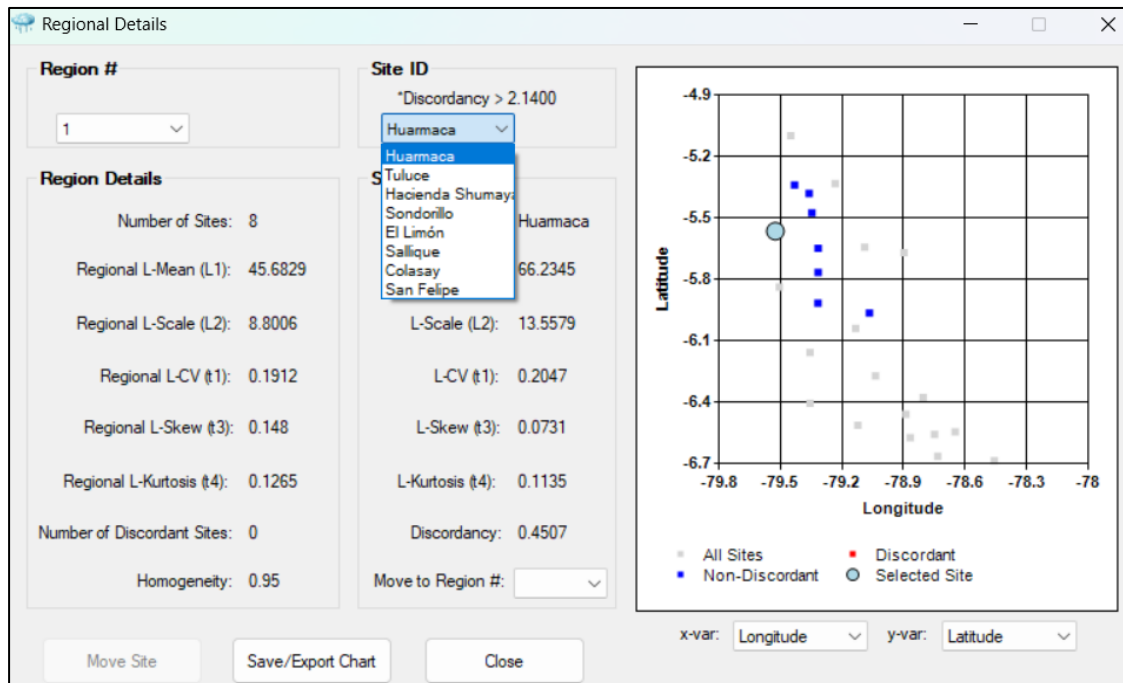


Figura 15

Detalles regionales ARF – Región 2: L-momentos regionales, discordancia de estaciones y distribución espacial (longitud–latitud)

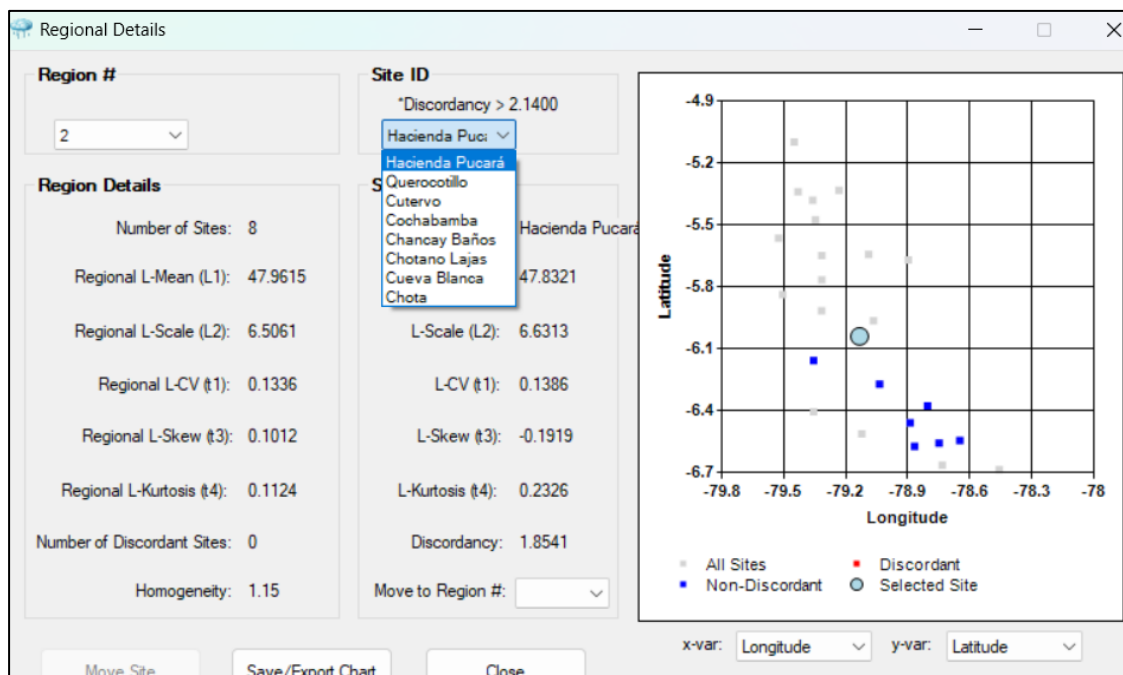


Figura 16

Detalles regionales ARF – Región 3: L-momentos regionales, discordancia de estaciones y distribución espacial (longitud–latitud)

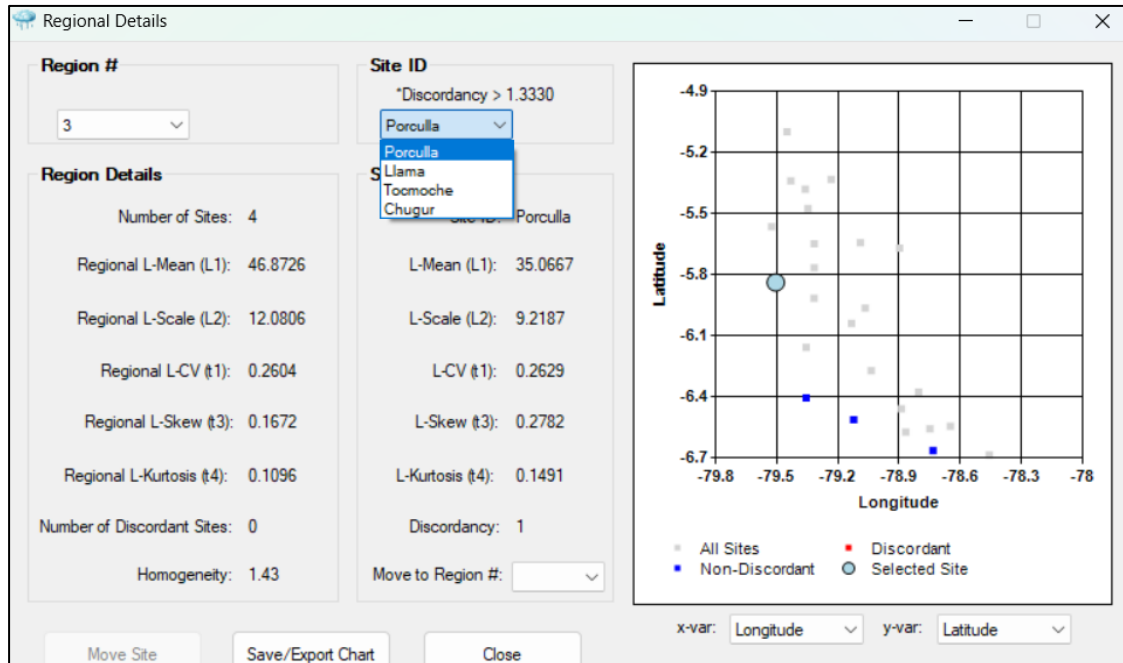
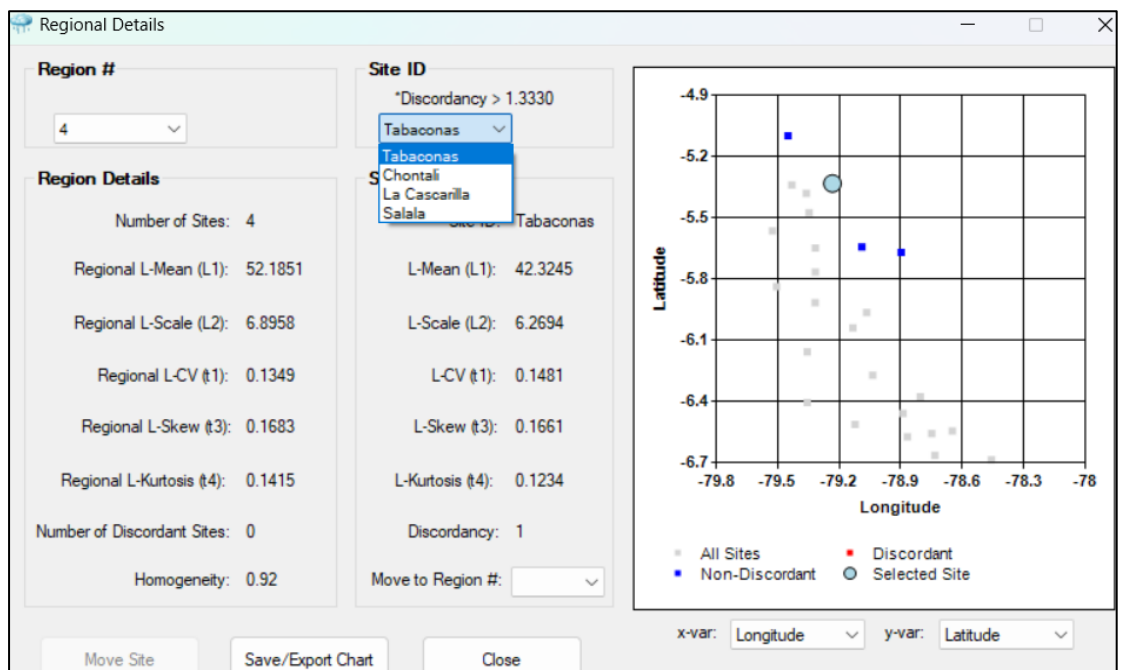


Figura 17

Detalles regionales ARF – Región 4: L-momentos regionales



3.6.6.6. Selección de distribución regional (prueba Z).

Para seleccionar la familia de distribución regional, se utilizó el estadístico:

$$Z_{DIST} = \frac{t_4^{obs} - t_4^{DIST}}{\sigma_{t_4}} \quad (84)$$

Criterio aplicado: $|Z_{DIST}| \leq 1.64$. Se evaluaron LNO, P3, GEV y GLO; se descartó GLO y se eligió LNO para todas las regiones por menor $|Z|$ y consistencia.

3.6.6.7. Estimación de cuantiles locales mediante ajustes probabilísticos por estación.

Para disponer de una referencia local independiente, se efectuó el análisis de frecuencia de cada estación a partir de su serie corregida de precipitaciones máximas diarias anuales. En cada caso se estimaron los L-momentos muestrales y se evaluaron distribuciones probabilísticas candidatas. La función con mejor ajuste se seleccionó en función de su correspondencia con los datos observados y de su comportamiento en los periodos de retorno analizados. Debido a que la distribución óptima no fue necesariamente la misma para todas las estaciones, los cuantiles obtenidos se interpretaron como estimaciones locales y se utilizaron como línea base para comparar el desempeño del Análisis Regional de Frecuencias y del enfoque de Región de Influencia.

La estimación local se expresó mediante:

$$P_{i,LOC}(T) = F_i^{-1} \left(1 - \frac{1}{T} \right) \quad (85)$$

Donde: $P_{i,LOC}(T)$: cuantil local estimado para la estación i y el periodo de retorno T , expresado en mm. F_i^{-1} : función cuantil inversa de la distribución probabilística seleccionada para la estación i . T : periodo de retorno, expresado en años. $1-1/T$: probabilidad de no excedencia asociada al periodo de retorno.

Los cuantiles locales se emplearon como resultados diagnósticos y comparativos. Cuando el ajuste produjo valores negativos en la cola inferior para

periodos de retorno extremadamente próximos a 1 año, estos se conservaron únicamente como evidencia del comportamiento matemático de la distribución y se excluyeron de cualquier aplicación hidráulica.

3.6.6.8. Curva regional adimensional y cuantiles regionalizados mediante ARF (Index-Flood).

Luego de verificar la homogeneidad de las regiones y seleccionar la distribución regional de mejor ajuste, se calcularon las curvas regionales adimensionales de crecimiento. Para las cuatro regiones se adoptó la distribución LNO, debido a que presentó los menores valores absolutos del estadístico Z_{DIST} y cumplió el criterio de aceptación establecido. Cada curva regional expresó la relación entre el periodo de retorno y el cuantil adimensional correspondiente.

Los cuantiles regionalizados se estimaron mediante el procedimiento index-flood. En este enfoque, las estaciones de una misma región comparten la forma de la distribución regional y difieren únicamente por un factor de escala local. Para las precipitaciones máximas diarias anuales, dicho factor correspondió al L-momento de primer orden $l_{1,i}$, equivalente a la media de la serie de la estación.

$$P_{i,ARF}(T) = l_{1,i} \times q_R(T) \quad (86)$$

Donde: $P_{i,ARF}(T)$: precipitación máxima diaria anual regionalizada para la estación i y el periodo de retorno T , expresada en mm. $l_{1,i}$: L-momento de primer orden o media de la serie PMDA de la estación i , expresado en mm. $q_R(T)$: cuantil regional adimensional de la región homogénea R para el periodo de retorno T . R : región homogénea a la que pertenece la estación. T : periodo de retorno, expresado en años.

Probabilidad de no excedencia y variable reducida

$$F(TR) = 1 - \frac{1}{TR} \quad (87)$$

$$y(TR) = -\ln [-\ln (F(TR))]$$

Finalmente, cuando fue necesario expresar los cuantiles como precipitaciones máximas acumuladas durante una ventana de observación continua de 24 horas, se aplicó el factor de corrección por intervalo fijo de observación.

Conversión opcional a precipitación máxima de 24 horas continuas

$$P_{24h,m}(T) = 1.13 \times P_m(TR) \quad (89)$$

Donde: $P_{24h,m}(TR)$: precipitación máxima de 24 horas continuas estimada mediante el método m , en mm. $P_m(TR)$: PMDA estimada mediante el ajuste local, ARF o ROI. m : método aplicado: local, ARF o ROI. 1.13: factor de corrección por intervalo fijo de observación.

Las precipitaciones máximas diarias anuales se obtuvieron a partir de registros acumulados en intervalos de observación fijos. Esta condición introduce una limitación: un evento de lluvia intensa puede iniciarse antes del cierre de una jornada de medición y continuar después de ese momento, quedando dividido entre dos registros diarios consecutivos. En ese caso, el mayor valor diario observado no representa necesariamente la precipitación acumulada máxima en una ventana de 24 horas continua. Para reducir esta subestimación asociada a la discretización temporal, se aplicó el factor de corrección por intervalo fijo de observación ($COITF = 1.13$). El SENAMHI emplea este coeficiente para convertir las precipitaciones máximas diarias observadas en valores representativos de precipitación máxima anual en 24 horas continuas; su magnitud se sustenta en las consideraciones teóricas desarrolladas por Weiss (1964). De manera concordante, la Organización Meteorológica Mundial recomienda incrementar los datos obtenidos en intervalos fijos mediante el factor 1.13 cuando se requiere estimar la precipitación acumulada máxima en 24 horas continuas (SENAMHI, 2014; WMO, 1973).

3.6.7. Regionalización mediante L-momentos por Región de Influencia (ROI/ RI)

En forma complementaria, se aplicó el enfoque de Región de Influencia (ROI) para estimar cuantiles en función de regiones “móviles” asociadas a cada estación objetivo. Para ello, se definió, para cada estación, una región de influencia conformada por estaciones vecinas, según su cercanía en el espacio de atributos (variables geográficas). Con la región ROI definida, se recalcularon los L-momentos regionales ponderados y se verificaron nuevamente los criterios de discordancia y heterogeneidad, a fin de asegurar consistencia regional en cada vecindario. Finalmente, se estandarizaron los productos para que la comparación ARF vs. ROI se realizara con criterios equivalentes y con trazabilidad completa de las estaciones incluidas en cada caso.

$$P_{i,ROI}(T) = l_{1,i} \times q_{ROI,i}(T) \quad (90)$$

Donde: $P_{i,ROI}(T)$: cuantil estimado mediante Región de Influencia para la estación objetivo i y el periodo de retorno T , expresado en mm. $l_{1,i}$: L-momento de primer orden o media PMDA de la estación objetivo. $q_{ROI,i}(T)$: cuantil adimensional calculado para el vecindario hidrológico específico de la estación i . T : periodo de retorno, expresado en años.

En forma complementaria, se estimaron cuantiles mediante el enfoque de Región de Influencia. Para cada estación objetivo se conformó un vecindario hidrológico específico utilizando estaciones con mayor similitud en el espacio de atributos. Luego se calculó una curva de crecimiento propia para dicho conjunto. A diferencia del ARF fijo, el ROI no impuso una única frontera regional: la composición del grupo varió según la estación analizada. Los cuantiles se obtuvieron escalando la curva adimensional del vecindario con el L-momento de primer orden de la estación objetivo.

a) Concepto operativo.

Para cada estación objetivo s , se define una región de influencia ROI(s) formada por estaciones con mayor similitud en un espacio de atributos (puede incluir estaciones fuera de la región ARF contigua).

b) Distancia en espacio de atributos (forma general).

Con un vector de atributos estandarizados

ai (p. ej., $[t, t3, t4, altitud, lat, lon]$), una distancia euclidiana típica es:

$$dsi = \|as - ai\|_2 \quad (91)$$

c) Procedimiento aplicado en la investigación.

- Selección de estaciones objetivo y conformación de 6 grupos ROI.
- Verificación de consistencia de cada grupo mediante ajuste de distribuciones y comparación gráfica de cuantiles vs TR.
- Evaluación de sensibilidad por distribución (divergencia en TR altos).
- Estimación de cuantiles de diseño con lógica index-flood (análoga al ARF), usando el grupo ROI como “región” específica.

3.6.8. Sistematización para cartografía regional de PMDA

Como etapa final del proceso para la entrega de productos espaciales, se integraron: (i) las regiones homogéneas verificadas, (ii) los atributos geográficos de estaciones y (iii) los cuantiles regionalizados por periodos de retorno definidos en el estudio. Esta integración permitió estructurar la información para elaborar mapas regionales de PMDA, caracterizar espacialmente las regiones homogéneas altoandinas e interpretar la variabilidad espacial en función de la regionalización alcanzada.

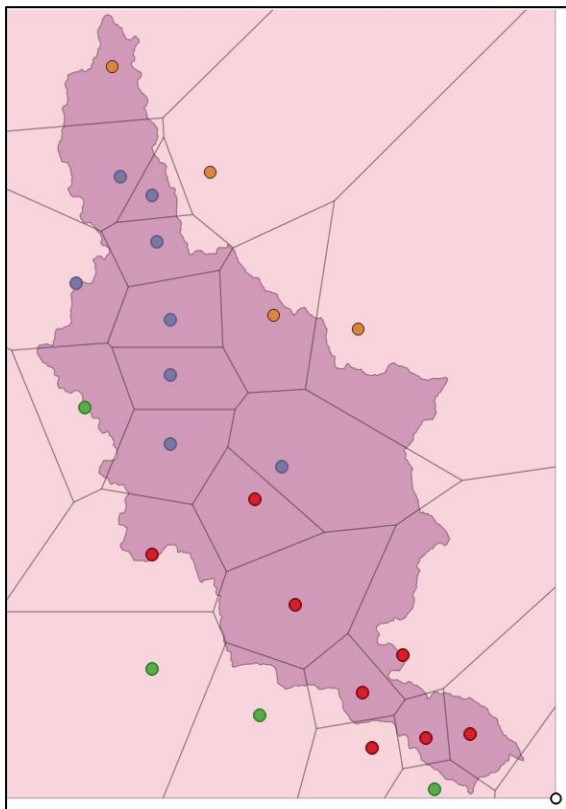
a) Polígonos de Thiessen (Voronoi) y áreas de influencia

- Procedimiento en QGIS/ArcGIS.
- Generar Voronoi/Thiessen a partir de los puntos de estaciones.

- Interceptar/recortar con el polígono de la cuenca Chamaya.
- Calcular el área de cada polígono A_i .

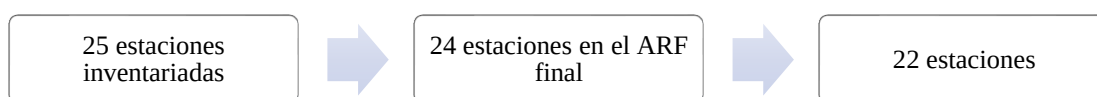
Figura 18

Polígono de Thiessen



Los polígonos de Thiessen se generaron a partir de las estaciones utilizadas como soporte espacial y, posteriormente, se interceptaron con el límite oficial de la cuenca del río Chamaya. Este recorte permitió cuantificar únicamente el área efectiva de cada estación dentro del ámbito de estudio. La regionalización estadística final mediante ARF incluyó 24 estaciones; sin embargo, la espacialización cartográfica contabilizó 22 estaciones con área efectiva positiva dentro de la cuenca. Tocmoche, asignada a la Región 3, y Tabaconas, perteneciente a la Región 4, se conservaron en el análisis estadístico regional, pero sus polígonos no aportaron superficie tras el recorte por el límite de Chamaya. La estación Quebrada Shugar permaneció en el inventario inicial, aunque no fue incorporada a una región homogénea definitiva.

La secuencia general de la aplicación de las estaciones fue la siguiente:



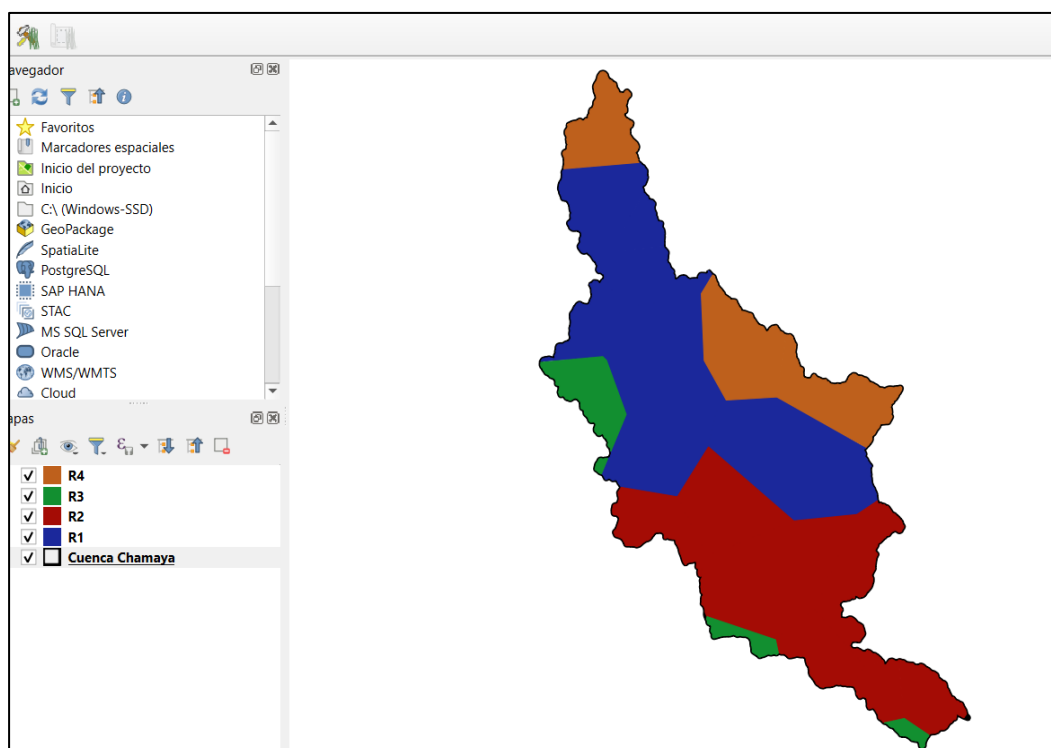
La reducción no obedeció a una pérdida arbitraria de información. Quebrada Shugar quedó documentada en el inventario inicial, pero no integró una región homogénea definitiva. Tocmoche y Tabaconas se conservaron en el ARF por su aporte estadístico, aunque no generaron superficie cartográfica efectiva dentro del límite de la cuenca tras el recorte espacial.

b) Construcción de mapa de regiones homogéneas ARF

- Se asignaron a cada estación su Región ARF (R1–R4).
- Se asociaron la región al polígono Thiessen correspondiente.
- Se disolvieron polígonos por región para obtener superficies continuas regionales (mapa de regiones).

Figura 19

Creación del mapa de regiones homogéneas



c) Mapas de PMDA regionalizada por periodo de retorno

Para cada Tr (por ejemplo 50, 100, 200, 500 años):

- Se obtuvo el cuantil regionalizado (por región o por estación según el producto final definido).
- Se asignaron valores a polígonos (Thiessen por estación).
- Se clasificaron en intervalos (mm) y simbolizar.
- Se exportaron mapas finales.

3.7. Procedimientos de análisis de datos

Para el procesamiento de los datos se organizó la información en Microsoft Excel 365, donde se consolidaron los registros diarios por estación, se estandarizaron los campos (ID, nombre, coordenadas, altitud y fechas), y se generó la serie de máximos anuales (AMS/PMDA) por año hidrológico, además de identificarse vacíos y posibles inconsistencias para su depuración preliminar. Posteriormente, las series AMS y sus metadatos se importaron a R (versión 4.2.2), empleando los paquetes *lmom* y *lmomco* para estructurar las bases por estación y por región, calcular estadísticos descriptivos y preparar insumos para el análisis regional. En paralelo, HydroOffice 2023 se utilizó para ordenar el flujo de regionalización (cálculo de métricas regionales y verificación de agrupaciones), mientras que QGIS y ArcMap 10.8 se emplearon para la gestión espacial (verificación de estaciones en el ámbito de cuenca, generación de capas temáticas y preparación de salidas cartográficas).

El análisis estadístico se desarrolló comparando el comportamiento empírico de las series AMS/PMDA con distribuciones teóricas mediante la prueba de Smirnov–Kolmogorov, permitiendo aceptar o rechazar el ajuste del modelo propuesto a un nivel de significancia del 0.05. Con las series depuradas, se aplicó el método de L-momentos para estimar parámetros de distribución y robustecer el análisis de extremos, dado su

desempeño ante heterogeneidad y presencia de valores extremos (Hosking J. R., 1990); en seguida, se ejecutó el esquema de regionalización por ARF, estructurando regiones homogéneas, estimando cuantiles regionales adimensionales y cuantiles locales por estación para distintos periodos de retorno, y contrastando resultados entre enfoques ARF y ROI. Finalmente, los productos del análisis (regiones, cuantiles y métricas de verificación) se integraron en el SIG para su representación espacial y comparación cartográfica, asegurando coherencia en la delimitación regional.

3.8. Material y equipos (descripción del uso)

Para la obtención de datos se emplearon insumos cartográficos y bases oficiales de precipitación diaria, además de equipos informáticos y software SIG–estadístico para depurar registros, construir las series de PMDA, verificar su consistencia y generar productos espaciales (mapas de estaciones, regiones ARF/ROI y superficies de influencia), la siguiente tabla resume los materiales, equipos y programas utilizados.

Tabla 13

Materiales y equipos utilizados para el proceso de obtención de datos

Material / equipo	Especificación / versión	Descripción del uso en el estudio
Computadora / laptop	Equipo de escritorio o portátil	Se utilizó para ejecutar el SIG, el procesamiento estadístico y la organización de bases de datos, asegurando trazabilidad de archivos y reproducibilidad de resultados.
Conexión a internet	Acceso estable	Se empleó para la descarga y/o consulta de información oficial de estaciones y registros de precipitación, así como para obtener cartografía base disponible.
Navegador web	(Chrome)	Se utilizó para acceder a SENAMHI Data y a repositorios institucionales para descargar series, metadatos y capas cartográficas.
Base de datos pluviométrica	Registros diarios por estación (SENAMHI)	Se recopiló la precipitación diaria para cada estación y se estructuró por periodos de registro para construir las series AMS/PMDA.
Inventario maestro de estaciones	Tabla (XLSX/CSV)	Se consolidaron atributos técnicos por estación y se verificó la completitud y consistencia preliminar del registro antes del análisis regional.
Polígono de cuenca Chamaya	Vector (SHP/GeoPackage)	Se utilizó como límite espacial para verificar la relación estación–cuenca, recortar capas y generar mapas oficiales del ámbito de estudio.
Cartografía base	Hidrografía, límites administrativos	Se incorporó para contextualización espacial, lectura cartográfica y control visual de coherencia geográfica del proyecto.
Modelo Digital de Elevación (DEM)	Raster (según disponibilidad)	Se empleó para contextualizar el gradiente altitudinal y apoyar la caracterización fisiográfica del ámbito, mediante productos derivados (sombreado/altitud).

Material / equipo	Especificación / versión	Descripción del uso en el estudio
QGIS	3.40	Se utilizó para gestión espacial (CRS, reproyección, corrección geométrica), importación de estaciones, verificación de cobertura, Voronoi/Thiessen, intersecciones, disoluciones y cálculo de áreas.
ArcMap	10.8	Se empleó como herramienta complementaria para geoprocursos equivalentes (Thiessen, disolución, recortes) y validación cruzada de resultados SIG cuando correspondió.
Microsoft Excel	365	Se utilizó para depuración preliminar de bases (formato de fechas, vacíos, control de rangos), organización tabular y verificación rápida de consistencia antes de SIG y R.
R	4.2.2	Se utilizó para el procesamiento estadístico: construcción masiva de series PMDA, cálculo de estadísticos descriptivos, L-momentos y razones-L, y soporte de rutinas de regionalización.
Paquetes de R	lmom y lmomco	Se emplearon para estimación de L-momentos/razones-L, ajuste de distribuciones y apoyo en la generación de cuantiles regionales y locales.
HydroOffice	2023	Se utilizó para sistematizar el flujo de regionalización y gestionar resultados operativos del análisis regional, según el esquema adoptado en el estudio.
Almacenamiento externo / respaldo	Disco externo o nube (según disponibilidad)	Se utilizó para respaldo de bases originales, versiones depuradas, proyectos SIG y resultados finales, evitando pérdida de información.
Exportador a PDF/imagen	Herramienta de exportación	Se empleó para generar figuras y mapas finales con calidad de presentación (layout, escala, norte, leyenda).

3.9. Aspectos éticos

En la investigación se tomaron en cuenta los planteamientos del código de ética de la investigación científica, según el capítulo II, artículo 5° de la actividad investigadora de la Universidad Nacional Autónoma de Chota.

- Formulación y ejecución de proyectos de investigación que contribuyan al bienestar de la población en general.
- Protección al medio ambiente y el respeto a la biodiversidad.
- Responsabilidad, rigor científico y veracidad. Se actuó con responsabilidad en relación con la pertinencia, alcances y repercusiones de la investigación.
- Investigación exhaustiva y desde diferentes enfoques, aplicando el pensamiento crítico a todas las consultas e investigaciones.
- Respeto a los derechos de propiedad intelectual de otros investigadores con el uso estricto de las normas de publicación pertinentes.
- Honestidad y veracidad científica al reportar los resultados de la investigación, garantizando la autenticidad y evitando el riesgo de plagio.

- Justicia y bien común. Se antepuso la justicia y el bien común al interés personal, evitando efectos nocivos que pueda generar la investigación en las personas, medio ambiente y sociedad en general.
- Divulgación responsable de la investigación. Es obligatorio para los docentes investigadores, docentes, estudiantes, graduados, autoridades y personal administrativo de la UNACH que realizan investigación, difundir y publicar los resultados de las investigaciones realizadas en un ambiente de ética, pluralismo ideológico y diversidad cultural.

CAPÍTULO IV: RESULTADOS

4.1. Descripción de resultados

4.1.1. *Análisis de las series de PMDA de las estaciones meteorológicas seleccionadas*

a) Inventario de estaciones, ubicación y fuente de registros PMDA (24 h)

La Tabla 14 evidencia que el análisis de PMDA se sustentó en 25 estaciones meteorológicas convencionales (CM) distribuidas principalmente en la cuenca del río Chamaya (Figura 20), con cobertura espacial entre longitudes -79.53 a -78.46 y latitudes -6.69 a -5.10 , y un rango altitudinal amplio de 1012 a 3300 msnm, lo cual permite evaluar máximos diarios anuales bajo condiciones topográficas contrastantes. En términos de representatividad hidrológica, la mayoría de estaciones se ubicó dentro de la cuenca Chamaya (19/25), mientras que 6 estaciones se incorporaron como soporte espacial por su proximidad y relación con cuencas colindantes (Chinchipe al este, Chancay Lambayeque al sur, Motupe al oeste, IAM III al este e IAM IV al sur), lo que refuerza el análisis regional en bordes hidrológicos. Respecto a la fuente de información, los registros se obtuvieron por dos vías: descarga directa desde SENAMHI Data (✓) y solicitud formal con periodos explícitos, predominando en esta última el intervalo 1995–2024, aunque también se observa una estación con 2008–2024 (Porculla), lo que confirma que la longitud de los registros no fue uniforme entre estaciones y que se requirió integrar fuentes para consolidar la base de datos de PMDA.

Figura 20

Ubicación de las estaciones meteorológicas

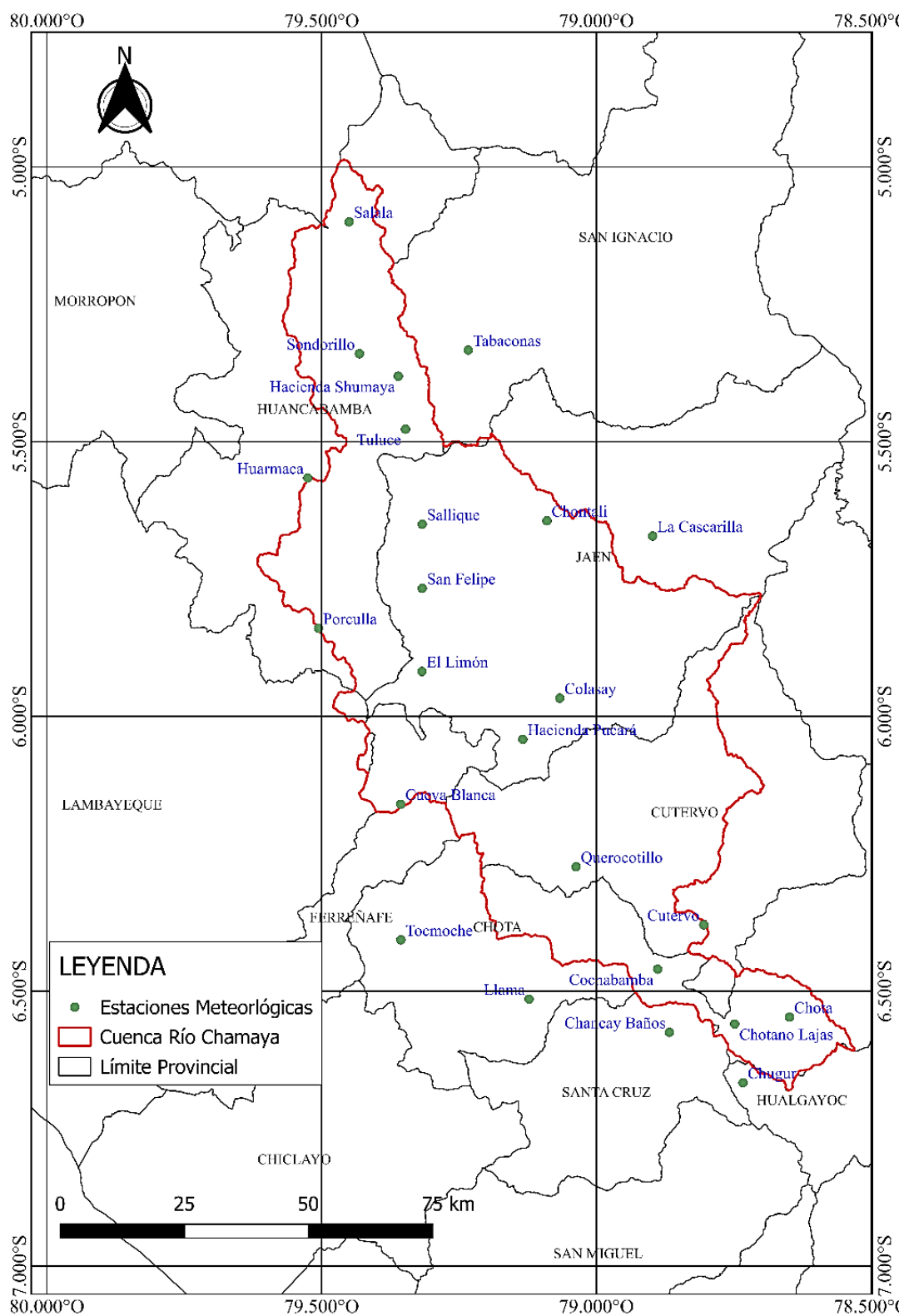


Tabla 14

Inventario, ubicación geográfica y fuente de datos de las estaciones meteorológicas empleadas (PMDA)

Id	Código	Estación	Región	Provincia	Distrito	Longitud	Latitud	Altitud (msnm)	Estación Tipo	Obtención de datos	Cuenca	Observación
1	qc00152132	Porculla	Piura	Huancabamba	Huarmaca	-79.51	-5.84	2142	CM	2008-2024	Chamaya	Dentro
2	qc00000248	Huarmaca	Piura	Huancabamba	Huarmaca	-79.53	-5.57	2232	CM	√	Chamaya	Dentro
3	qc00152127	Tuluze	Piura	Huancabamba	Sondor	-79.35	-5.48	2209	CM	√	Chamaya	Dentro
4	qc00152128	Hacienda Shumaya	Piura	Huancabamba	Sondor	-79.36	-5.38	1975	CM	√	Chamaya	Dentro
5	qc00152126	Sondorillo	Piura	Huancabamba	Sondorillo	-79.43	-5.34	1917	CM	1995-2024	Chamaya	Dentro
6	qc00000240	Tabaconas	Cajamarca	San Ignacio	Tabaconas	-79.23	-5.33	1902	CM	√	Chinchipe	Este
7	qc00000241	El Limón	Cajamarca	Jaén	Pomahuaca	-79.32	-5.92	1110	CM	1995-2024	Chamaya	Dentro
8	qc00152129	Sallique	Cajamarca	Jaén	Sallique	-79.32	-5.65	1804	CM	√	Chamaya	Dentro
9	qc00153107	Hacienda Pucará	Cajamarca	Jaén	Pucará	-79.13	-6.04	1012	CM	1995-2024	Chamaya	Dentro
10	qc00153109	Querocotillo	Cajamarca	Cutervo	Querocotillo	-79.04	-6.27	1970	CM	1995-2024	Chamaya	Dentro
11	qc00000352	Cutervo	Cajamarca	Cutervo	Cutervo	-78.8	-6.38	2668	CM	1995-2024	Chamaya	Dentro
12	qc00000353	Cochabamba	Cajamarca	Chota	Cochabamba	-78.89	-6.46	1653	CM	1995-2024	Chamaya	Dentro
13	qc00000341	Llama	Cajamarca	Chota	Llama	-79.12	-6.51	2096	CM	√	Chancay Lambayeque	Sur
14	qc00000250	Chontalí	Cajamarca	Jaén	Chontalí	-79.09	-5.64	1627	CM	√	Chamaya	Dentro
15	qc00000349	La Cascarilla	Cajamarca	Jaén	Jaén	-78.9	-5.67	1991	CM	√	IAM III	Este
16	qc00000395	Chancay Baños	Cajamarca	Santa Cruz	Chancay Baños	-78.87	-6.58	1638	CM	√	Chancay Lambayeque	Sur
17	qc00153235	Chotano Lajas	Cajamarca	Chota	Lajas	-78.75	-6.56	2163	CM	√	Chamaya	Dentro
18	qc00153103	Tocmoche	Cajamarca	Chota	Tocmoche	-79.36	-6.41	1399	CM	1995-2024	Motupe	Oeste
19	qc00153108	Quebrada Shugar	Cajamarca	Hualgayoc	Bambamarca	-78.46	-6.69	3292	CM	√	IAM IV	Sur
20	qc00153102	Cueva Blanca	Lambayeque	Ferreñafe	Incahuasi	-79.36	-6.16	3300	CM	1995-2024	Chamaya	Dentro
21	qc00153208	Chugur	Cajamarca	Hualgayoc	Chugur	-78.73	-6.67	2748	CM	√	Chancay Lambayeque	Sur
22	303	Chota	Cajamarca	Chota	Chota	-78.65	-6.55	2468	CM	1995-2024	Chamaya	Dentro
23	203	Salala	Piura	Huancabamba	El Carmen de la Frontera	-79.45	-5.1	2974	CM	√	Chamaya	Dentro
24	152134	Colasay	Cajamarca	Jaén	Colasay	-79.07	-5.97	1950	CM	√	Chamaya	Dentro
25	152130	San Felipe	Cajamarca	Jaén	San Felipe	-79.32	-5.77	1925	CM	√	Chamaya	Dentro

Nota. Aquellos datos con el símbolo √ fueron obtenidos de la data de SENAMHI y aquellos con fechas de datos fueron obtenido de solicitud.

b) Cobertura y proximidad espacial de la red de estaciones (distancias entre estaciones)

La Tabla 15 resume, para cada estación, sus cinco vecinos más cercanos, mostrando que en muchos casos las menores distancias se ubican en el rango aproximado de 10 a 30 km, lo que refleja existencia de grupos de estaciones con proximidad espacial. Sin embargo, también se identifican estaciones para las cuales la distancia al cuarto o quinto vecino se incrementa a valores cercanos a 40–50 km (y en algunos casos superiores), indicando que no todas las estaciones cuentan con la misma densidad de estaciones cercanas; por tanto, el soporte espacial inmediato varía entre estaciones, lo que resulta consistente con la distribución no uniforme observada.

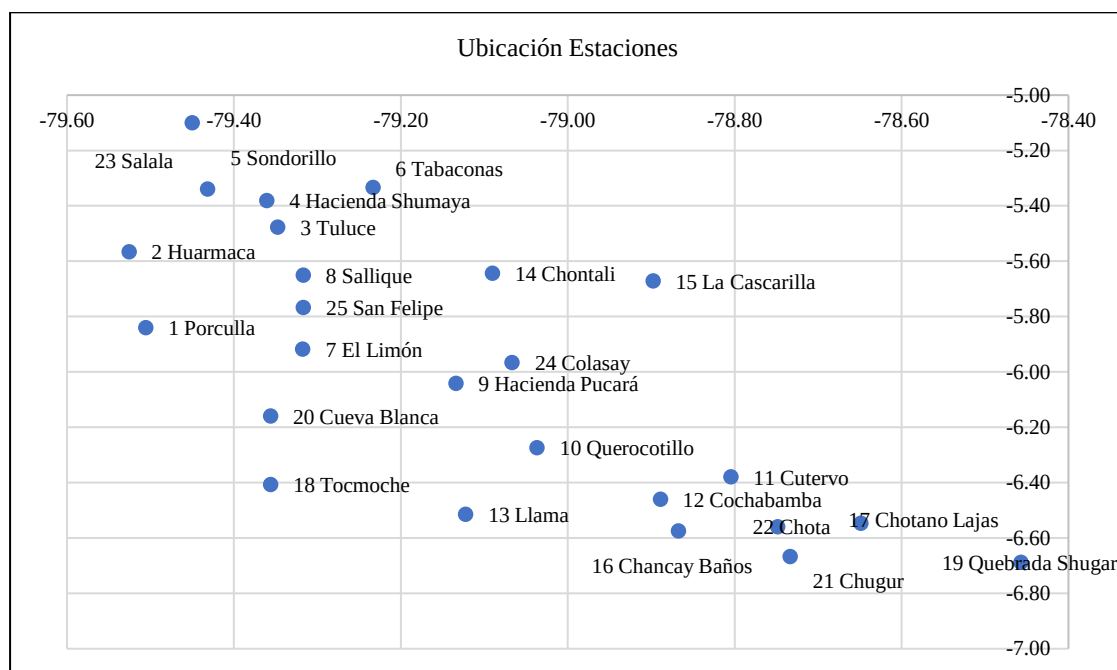
Tabla 15

Estaciones más próximas por estación (cinco menores distancias, km)

Estación	L1	L2	L3	L4	L5
1	22.4	22.6	29.7	30.5	39.3
2	22.1	24.9	27.3	27.5	30.5
3	10.8	17.9	19.6	20.4	22.1
4	9.1	10.8	15.1	27.5	30.4
5	9.1	17.9	22.0	26.8	27.3
6	15.1	20.4	22.0	35.4	36.5
7	16.8	22.6	24.5	27.2	28.3
8	13.0	19.6	24.9	25.2	29.7
9	11.2	24.5	27.9	28.0	36.7
10	26.5	28.0	28.3	28.4	34.3
11	12.9	21.0	22.9	25.4	28.3
12	12.9	13.0	19.1	26.5	26.6
13	26.6	28.4	28.5	29.0	38.3
14	21.5	25.2	28.6	34.0	36.0
15	21.5	37.8	46.5	47.6	48.8
16	13.0	13.3	18.0	22.9	24.4
17	11.1	12.0	13.3	19.1	21.0
18	27.5	28.5	38.3	47.5	
19	26.3	30.6	35.2	47.1	
20	27.2	27.5	27.9	37.5	38.6
21	12.0	16.3	18.0	28.7	30.6
22	11.1	16.3	24.4	25.4	26.3
23	26.8	32.8	35.4	43.5	
24	11.2	28.3	34.3	35.5	36.0
25	13.0	16.8	22.4	28.6	32.2

Figura 21

Distribución espacial de las estaciones meteorológicas utilizadas



Nota. Las 25 estaciones se distribuyen dentro del dominio espacial del estudio con cobertura longitudinal aproximada entre -79.53 y -78.46 y latitudinal entre -6.69 y -5.10 , evidenciando despliegue no uniforme, con concentraciones de puntos en el sector occidental-noroccidental y otra concentración hacia el sector oriental-suroriental. Asimismo, se observan estaciones relativamente más aisladas en los extremos del dominio, lo que indica que la cobertura espacial combinó zonas con alta densidad de estaciones y zonas con mayor separación entre puntos.

La Tabla 16 evidencia amplia variación de separaciones entre estaciones, con distancias mínimas del orden de 9.1 km (entre estaciones cercanas) y distancias máximas que alcanzan aproximadamente 208.2 km, lo que confirma que la red utilizada integra estaciones muy próximas entre sí y otras ampliamente separadas. En conjunto, se observa que una fracción importante de pares presenta distancias menores a 50 km, mientras que los valores más altos se concentran en combinaciones que involucran estaciones ubicadas en extremos opuestos del dominio espacial, lo cual caracteriza la heterogeneidad espacial del muestreo meteorológico para el análisis de PMDA.

Tabla 16

Matriz de distancias entre estaciones meteorológicas (km)

Estación	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1		30.5	44.0	53.5	56.2	63.9	22.6	29.7	46.9	70.9	98.1	97.1	86.3	50.9	69.9	108.1	116.0	65.3	149.6	39.3	125.6	123.3	82.6	50.6	22.4
2	30.5		22.1	27.5	27.3	41.5	45.5	24.9	68.4	95.6	120.7	122.0	114.6	49.0	70.5	133.9	140.1	95.4	172.0	68.7	150.7	146.1	52.5	67.6	32.2
3	44.0	22.1		10.8	17.9	20.4	49.2	19.6	67.1	95.1	117.0	120.6	118.1	34.0	54.3	133.3	137.6	103.5	167.0	76.0	148.8	142.0	43.5	62.7	32.4
4	53.5	27.5	10.8		9.1	15.1	60.0	30.4	77.7	105.7	127.1	131.0	128.9	41.9	60.7	143.7	147.7	114.2	176.6	86.7	159.1	151.9	32.8	72.9	43.3
5	56.2	27.3	17.9	9.1		22.0	65.6	36.8	84.8	112.8	134.9	138.4	135.2	50.8	69.7	151.0	155.4	119.1	184.8	91.7	166.7	159.9	26.8	80.6	49.2
6	63.9	41.5	20.4	15.1	22.0		65.8	36.5	79.6	106.9	125.7	131.1	132.1	38.0	52.9	144.0	146.7	120.3	173.6	93.0	158.4	149.8	35.4	72.9	49.2
7	22.6	45.5	49.2	60.0	65.6	65.8		29.8	24.5	50.3	76.5	76.8	69.8	39.6	54.0	88.5	95.2	54.6	128.1	27.2	105.5	101.9	92.2	28.3	16.8
8	29.7	24.9	19.6	30.4	36.8	36.5	29.8		48.1	76.0	99.0	101.8	98.6	25.2	46.5	114.3	119.2	84.3	149.7	56.9	130.3	124.3	63.0	44.8	13.0
9	46.9	68.4	67.1	77.7	84.8	79.6	24.5	48.1		28.0	52.4	53.9	52.6	44.5	48.8	66.3	71.7	47.5	103.8	27.9	82.5	77.8	110.5	11.2	36.7
10	70.9	95.6	95.1	105.7	112.8	106.9	50.3	76.0	28.0		28.3	26.5	28.4	70.4	68.8	38.4	45.1	38.3	79.0	37.5	55.2	52.7	138.4	34.3	64.4
11	98.1	120.7	117.0	127.1	134.9	125.7	76.5	99.0	52.4	28.3		12.9	38.3	87.8	79.4	22.9	21.0	61.1	51.5	65.7	32.9	25.4	159.3	54.3	88.7
12	97.1	122.0	120.6	131.0	138.4	131.1	76.8	101.8	53.9	26.5	12.9		26.6	93.5	87.8	13.0	19.1	52.0	54.1	61.6	28.7	28.3	163.7	58.3	90.6
13	86.3	114.6	118.1	128.9	135.2	132.1	69.8	98.6	52.6	28.4	38.3	26.6		97.0	97.1	29.0	41.7	28.5	76.1	47.2	46.3	52.5	161.6	61.3	85.9
14	50.9	49.0	34.0	41.9	50.8	38.0	39.6	25.2	44.5	70.4	87.8	93.5	97.0		21.5	106.5	108.7	89.9	135.7	64.5	120.5	111.8	72.5	36.0	28.6
15	69.9	70.5	54.3	60.7	69.7	52.9	54.0	46.5	48.8	68.8	79.4	87.8	97.1	21.5		100.6	100.2	96.3	123.2	74.3	112.2	101.3	88.3	37.8	47.6
16	108.1	133.9	133.3	143.7	151.0	144.0	88.5	114.3	66.3	38.4	22.9	13.0	29.0	106.5	100.6		13.3	57.2	47.1	71.2	18.0	24.4	176.4	71.2	102.8
17	116.0	140.1	137.6	147.7	155.4	146.7	95.2	119.2	71.7	45.1	21.0	19.1	41.7	108.7	100.2	13.3		69.3	35.2	80.6	12.0	11.1	180.1	74.8	108.4
18	65.3	95.4	103.5	114.2	119.1	120.3	54.6	84.3	47.5	38.3	61.1	52.0	28.5	89.9	96.3	57.2	69.3		104.2	27.5	74.7	79.8	145.8	58.5	71.3
19	149.6	172.0	167.0	176.6	184.8	173.6	128.1	149.7	103.8	79.0	51.5	54.1	76.1	135.7	123.2	47.1	35.2	104.2		115.5	30.6	26.3	208.2	104.9	139.9
20	39.3	68.7	76.0	86.7	91.7	93.0	27.2	56.9	27.9	37.5	65.7	61.6	47.2	64.5	74.3	71.2	80.6	27.5	115.5		89.1	89.4	118.4	38.6	43.9
21	125.6	150.7	148.8	159.1	166.7	158.4	105.5	130.3	82.5	55.2	32.9	28.7	46.3	120.5	112.2	18.0	12.0	74.7	30.6	89.1		16.3	191.6	86.2	119.2
22	123.3	146.1	142.0	151.9	159.9	149.8	101.9	124.3	77.8	52.7	25.4	28.3	52.5	111.8	101.3	24.4	11.1	79.8	26.3	89.4	16.3		183.9	79.5	114.1
23	82.6	52.5	43.5	32.8	26.8	35.4	92.2	63.0	110.5	138.4	159.3	163.7	161.6	72.5	88.3	176.4	180.1	145.8	208.2	118.4	191.6	183.9		105.4	75.7
24	50.6	67.6	62.7	72.9	80.6	72.9	28.3	44.8	11.2	34.3	54.3	58.3	61.3	36.0	37.8	71.2	74.8	58.5	104.9	38.6	86.2	79.5	105.4		35.5
25	22.4	32.2	32.4	43.3	49.2	49.2	16.8	13.0	36.7	64.4	88.7	90.6	85.9	28.6	47.6	102.8	108.4	71.3	139.9	43.9	119.2	114.1	75.7		

c) Estadística descriptiva, completitud y distribución de PMDA por estación

La Tabla 17 muestra que las series de PMDA utilizadas presentan longitudes heterogéneas, con registros entre 21 y 60 años y porcentajes de completitud entre 72.7% y 98.3%, lo que evidencia distinta disponibilidad de información entre estaciones. En términos de magnitud, los promedios oscilaron entre 29.0 mm (Sondorillo) y 75.3 mm (La Cascarilla), mientras que los máximos anuales alcanzaron valores elevados en estaciones puntuales, destacando 145.2 mm (Huarmaca) y 130.8 mm (Tocmoche), lo que confirma la presencia de eventos extremos. Los estadísticos de posición (Q1, mediana y Q3) y el IQR evidencian contrastes en dispersión, con rangos intercuartílicos más amplios en estaciones como Huarmaca (IQR = 37.2) y Tocmoche (IQR = 35.8), frente a estaciones más concentradas como Cutervo (IQR = 10.9). Finalmente, el conteo de valores atípicos fue bajo en la mayoría de estaciones, siendo la estación con mayor presencia la Quebrada Shugar (3 outliers).

El box plot sintetiza que la distribución de PMDA difiere de forma marcada entre estaciones tanto en nivel central como en dispersión: se observan estaciones con medianas más altas, coherentes con valores centrales elevados (por ejemplo, La Cascarilla y Huarmaca), y estaciones con medianas menores (por ejemplo, Sondorillo y San Felipe), lo cual refleja distintos regímenes locales de máximos diarios. Asimismo, la amplitud de las cajas confirma que algunas series presentan mayor variabilidad interanual (cajas y bigotes más extendidos, como en Tocmoche y Huarmaca), mientras que otras son más concentradas (como Cutervo). La presencia de puntos fuera de los bigotes es limitada y aparece en estaciones específicas, lo que coincide con el número reducido de outliers reportado en la Tabla 17.

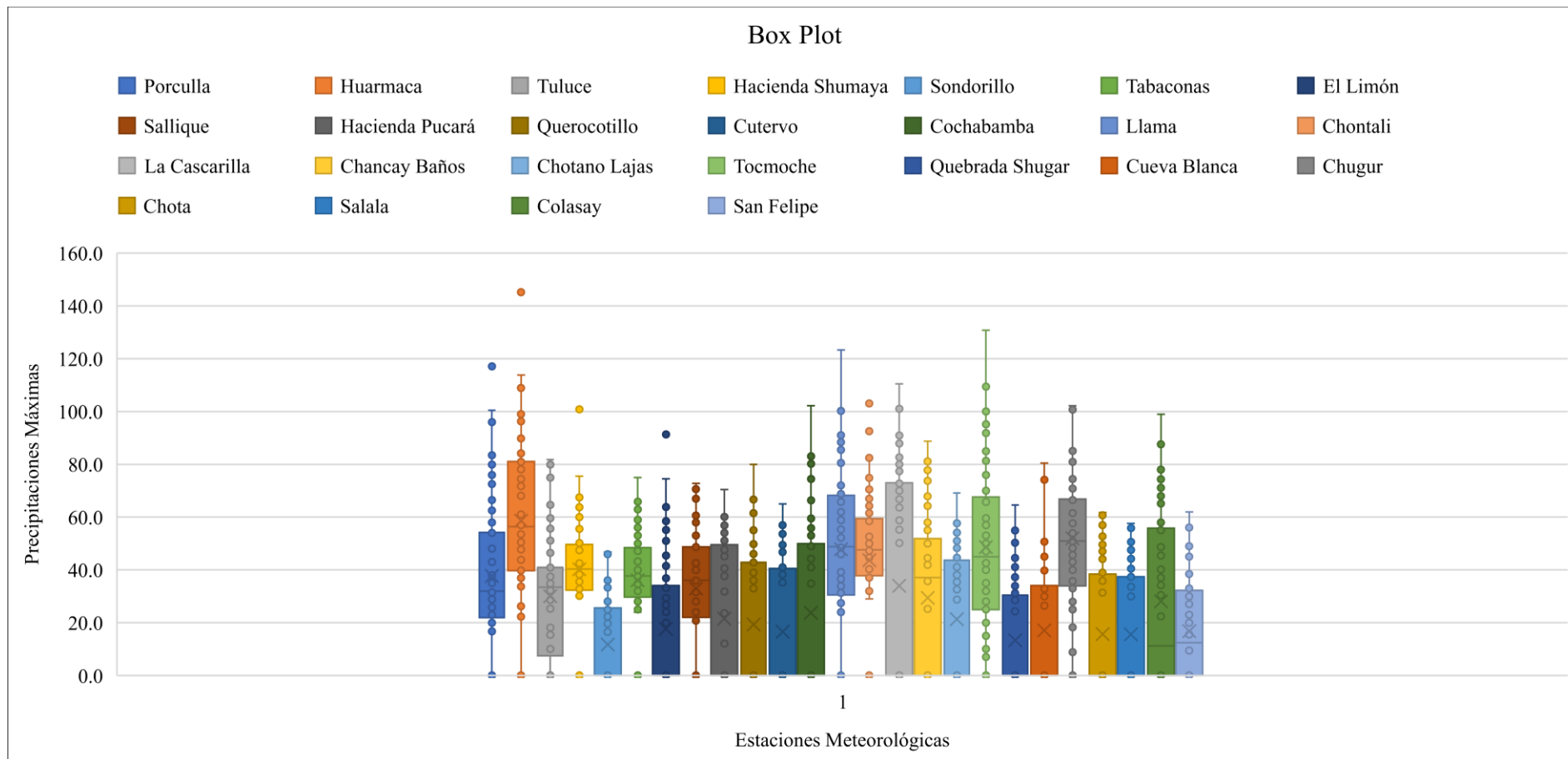
Tabla 17

Estadística descriptiva de PMDA por estación

N°	Estación	N° de años	Cumple	Años Faltantes	% Faltantes	% Completos	Promedio	Desv. Estándar	C.V.	Máximo	Año máx.	Mínimo	Año mín.	Q1	Mediana	Q3	IQR	Límite inferior	Límite superior	N° de Outliers
1	Porculla	56	Si	3	5.4	94.6	41.7	23.0	0.55	117.1	2008	16.7	2004	25.0	34.4	54.8	29.8	0.0	99.5	2
2	Huarmaca	55	Si	6	10.9	89.1	66.1	26.3	0.40	145.2	2010	22.3	1966	45.0	61.7	82.1	37.2	0.0	137.8	1
3	Tuluce	47	Si	3	6.4	93.6	39.6	18.4	0.47	81.8	2008	10.0	1964, 1965	28.7	37.3	49.3	20.6	0.0	80.2	1
4	Hacienda Shumaya	56	Si	3	5.4	94.6	44.5	13.5	0.30	100.8	2001	30.2	2013, 2020, 2022	36.5	40.5	50.0	13.6	16.1	70.3	2
5	Sondorillo	25	Si	5	20.0	80.0	29.0	7.9	0.27	46.8	2014	16.5	2009	24.2	27.4	35.5	11.3	7.3	52.5	0
6	Tabaconas	53	Si	7	13.2	86.8	42.2	11.2	0.27	75.0	1998	24.0	1974	35.0	39.0	50.0	15.0	12.5	72.5	1
7	El Limón	26	Si	4	15.4	84.6	42.2	17.5	0.41	91.3	2010	20.0	2024	27.2	41.6	52.2	25.0	0.0	89.7	1
8	Sallique	49	Si	12	24.5	75.5	41.6	13.8	0.33	72.8	1996	20.8	2013	32.3	40.0	51.2	18.9	3.9	79.6	0
9	Hacienda Pucará	28	Si	2	7.1	92.9	47.7	12.2	0.26	70.5	2000	12.1	1997	40.7	50.2	54.2	13.5	20.4	74.5	1
10	Querocotillo	24	Si	1	4.2	95.8	49.9	12.2	0.24	80.0	1998	33.0	1997	40.1	48.8	58.0	17.9	13.2	84.9	0
11	Cutervo	22	Si	6	27.3	72.7	46.8	8.1	0.17	65.0	2014	35.2	2011	40.1	47.9	51.0	10.9	23.8	67.3	0
12	Cochabamba	25	Si	5	20.0	80.0	58.9	17.9	0.30	102.2	2008	34.8	2017	45.0	53.0	74.4	29.4	0.9	118.5	0
13	Llama	51	Si	6	11.8	88.2	58.1	21.6	0.37	123.3	2002	24.0	2020	40.3	55.5	71.4	31.1	0.0	117.9	1
14	Chontalí	50	Si	11	22.0	78.0	54.2	15.4	0.28	103.1	2022	29.0	2002	44.7	51.5	62.8	18.1	17.5	90.0	2
15	La Cascarilla	28	Si	4	14.3	85.7	75.3	13.7	0.18	110.5	2001	50.2	2010	67.0	75.7	82.7	15.8	43.3	106.3	1
16	Chancay Baños	35	Si	2	5.7	94.3	52.1	14.2	0.27	88.8	2010	25.2	1990	42.8	50.5	57.2	14.4	21.2	78.8	2
17	Chotano Lajas	29	Si	1	3.4	96.6	45.6	9.2	0.20	69.1	1998	28.8	2011	39.5	44.6	51.0	11.5	22.3	68.3	1
18	Tocmoche	56	Si	5	8.9	91.1	53.8	29.5	0.55	130.8	2023	7.0	1995	32.8	49.0	68.5	35.8	0.0	122.1	1
19	Quebrada Shugar	23	Si	2	8.7	91.3	35.9	9.8	0.27	64.6	2022	24.3	2001	29.9	34.2	37.4	7.5	18.7	48.7	3
20	Cueva Blanca	26	Si	4	15.4	84.6	40.7	13.2	0.32	80.5	2023	26.4	2004	31.9	37.6	46.2	14.4	10.3	67.7	2
21	Chugur	60	Si	1	1.7	98.3	53.6	22.0	0.41	102.2	1966	8.8	1992	38.9	52.0	66.8	27.9	0.0	108.7	0
22	Chota	21	Si	3	14.3	85.7	46.3	10.0	0.22	61.8	2006, 2024	31.4	2011	38.2	44.1	54.2	16.0	14.2	78.2	0
23	Salala	23	Si	4	17.4	82.6	42.0	8.6	0.20	57.6	2013	30.0	2020	34.7	40.5	49.1	14.5	13.0	70.8	0
24	Colasay	31	Si	4	12.9	87.1	56.1	18.8	0.33	99.0	1973	22.4	1965	39.6	55.0	71.6	32.0	0.0	119.5	0
25	San Felipe	32	Si	2	6.3	93.8	32.5	11.5	0.35	62.0	1992	9.4	1974	26.3	32.0	38.6	12.4	7.7	57.2	1

Figura 22

Boxplot de PMDA por estación (comparación rápida de magnitudes y dispersión)



d) Exploración gráfica de las series PMDA: Series temporales de PMDA por estación (variabilidad interanual)

Las series temporales evidencian que los PMDA tienen fluctuación interanual en todas las estaciones, con ocurrencia de picos aislados que explican los máximos y outliers identificados en la estadística descriptiva. En conjunto, las figuras permiten verificar visualmente la alternancia de años con PMDA moderadas frente a años con incrementos pronunciados, consolidando la necesidad de tratar el conjunto como muestra regional donde la magnitud y dispersión no son uniformes entre estaciones.

Figura 23

Precipitaciones de la Estación 01

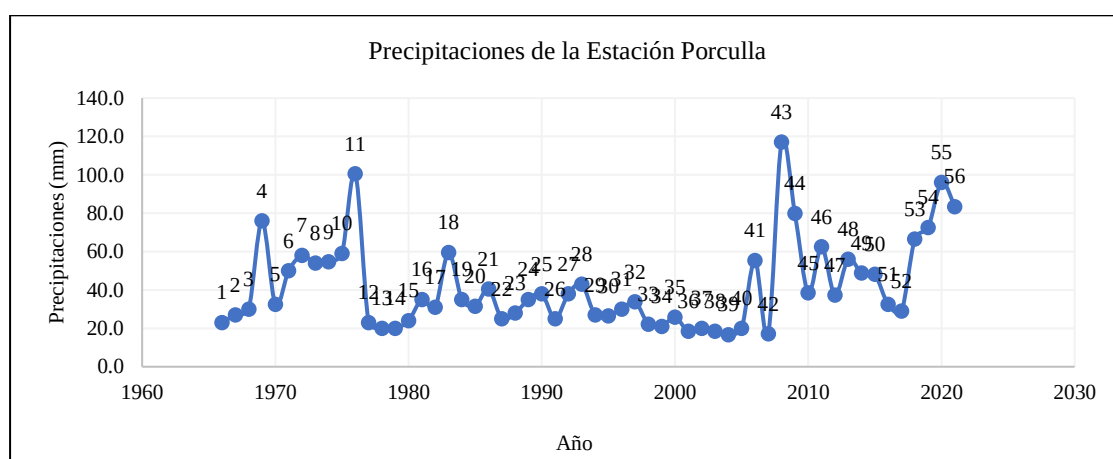


Figura 24

Precipitaciones de la Estación 02

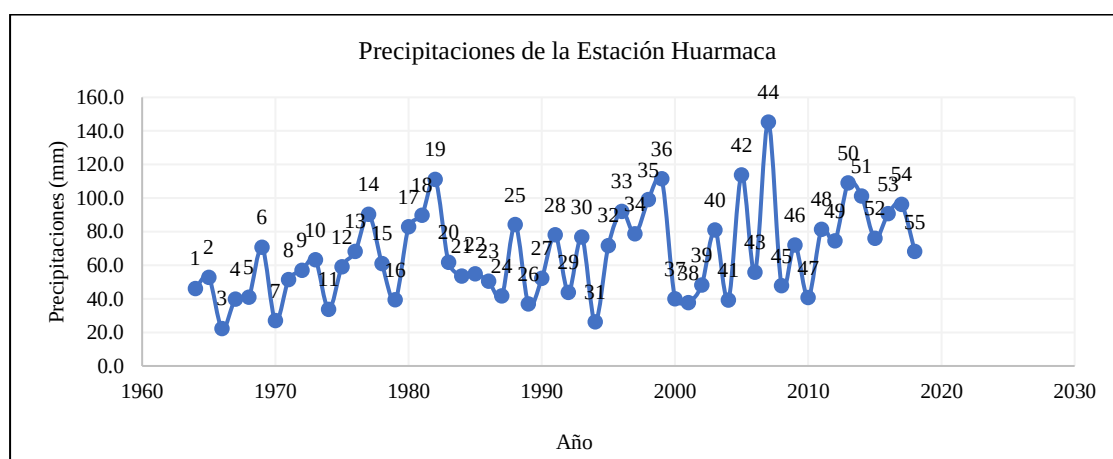


Figura 25

Precipitaciones de la Estación 03

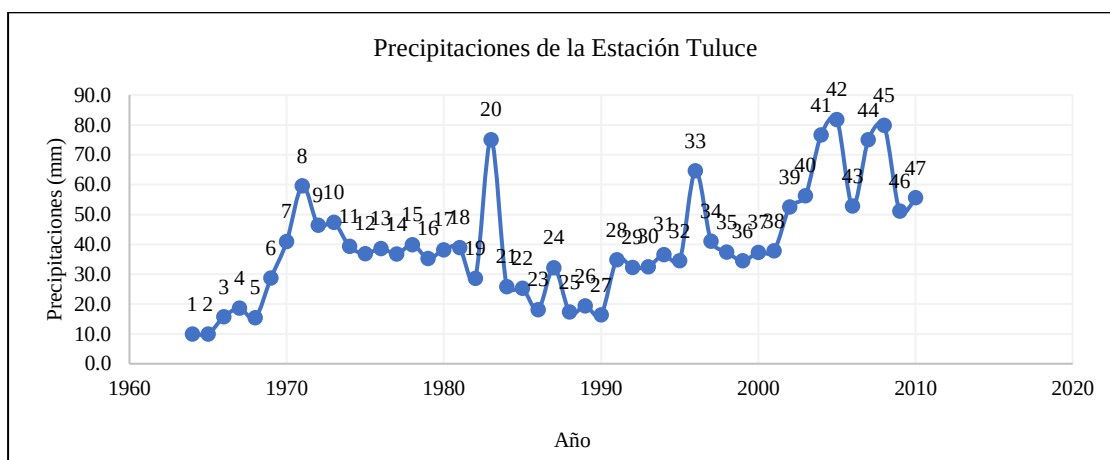


Figura 26

Precipitaciones de la Estación 04

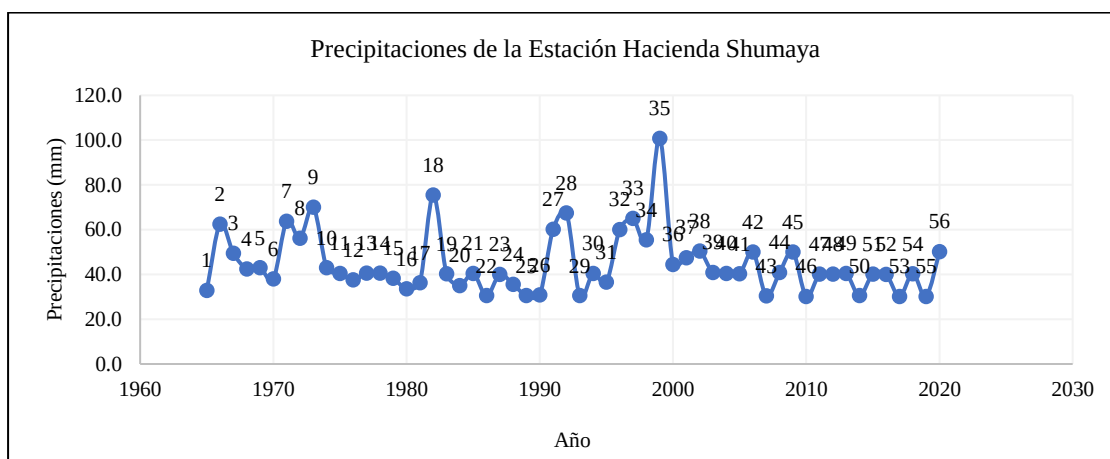


Figura 27

Precipitaciones de la Estación 05

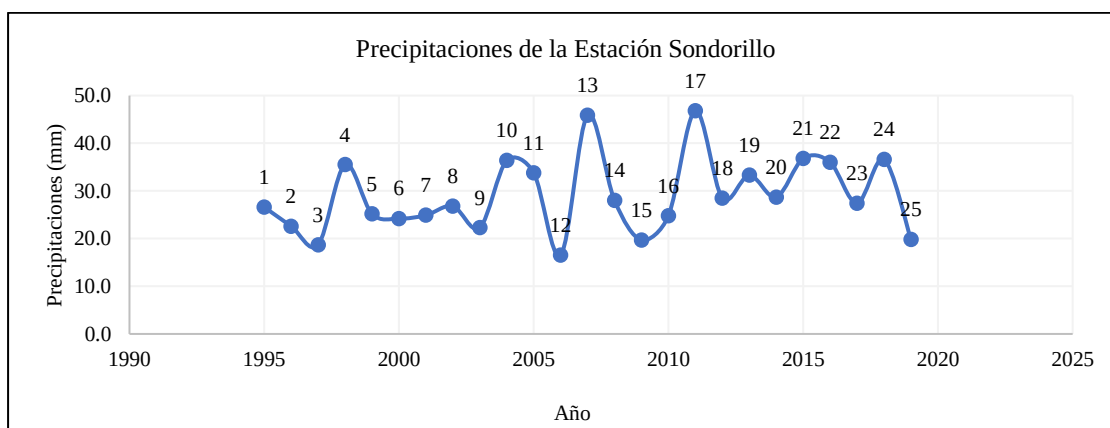


Figura 28

Precipitaciones de la Estación 06

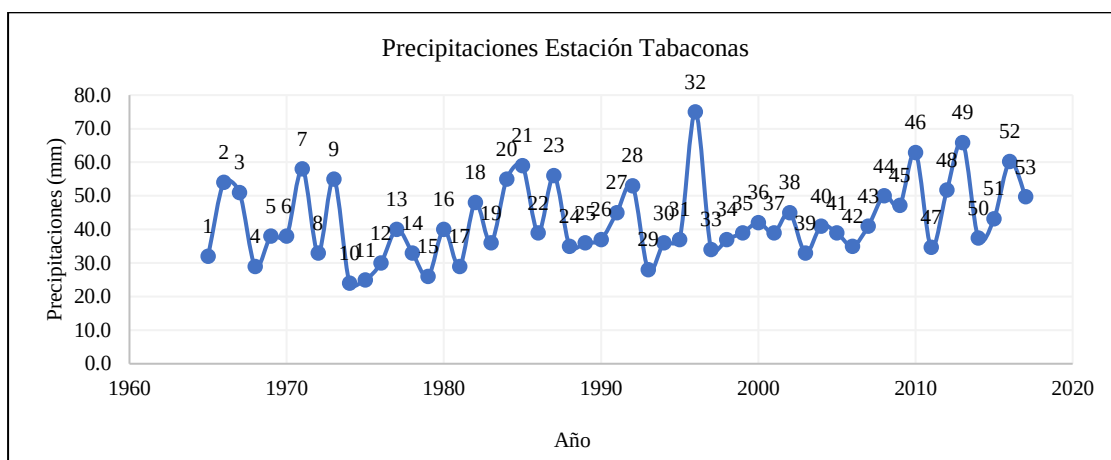


Figura 29

Precipitaciones de la Estación 07

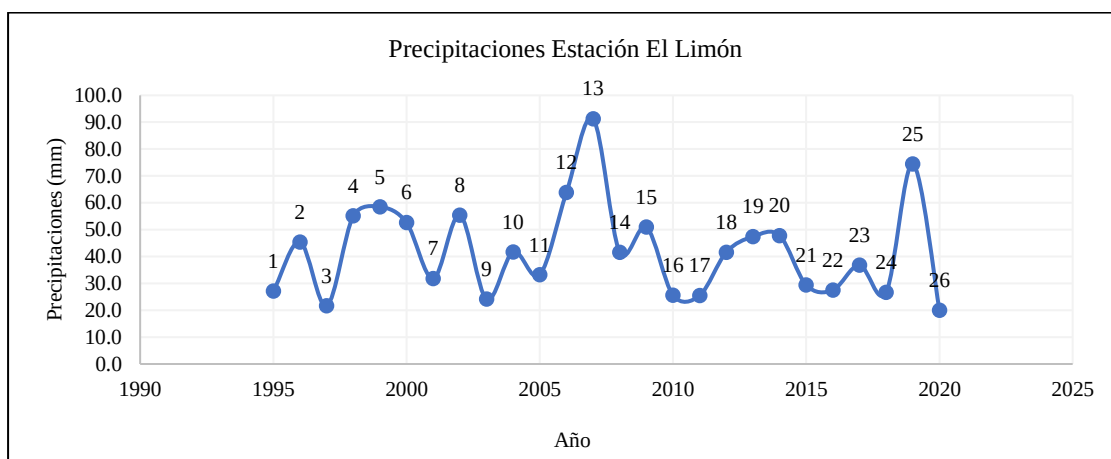


Figura 30

Precipitaciones de la Estación 08

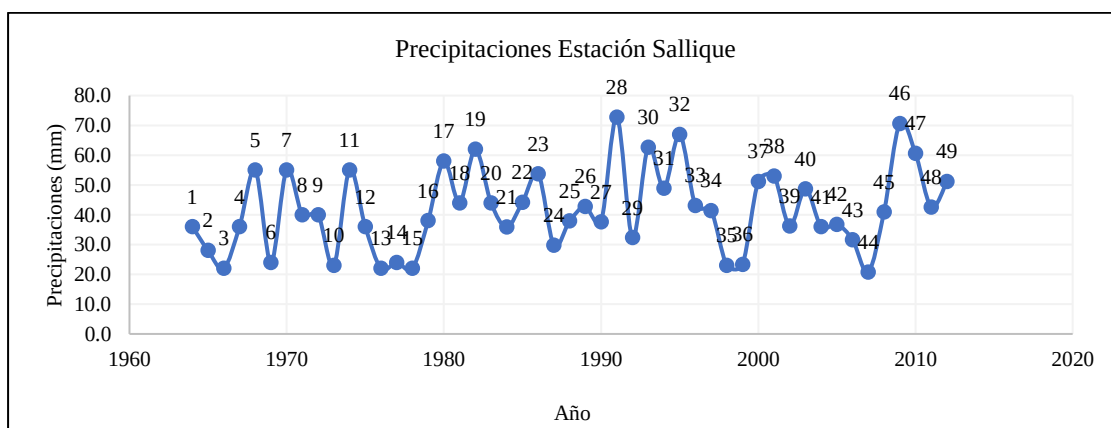


Figura 31

Precipitaciones de la Estación 09

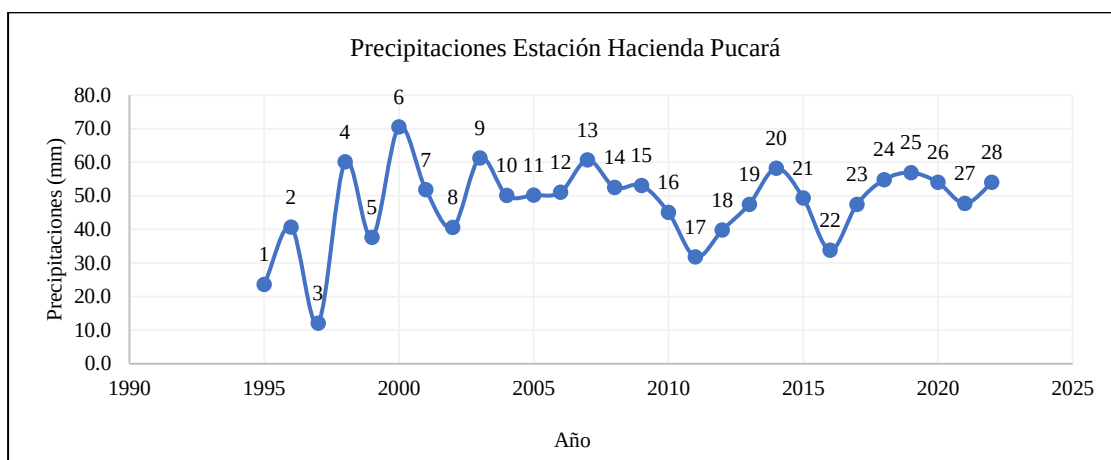


Figura 32

Precipitaciones de la Estación 10

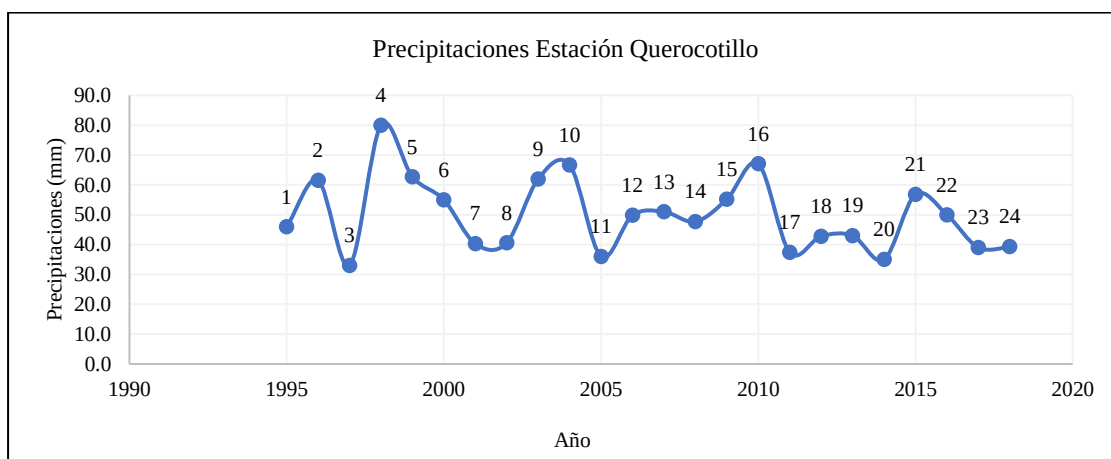


Figura 33

Precipitaciones de la Estación 11

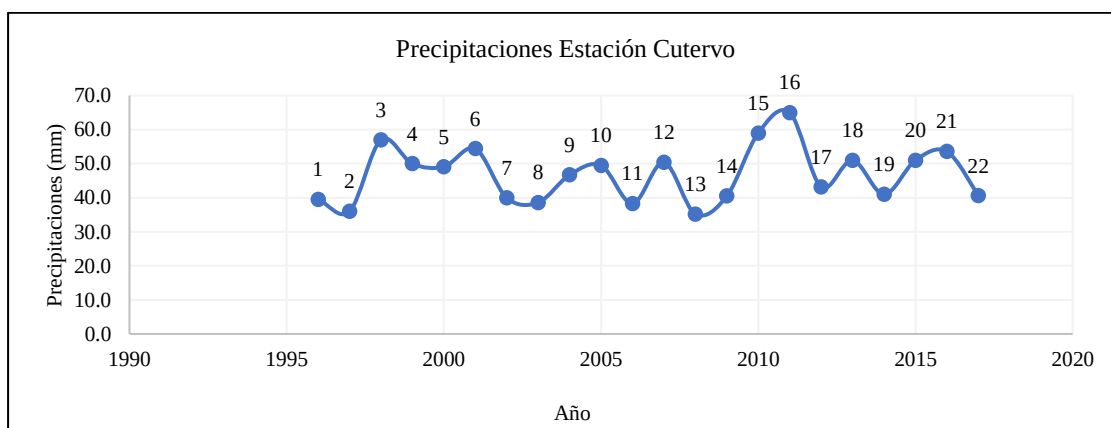


Figura 34

Precipitaciones de la Estación 12

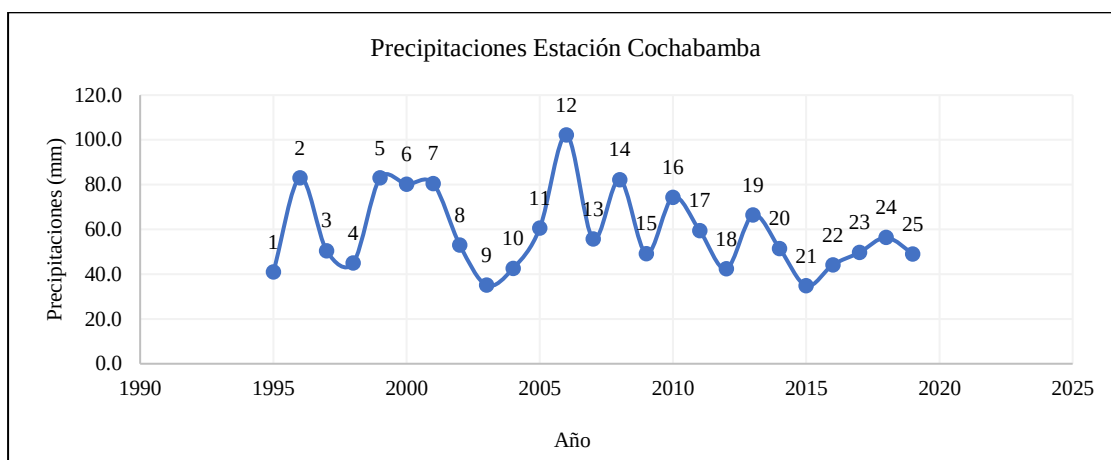


Figura 35

Precipitaciones de la Estación 13

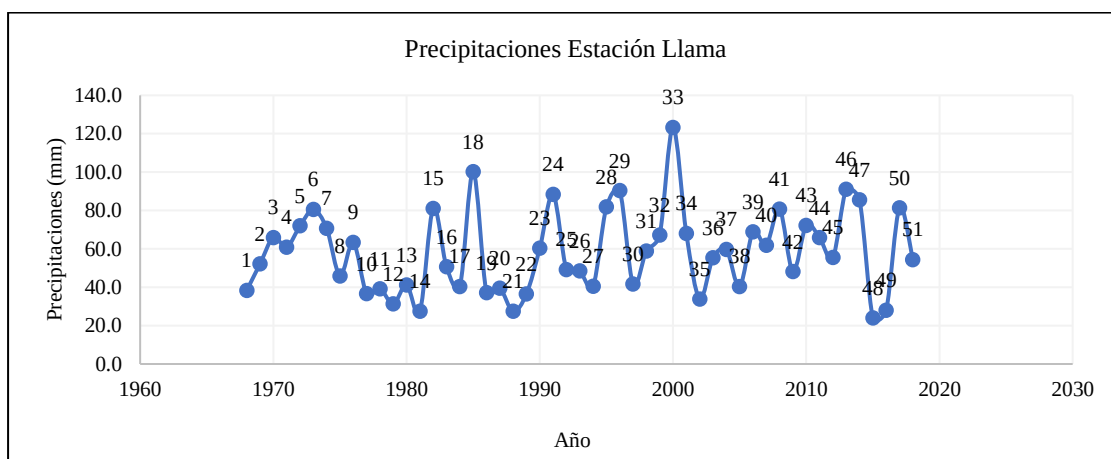


Figura 36

Precipitaciones de la Estación 14

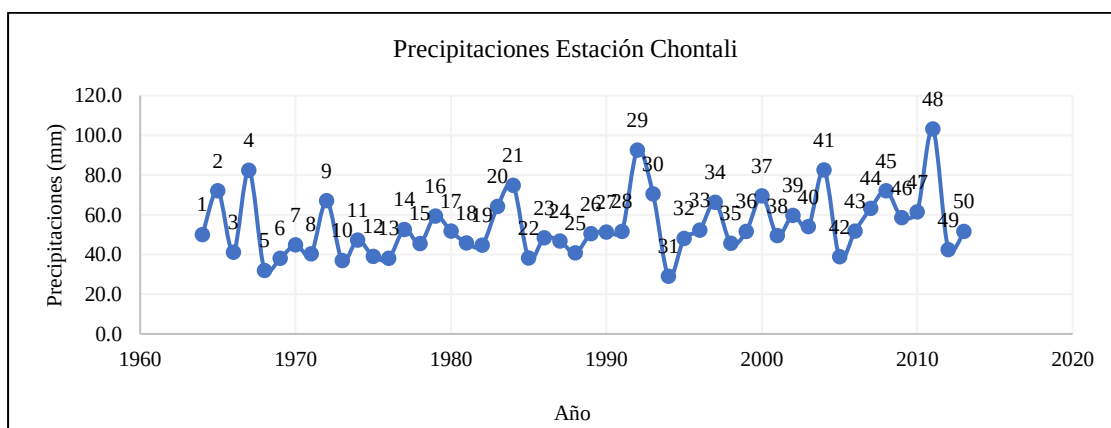


Figura 37

Precipitaciones de la Estación 15

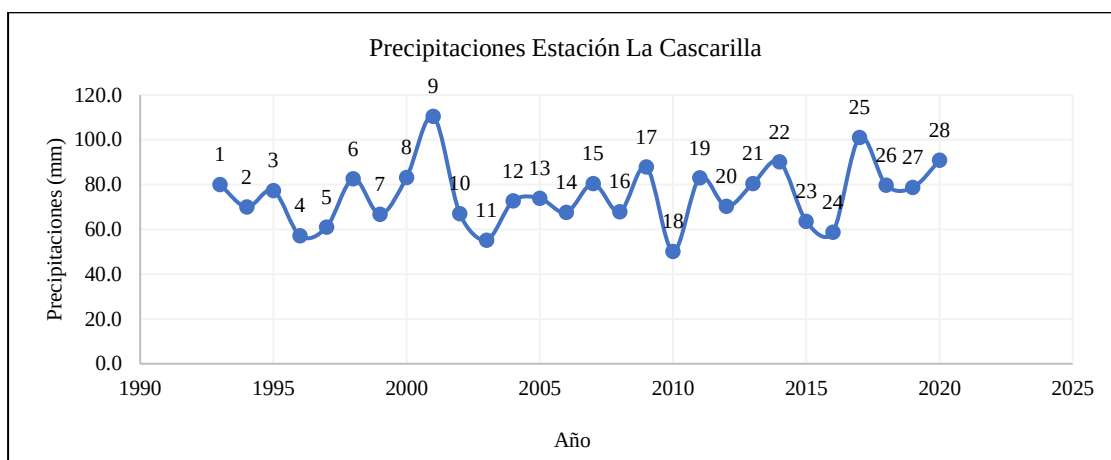


Figura 38

Precipitaciones de la Estación 16

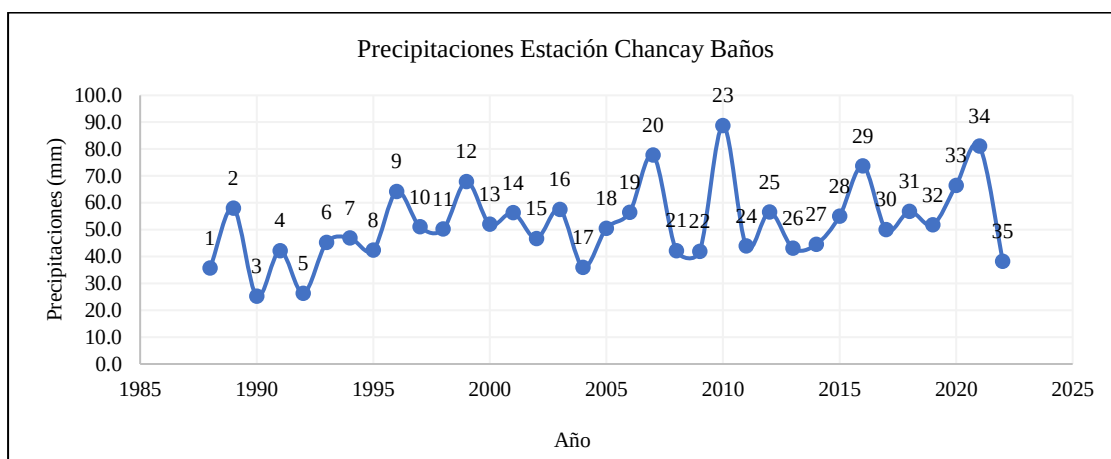


Figura 39

Precipitaciones de la Estación 17

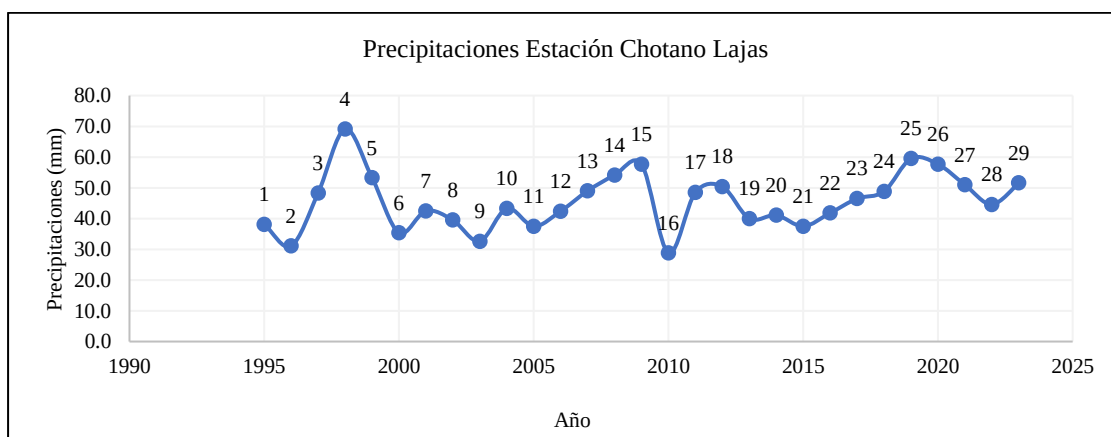


Figura 40

Precipitaciones de la Estación 18

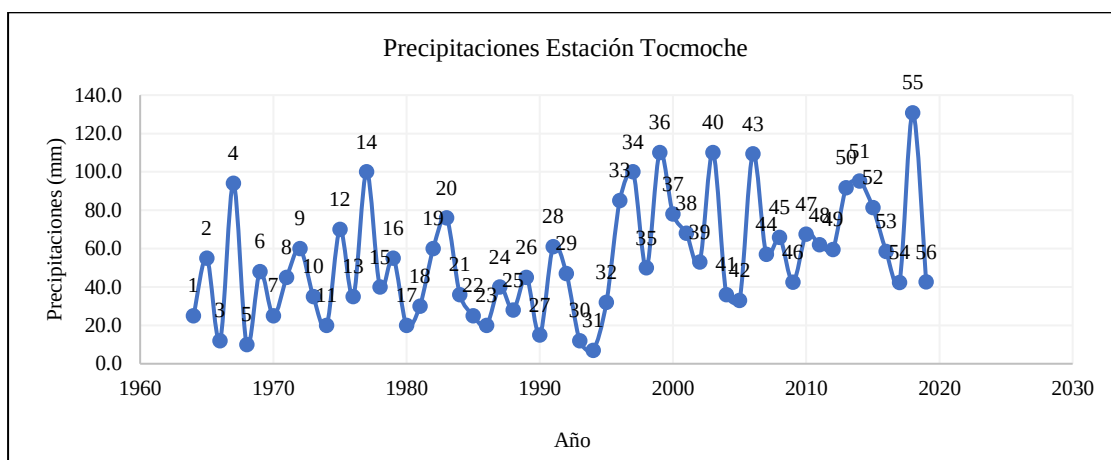


Figura 41

Precipitaciones de la Estación 19

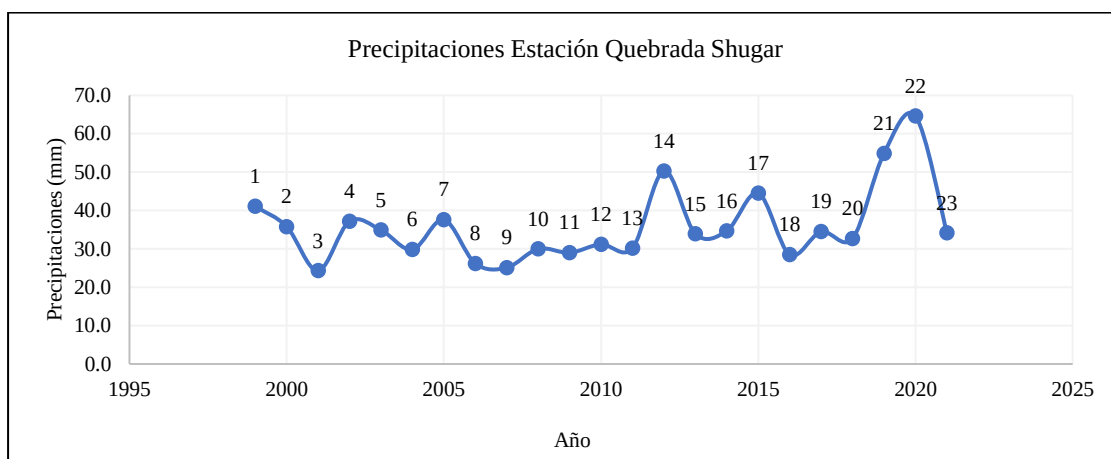


Figura 42

Precipitaciones de la Estación 20

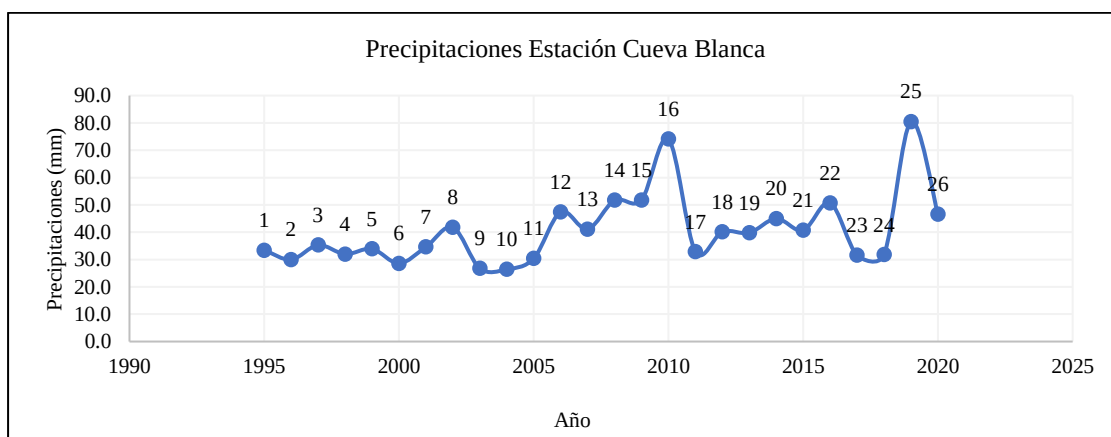


Figura 43

Precipitaciones de la Estación 21

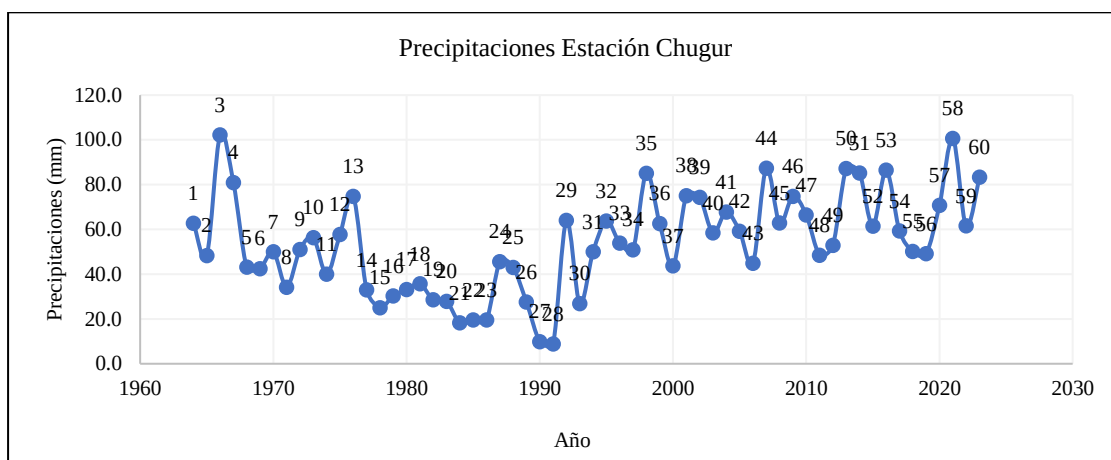


Figura 44

Precipitaciones de la Estación 22

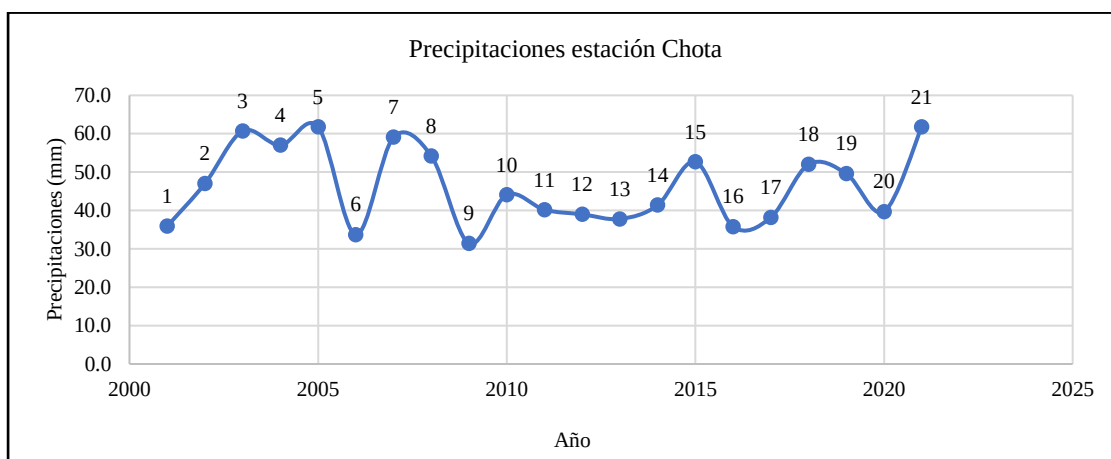


Figura 45

Precipitaciones de la Estación 23

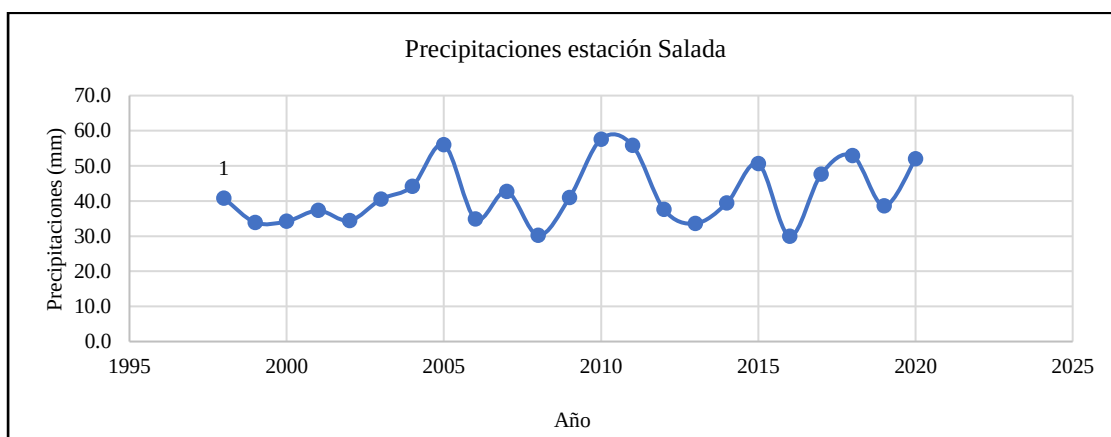


Figura 46

Precipitaciones de la Estación 24

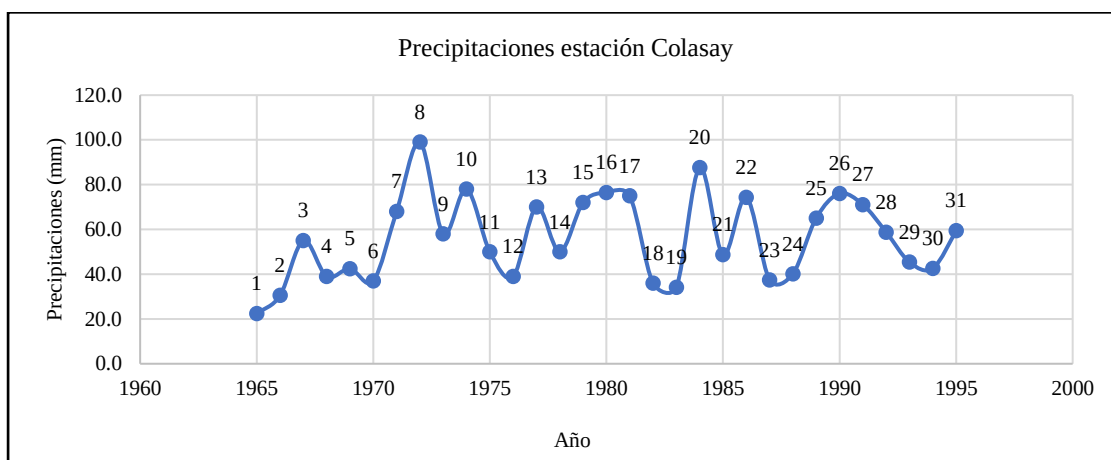
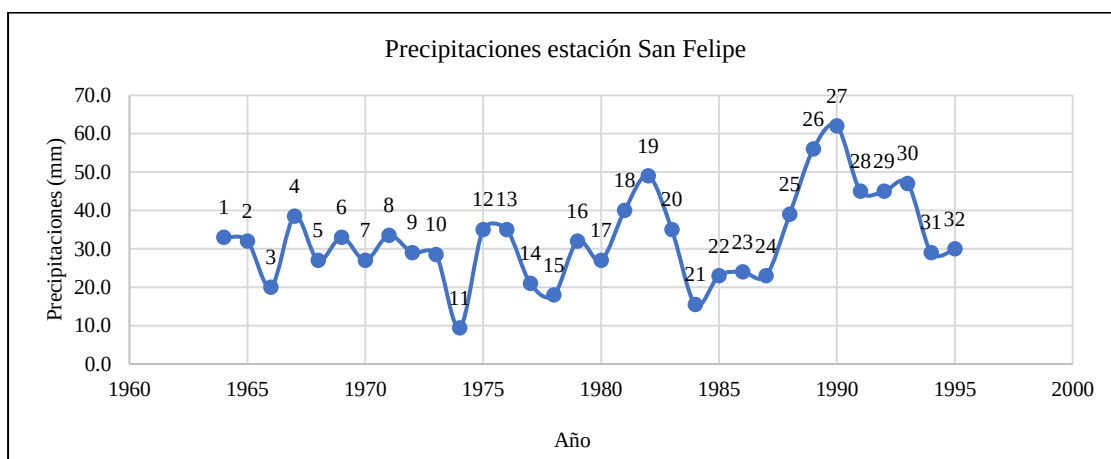


Figura 47

Precipitaciones de la Estación 25



e) Evaluación de tendencia y estacionariedad de las series PMDA

Para las 25 estaciones, se aplicaron tres verificaciones de tendencia/estacionariedad (Mann–Kendall, Spearman y tendencia en la media) comparando el estadístico de prueba con su valor crítico; en la mayoría de estaciones se aceptó la hipótesis de no tendencia, lo que se refleja en decisiones de “se acepta” y en interpretaciones de estacionariedad. Sin embargo, se identificaron estaciones con evidencia consistente de tendencia al rechazarse los contrastes ($|\text{estadístico}| > \text{valor}$

crítico) en los tres bloques, destacando Huarmaca (ID 2), Tuluze (ID 3), Chontalí (ID 14) y Tocmoche (ID 18), donde se reporta “hay tendencia / no hay estacionariedad” y, por ello, el resultado final fue “No” y se definió corrección de datos = Sí. Adicionalmente, se observan casos puntuales de resultados mixtos entre pruebas: Tabaconas (ID 6) rechazó Mann–Kendall y Spearman, pero aceptó la tendencia en la media y el resultado final se mantuvo como “Sí” sin corrección; Chancay Baños (ID 16) rechazó Mann–Kendall y tendencia en la media, pero aceptó Spearman, aun así, el resultado final se registró como “Sí” sin corrección; y San Felipe (ID 25) solo rechazó la prueba de tendencia en la media, manteniendo “Sí” como resultado final. En conjunto, la tabla indica que la base de PMDA fue mayoritariamente estacionaria, aunque se requirió corrección específica en cuatro estaciones para homogenizar el comportamiento temporal antes de continuar con el análisis posterior.

Tabla 18

Estaciones que necesitan corrección por presentar tendencia en la serie de datos

Estación	Resultado Final	Corrección de Datos
Huarmaca	No	Si
Tuluze	No	Si
Chontalí	No	Si
Tocmoche	No	Si

Tabla 19

Resultados del análisis de tendencia/estacionariedad en series PMDA (Mann–Kendall, Spearman y tendencia en la media)

ID	Estación	Test de Mann Kendall				Test de Spearman				Tendencia en la media				Decisión Final luego de los 3 test	
		Test Estadístico	Valor Crítico	Decisión	Interpretación	Test Estadístico	Valor Crítico	Decisión	Interpretación	Test Estadístico	Valor Crítico	Decisión	Interpretación	Resultado Final	Corrección de Datos
1	Porculla	-1.94	1.96	Se acepta	No hay tendencia	-1.83	2.00	Se acepta	Hay estacionariedad	-1.81	1.96	Se acepta	No hay tendencia	Si	No
2	Huarmaca	2.93	1.96	Se rechaza	Hay tendencia	3.17	2.01	Se rechaza	No hay estacionariedad	3.28	1.96	Se rechaza	Hay tendencia	No	Si
3	Tuluce	3.48	1.96	Se rechaza	Hay tendencia	3.95	2.01	Se rechaza	No hay estacionariedad	4.45	1.96	Se rechaza	Hay tendencia	No	Si
4	Hacienda Shumaya	-1.72	1.96	Se acepta	No hay tendencia	-1.42	2.00	Se acepta	Hay estacionariedad	-1.09	1.96	Se acepta	No hay tendencia	Si	No
5	Sondorillo	1.66	1.96	Se acepta	No hay tendencia	1.84	2.07	Se acepta	Hay estacionariedad	1.34	1.96	Se acepta	No hay tendencia	Si	No
6	Tabaconas	2.34	1.96	Se rechaza	Hay tendencia	2.18	2.01	Se rechaza	No hay estacionariedad	1.92	1.96	Se acepta	No hay tendencia	Si	No
7	El Limón	-0.75	1.96	Se acepta	No hay tendencia	-0.74	2.06	Se acepta	Hay estacionariedad	-0.41	1.96	Se acepta	No hay tendencia	Si	No
8	Sallique	1.43	1.96	Se acepta	No hay tendencia	1.57	2.01	Se acepta	Hay estacionariedad	1.63	1.96	Se acepta	No hay tendencia	Si	No
9	Hacienda Pucará	0.93	1.96	Se acepta	No hay tendencia	0.96	2.06	Se acepta	Hay estacionariedad	1.31	1.96	Se acepta	No hay tendencia	Si	No
10	Querocotillo	-1.17	1.96	Se acepta	No hay tendencia	-1.30	2.07	Se acepta	Hay estacionariedad	-1.50	1.96	Se acepta	No hay tendencia	Si	No
11	Cutervo	0.93	1.96	Se acepta	No hay tendencia	1.07	2.09	Se acepta	Hay estacionariedad	0.64	1.96	Se acepta	No hay tendencia	Si	No
12	Cochabamba	-1.12	1.96	Se acepta	No hay tendencia	-1.03	2.07	Se acepta	Hay estacionariedad	-1.18	1.96	Se acepta	No hay tendencia	Si	No
13	Llama	1.15	1.96	Se acepta	No hay tendencia	1.12	2.01	Se acepta	Hay estacionariedad	1.04	1.96	Se acepta	No hay tendencia	Si	No
14	Chontalí	2.57	1.96	Se rechaza	Hay tendencia	2.45	2.01	Se rechaza	No hay estacionariedad	2.00	1.96	Se rechaza	Hay tendencia	No	Si
15	La Cascarilla	1.22	1.96	Se acepta	No hay tendencia	1.20	2.06	Se acepta	Hay estacionariedad	0.96	1.96	Se acepta	No hay tendencia	Si	No
16	Chancay Baños	2.06	1.96	Se rechaza	Hay tendencia	1.92	2.03	Se acepta	Hay estacionariedad	2.27	1.96	Se rechaza	Hay tendencia	SI	No

ID	Test de Mann Kendall					Test de Spearman				Tendencia en la media				Decisión Final luego de los 3 test	
	Estación	Test Estadístico	Valor Crítico	Decisión	Interpretación	Test Estadístico	Valor Crítico	Decisión	Interpretación	Test Estadístico	Valor Crítico	Decisión	Interpretación	Resultado Final	Corrección de Datos
17	Chotano Lajas	1.82	1.96	Se acepta	No hay tendencia	1.77	2.05	Se acepta	Hay estacionariedad	1.24	1.96	Se acepta	No hay tendencia	Si	No
18	Tocmoche	2.94	1.96	Se rechaza	Hay tendencia	3.26	2.00	Se rechaza	No hay estacionariedad	3.16	1.96	Se rechaza	Hay tendencia	No	Si
19	Quebrada Shugar	1.16	1.96	Se acepta	No hay tendencia	0.98	2.08	Se acepta	Hay estacionariedad	1.88	1.96	Se acepta	No hay tendencia	Si	No
20	Cueva Blanca	-0.35	1.96	Se acepta	No hay tendencia	-0.60	2.06	Se acepta	Hay estacionariedad	-0.09	1.96	Se acepta	No hay tendencia	Si	No
21	Chugur	-0.86	1.96	Se acepta	No hay tendencia	-0.68	2.00	Se acepta	Hay estacionariedad	-0.69	1.96	Se acepta	No hay tendencia	Si	No
22	Chota	-0.36	1.96	Se acepta	No hay tendencia	-0.27	2.09	Se acepta	Hay estacionariedad	-0.39	1.96	Se acepta	No hay tendencia	Si	No
23	Salala	1.27	1.96	Se acepta	No hay tendencia	1.28	2.08	Se acepta	Hay estacionariedad	1.43	1.96	Se acepta	No hay tendencia	Si	No
24	Colasay	1.12	1.96	Se acepta	No hay tendencia	1.23	2.05	Se acepta	Hay estacionariedad	0.97	1.96	Se acepta	No hay tendencia	Si	No
25	San Felipe	1.65	1.96	Se acepta	No hay tendencia	1.74	2.04	Se acepta	Hay estacionariedad	2.04	1.96	Se rechaza	Hay tendencia	Si	No

f) Evaluación de homogeneidad (cambios en media/varianza) de las series PMDA

La evaluación de homogeneidad de las series de PMDA en 25 estaciones se realizó mediante tres contrastes (Mann–Whitney, prueba t y prueba F), donde se comparó el estadístico con su valor crítico para decidir aceptación o rechazo de homogeneidad. La decisión final indica que 22 estaciones fueron clasificadas como homogéneas (Sí) y 3 estaciones como no homogéneas (No), requiriendo estas últimas corrección: Porculla (ID 1), Cueva Blanca (ID 20) y Chugur (ID 21), que presentaron rechazo en los tres tests. En el resto de estaciones, aunque varias registraron rechazos en uno o dos tests, la decisión final se mantuvo como homogénea, evidenciando resultados mixtos según el tipo de prueba; por ejemplo, casos donde se rechazó Mann–Whitney y prueba t, pero se aceptó prueba F (Huarmaca, Tuluze, Tabaconas, Chontalí, Chancay Baños, Tocmoche y San Felipe), o donde se aceptaron Mann–Whitney y t, pero se rechazó F (Hacienda Pucará y Cochabamba). La mayoría de series mantuvo consistencia interna bajo al menos parte de los contrastes aplicados, mientras que tres estaciones mostraron señales sistemáticas de no homogeneidad y, por tanto, fueron tratadas mediante corrección para mantener coherencia en el análisis posterior.

Tabla 20

Estaciones que necesitan corrección por presentar saltos en la media

Estación	Es homogénea	Corrección
Porculla	No	Si
Cueva Blanca	No	Si
Chugur	No	Si

Tabla 21

Número de estaciones que cumplen y no cumplen las pruebas

	Mann-Whitney	Prueba t	Prueba F
Cumplen	12	11	18
No cumplen	13	14	7

Tabla 22

Resultados del análisis de homogeneidad de las series PMDA (Mann–Whitney, prueba t y prueba F)

Id	Estación	Mann-Whitney				Prueba t				Prueba F				Decisión Final luego de los 3 test	
		Test Estadístico	Valor Crítico	Decisión	Interpretación	Test Estadístico	Valor Crítico	Decisión	Interpretación	Test Estadístico	Valor Crítico	Decisión	Interpretación	Es Homogénea	Corrección
1	Porculla	-3.75	1.96	Se rechaza	No Existe Homogeneidad	4.4	2.0	Se rechaza	No Existe Homogeneidad	2.2	2.0	Se rechaza	No Existe Homogeneidad	No	Si
2	Huarmaca	-2.72	1.96	Se rechaza	No Existe Homogeneidad	3.2	2.0	Se rechaza	No Existe Homogeneidad	1.8	1.9	Se acepta	Existe Homogeneidad	Si	No
3	Tuluce	-4.07	1.96	Se rechaza	No Existe Homogeneidad	5.1	2.0	Se rechaza	No Existe Homogeneidad	1.4	2.0	Se acepta	Existe Homogeneidad	Si	No
4	Hacienda Shumaya	-1.82	1.96	Se acepta	Existe Homogeneidad	2.1	2.0	Se rechaza	No Existe Homogeneidad	5.2	2.1	Se rechaza	No Existe Homogeneidad	Si	No
5	Sondorillo	-1.90	1.96	Se acepta	Existe Homogeneidad	1.8	2.1	Se acepta	Existe Homogeneidad	1.9	2.8	Se acepta	Existe Homogeneidad	Si	No
6	Tabaconas	-2.25	1.96	Se rechaza	No Existe Homogeneidad	2.2	2.0	Se rechaza	No Existe Homogeneidad	1.0	2.1	Se acepta	Existe Homogeneidad	Si	No
7	El Limón	-1.53	1.96	Se acepta	Existe Homogeneidad	1.4	2.1	Se acepta	Existe Homogeneidad	1.3	2.9	Se acepta	Existe Homogeneidad	Si	No
8	Sallique	-2.43	1.96	Se rechaza	No Existe Homogeneidad	2.6	2.0	Se rechaza	No Existe Homogeneidad	1.2	2.2	Se acepta	Existe Homogeneidad	Si	No
9	Hacienda Pucará	-0.82	1.96	Se acepta	Existe Homogeneidad	1.4	2.1	Se acepta	Existe Homogeneidad	6.6	2.6	Se rechaza	No Existe Homogeneidad	Si	No
10	Querocotillo	-1.96	1.96	Se acepta	Existe Homogeneidad	2.1	2.1	Se rechaza	No Existe Homogeneidad	3.2	3.5	Se acepta	Existe Homogeneidad	Si	No
11	Cutervo	-1.98	1.96	Se rechaza	No Existe Homogeneidad	1.7	2.1	Se acepta	Existe Homogeneidad	1.5	2.8	Se acepta	Existe Homogeneidad	Si	No
12	Cochabamba	-1.75	1.96	Se acepta	Existe Homogeneidad	1.9	2.1	Se acepta	Existe Homogeneidad	4.1	3.5	Se rechaza	No Existe Homogeneidad	Si	No
13	Llama	-1.95	1.96	Se acepta	Existe Homogeneidad	1.9	2.0	Se acepta	Existe Homogeneidad	1.4	2.0	Se acepta	Existe Homogeneidad	Si	No
14	Chontalí	-2.48	1.96	Se rechaza	No Existe Homogeneidad	2.4	2.0	Se rechaza	No Existe Homogeneidad	2.0	2.0	Se acepta	Existe Homogeneidad	Si	No
15	La Cascarilla	-1.03	1.96	Se acepta	Existe Homogeneidad	1.1	2.1	Se acepta	Existe Homogeneidad	1.3	2.6	Se acepta	Existe Homogeneidad	Si	No
16	Chancay Baños	-2.06	1.96	Se rechaza	No Existe Homogeneidad	2.4	2.0	Se rechaza	No Existe Homogeneidad	1.1	3.1	Se acepta	Existe Homogeneidad	Si	No

Id	Estación	Mann-Whitney				Prueba t				Prueba F				Decisión Final luego de los 3 test	
		Test Estadístico	Valor Crítico	Decisión	Interpretación	Test Estadístico	Valor Crítico	Decisión	Interpretación	Test Estadístico	Valor Crítico	Decisión	Interpretación	Es Homogénea	Corrección
17	Chotano Lajas	-2.27	1.96	Se rechaza	No Existe Homogeneidad	2.0	2.1	Se acepta	Existe Homogeneidad	3.0	3.9	Se acepta	Existe Homogeneidad	Si	No
18	Tocmoche	-4.09	1.96	Se rechaza	No Existe Homogeneidad	4.7	2.0	Se rechaza	No Existe Homogeneidad	1.3	1.9	Se acepta	Existe Homogeneidad	Si	No
19	Quebrada Shugar	-1.86	1.96	Se acepta	Existe Homogeneidad	2.6	2.1	Se rechaza	No Existe Homogeneidad	5.1	2.8	Se rechaza	No Existe Homogeneidad	Si	No
20	Cueva Blanca	-3.19	1.96	Se rechaza	No Existe Homogeneidad	3.4	2.1	Se rechaza	No Existe Homogeneidad	10.1	2.9	Se rechaza	No Existe Homogeneidad	No	Si
21	Chugur	-4.65	1.96	Se rechaza	No Existe Homogeneidad	5.3	2.0	Se rechaza	No Existe Homogeneidad	1.9	1.9	Se rechaza	No Existe Homogeneidad	No	Si
22	Chota	-1.48	1.96	Se acepta	Existe Homogeneidad	1.8	2.1	Se acepta	Existe Homogeneidad	1.8	2.9	Se acepta	Existe Homogeneidad	Si	No
23	Salala	-1.23	1.96	Se acepta	Existe Homogeneidad	1.7	2.1	Se acepta	Existe Homogeneidad	2.0	2.9	Se acepta	Existe Homogeneidad	Si	No
24	Colasay	-1.56	1.96	Se acepta	Existe Homogeneidad	1.2	2.0	Se acepta	Existe Homogeneidad	2.3	2.5	Se acepta	Existe Homogeneidad	Si	No
25	San Felipe	-2.98	1.96	Se rechaza	No Existe Homogeneidad	4.0	2.0	Se rechaza	No Existe Homogeneidad	1.8	2.4	Se acepta	Existe Homogeneidad	Si	No

g) Selección de estaciones: Series PMDA corregidas para el análisis regional

Luego de aplicar las correcciones derivadas de los análisis de tendencia y homogeneidad, se consolidó una base final con 25 estaciones marcadas para uso (“Sí”), manteniendo tamaños muestrales heterogéneos entre 21 y 60 años (n), lo que confirma que la depuración se enfocó en ajustar series y no en excluir estaciones. Los promedios de PMDA se ubicaron entre 29.0 mm (Sondorillo) y 75.3 mm (La Cascarilla). En conjunto, los resultados reflejan una base de datos regional corregida y utilizable, con diferencias en nivel central y variabilidad entre estaciones, características que resultan relevantes para el análisis posterior de regionalización.

Tabla 23

Selección de estaciones con datos corregidos

N°	Uso	Estación	n	Promedio	Desv.	C.V.
1	Si	Porculla	56	35.0	17.4	0.50
2	Si	Huarmaca	55	66.1	23.9	0.36
3	Si	Tuluze	47	39.6	15.2	0.38
4	Si	Hacienda Shumaya	56	44.5	13.5	0.30
5	Si	Sondorillo	25	29.0	7.9	0.27
6	Si	Tabaconas	53	42.2	11.2	0.27
7	Si	El Limón	26	42.2	17.5	0.41
8	Si	Sallique	49	41.6	13.8	0.33
9	Si	Hacienda Pucará	28	47.7	12.2	0.26
10	Si	Querocotillo	24	49.9	12.2	0.24
11	Si	Cutervo	22	46.8	8.1	0.17
12	Si	Cochabamba	25	58.9	17.9	0.30
13	Si	Llama	51	58.1	21.6	0.37
14	Si	Chontalí	50	54.2	14.8	0.27
15	Si	La Cascarilla	28	75.3	13.7	0.18
16	Si	Chancay Baños	35	52.1	14.2	0.27
17	Si	Chotano Lajas	29	45.6	9.2	0.20
18	Si	Tocmoche	56	53.8	27.0	0.50
19	Si	Quebrada Shugar	23	35.9	9.8	0.27
20	Si	Cueva Blanca	26	34.5	7.5	0.22
21	Si	Chugur	60	41.6	20.7	0.50
22	Si	Chota	21	46.3	10.0	0.22
23	Si	Salala	23	42.0	8.6	0.20
24	Si	Colasay	31	56.1	18.8	0.33
25	Si	San Felipe	32	32.5	11.5	0.35

4.1.2. Análisis regional de frecuencia y enfoque de Región de Influencia (ROI)

4.1.2.1. Análisis regional de frecuencia.

a) Agrupación preliminar (Ward y K-means)

La Tabla 24 muestra la conformación de 4 regiones preliminares a partir de un procedimiento combinado: primero, un agrupamiento jerárquico (Ward) para identificar estructuras naturales en los datos, y luego K-means para consolidar la asignación final. La agrupación preliminar incorpora las 25 estaciones, pero presenta tamaños desbalanceados, destacando una región con muy pocas estaciones (R-2), lo cual anticipa la necesidad de ajustes posteriores para robustecer la verificación de homogeneidad regional.

Tabla 24

Grupos iniciales se obtuvieron a partir de los métodos de agrupación de Ward y K-means

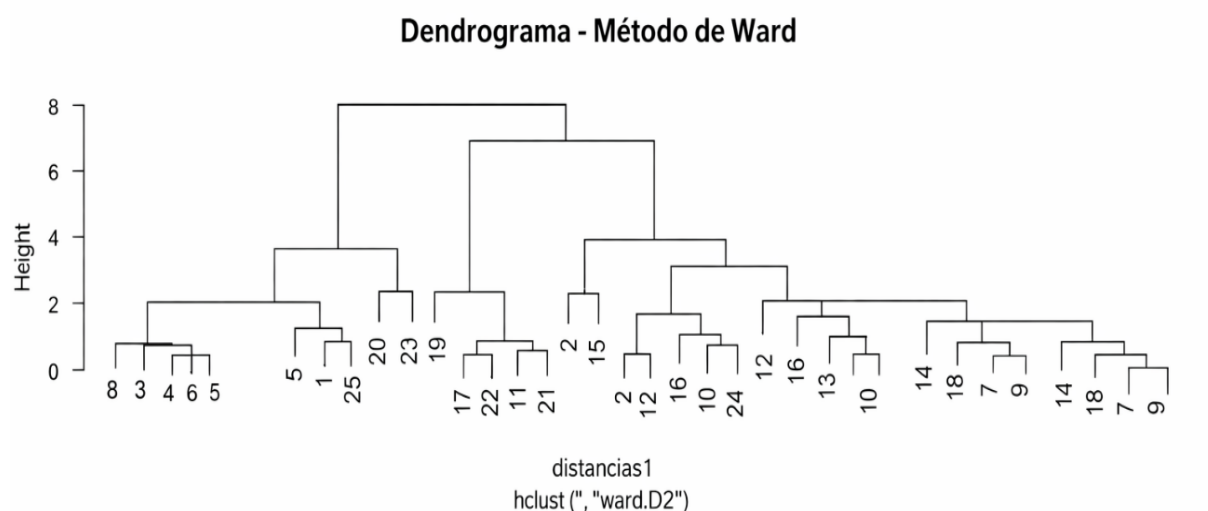
Nº	Región 1	Región 2	Región 3	Región 4
1	Porculla	Huarmaca	El Limón	Cutervo
2	Tuluce	La Cascarilla	Hacienda Pucará	Chotano Lajas
3	Hacienda Shumaya		Querocotillo	Quebrada Shugar
4	Sondorillo		Cochabamba	Chugur
5	Tabaconas		Llama	Chota
6	Sallique		Chontalí	
7	Cueva Blanca		Chancay Baños	
8	Salala		Tocmoche	
9	San Felipe		Colasay	

El dendrograma (Figura 48) sintetiza la jerarquía de similitud entre estaciones: las ramas que se fusionan a menor altura representan estaciones más parecidas, mientras que las uniones a mayor altura indican diferencias más marcadas. En consecuencia, los puntos de corte seleccionados permiten sustentar la decisión de trabajar con cuatro conglomerados, ya que se distinguen subconjuntos con separación relativa entre ramas,

lo cual respalda técnicamente la delimitación inicial de regiones para el análisis regional de frecuencias.

Figura 48

Dendrograma (agrupación jerárquica) de los grupos formados tentativamente



Nota. Dendrograma de agrupación jerárquica (método de Ward, ward.D2) aplicado a las 25 estaciones; la altura de enlace representa el incremento de disimilitud al fusionar grupos y el punto de corte seleccionado sustenta la conformación preliminar de cuatro conglomerados (regiones) para el análisis regional posterior.

La Figura 49 y Figura 50 muestran la distribución espacial de los grupos preliminares y permiten verificar que la regionalización inicial no solo responde a similitud estadística, sino que también mantiene coherencia geográfica. No obstante, la presencia de estaciones agrupadas con separaciones amplias y la heterogeneidad altitudinal dentro de algunas regiones sugieren que la partición inicial debe evaluarse con pruebas formales de discordancia y heterogeneidad.

Figura 49

Agrupación preliminar con el método Ward

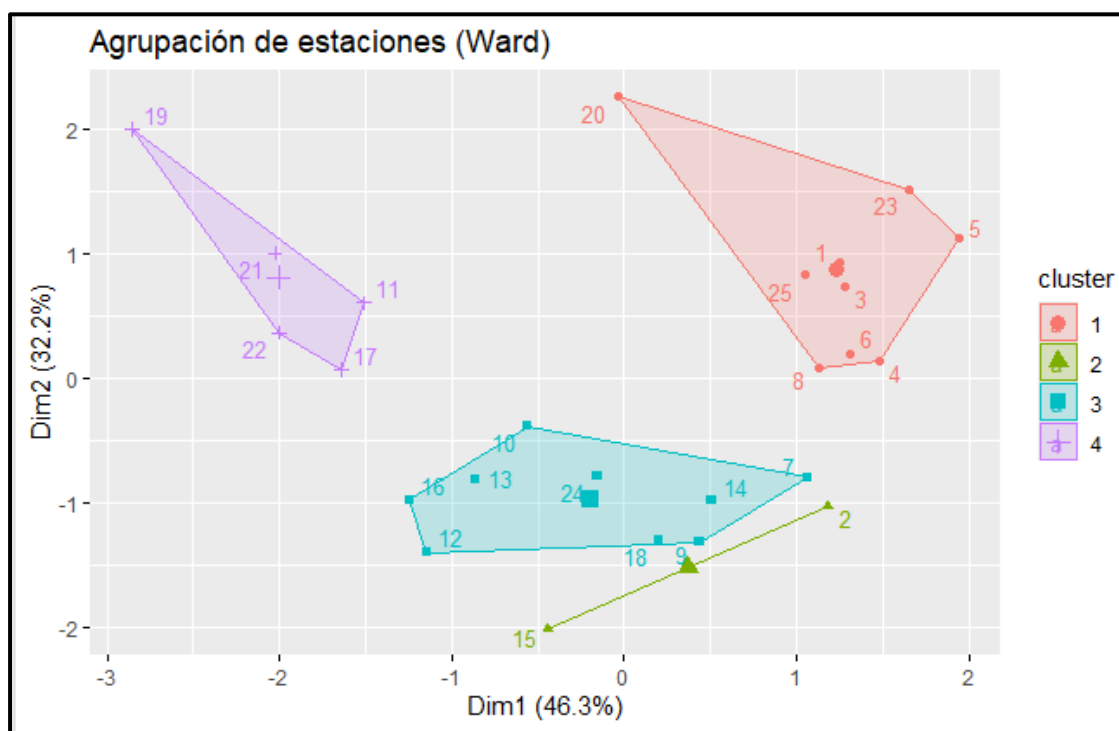
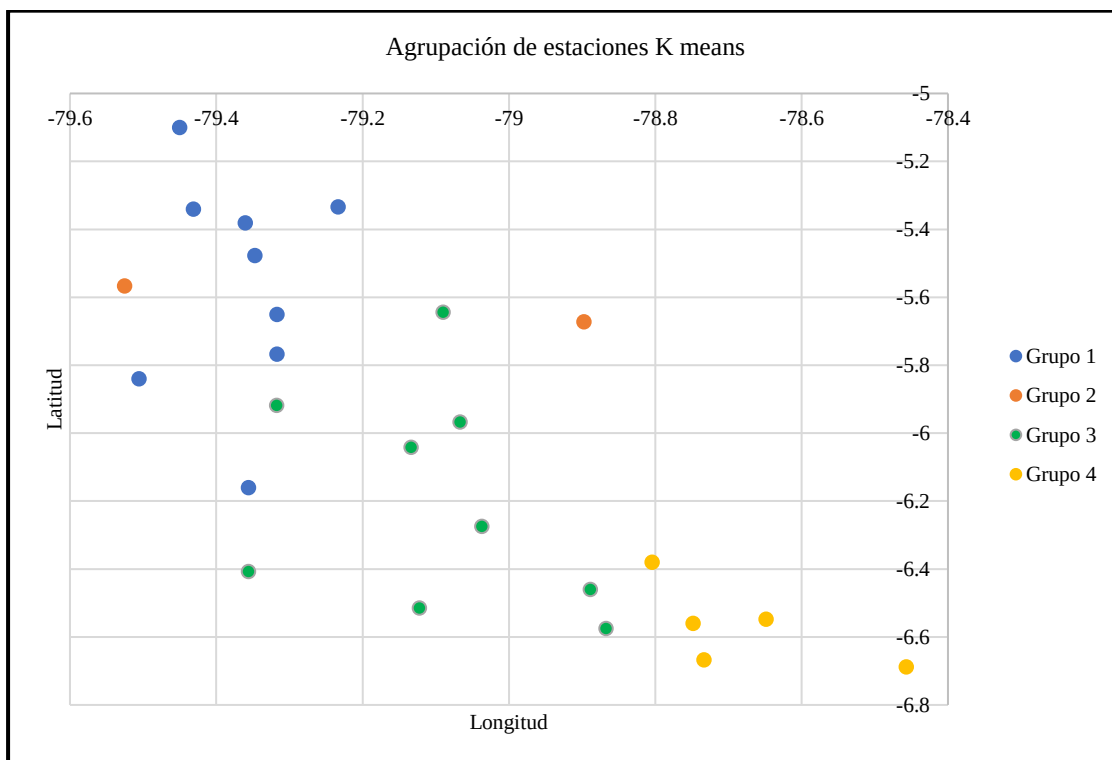


Figura 50

Agrupación preliminar de estaciones por el método K-means



Los valores máximos de discordancia por región preliminar son menores que 3, por lo que se cumple el criterio $D < 3$ en todas las regiones. Esto indica que, en esta etapa, no se identifican estaciones extremadamente discordantes dentro de su grupo preliminar y que el problema principal no es la presencia de “outliers regionales” por discordancia.

Tabla 25

Máximos valores de discordancia en las regiones preliminares formadas

R-1	R-2	R-3	R-4
2.04	0.41	2.43	1.31

Los resultados de H1 muestran valores muy superiores a 2 (y un caso no reportado por tamaño reducido del grupo), lo que confirma que las regiones preliminares no son homogéneas y que, aunque internamente no tengan estaciones discordantes fuertes, mezclan regímenes de precipitación máxima diaria. En consecuencia, se requiere un ajuste iterativo en la composición de las regiones para lograr $H < 2$.

Tabla 26

Valores de heterogeneidad en las regiones preliminares formadas

H\Región	R-1	R-2	R-3	R-4
H1	5.82	-	5.87	9.8
Resultado	Heterogénea	Heterogénea	Heterogénea	Heterogénea

b) Discordancia y homogeneidad (con L-momentos)

La Tabla 27 resume, por estación, el primer L-momento (l_1 , asociado al nivel medio de PMDA) y las razones L-momentos t ($L-C_v$), t_3 (L -asimetría) y t_4 (L -curtosis), que describen la forma de la distribución y son la base del ARF. Se observa variabilidad apreciable en t (aprox. 0.10 a 0.28), así como contrastes en t_3 y t_4 , lo que explica por

qué la agrupación preliminar no alcanzó homogeneidad: estaciones con distinta asimetría/curtosis tienden a generar H1 elevados cuando se combinan en una misma región.

Tabla 27

Estaciones y sus L-momentos

Id	Estación	l1	t	t3	t4	n
1	Porculla	35.0	0.2636	0.2782	0.1491	56
2	Huarmaca	66.1	0.2050	0.0731	0.1135	55
3	Tuluca	39.6	0.2158	0.1700	0.1083	47
4	Hacienda Shumaya	44.5	0.1544	0.3113	0.2238	56
5	Sondorillo	29.0	0.1554	0.1294	0.1098	25
6	Tabaconas	42.2	0.1485	0.1661	0.1234	53
7	El Limón	42.2	0.2309	0.2009	0.1014	26
8	Sallique	41.6	0.1889	0.0853	0.0890	49
9	Hacienda Pucará	47.7	0.1389	-0.1919	0.2326	28
10	Querocotillo	49.9	0.1399	0.1518	0.0584	24
11	Cutervo	46.8	0.0995	0.1048	0.0462	22
12	Cochabamba	58.9	0.1711	0.1893	0.0664	25
13	Llama	58.1	0.2102	0.1153	0.0784	51
14	Chontalí	54.2	0.1470	0.2398	0.2022	50
15	La Cascarilla	75.3	0.1035	0.0614	0.1574	28
16	Chancay Baños	52.1	0.1521	0.1274	0.1967	35
17	Chotano Lajas	45.6	0.1156	0.0606	0.1314	29
18	Tocmoche	53.8	0.2838	0.1429	0.0994	56
19	Quebrada Shugar	35.9	0.1423	0.3148	0.2646	23
20	Cueva Blanca	34.5	0.1176	0.3046	0.1253	26
21	Chugur	41.6	0.2804	0.1304	0.1088	60
22	Chota	46.3	0.1264	0.0859	-0.0450	21
23	Salala	42.0	0.1175	0.1480	0.0317	23
24	Colasay	56.1	0.1930	0.0673	0.0266	31
25	San Felipe	32.5	0.1980	0.1041	0.1924	32

Nota. Id: número de la estación, l1: L momento 1, t: L coeficiente de variación, t3: L coeficiente de asimetría, t4: L curtosis, n: número de datos de cada estación.

c) Ajuste iterativo y regiones definitivas

El objetivo consiste en satisfacer los criterios de discordancia y homogeneidad mediante la conformación de regiones hidrológicas dentro del área de estudio, tomando como punto de partida los cuatro grupos previamente establecidos. Para alcanzar este propósito, se llevaron a cabo los siguientes ajustes:

- Trasladar una o varias estaciones de una región a otra;
- Excluir una o varias estaciones del conjunto de datos;
- Reorganizar una región redistribuyendo sus estaciones hacia otras regiones.

Como resultado de esta rutina, se delimitaron cuatro regiones homogéneas, cada una integrada por 24 estaciones meteorológicas. La Región 1 y la Región 2 quedaron conformadas por 8 estaciones cada una, mientras que la Región 3 y la Región 4 reunieron 4 estaciones cada una. La estación Quebrada Shugar, incorporada en el inventario inicial, no quedó asignada a una región homogénea definitiva; por ello, se mantuvo documentada como parte de la base observacional, pero no intervino en la estimación regional final mediante ARF. La Tabla 28 presenta la composición estadística empleada para verificar la discordancia, la heterogeneidad y el ajuste probabilístico regional.

Tabla 28

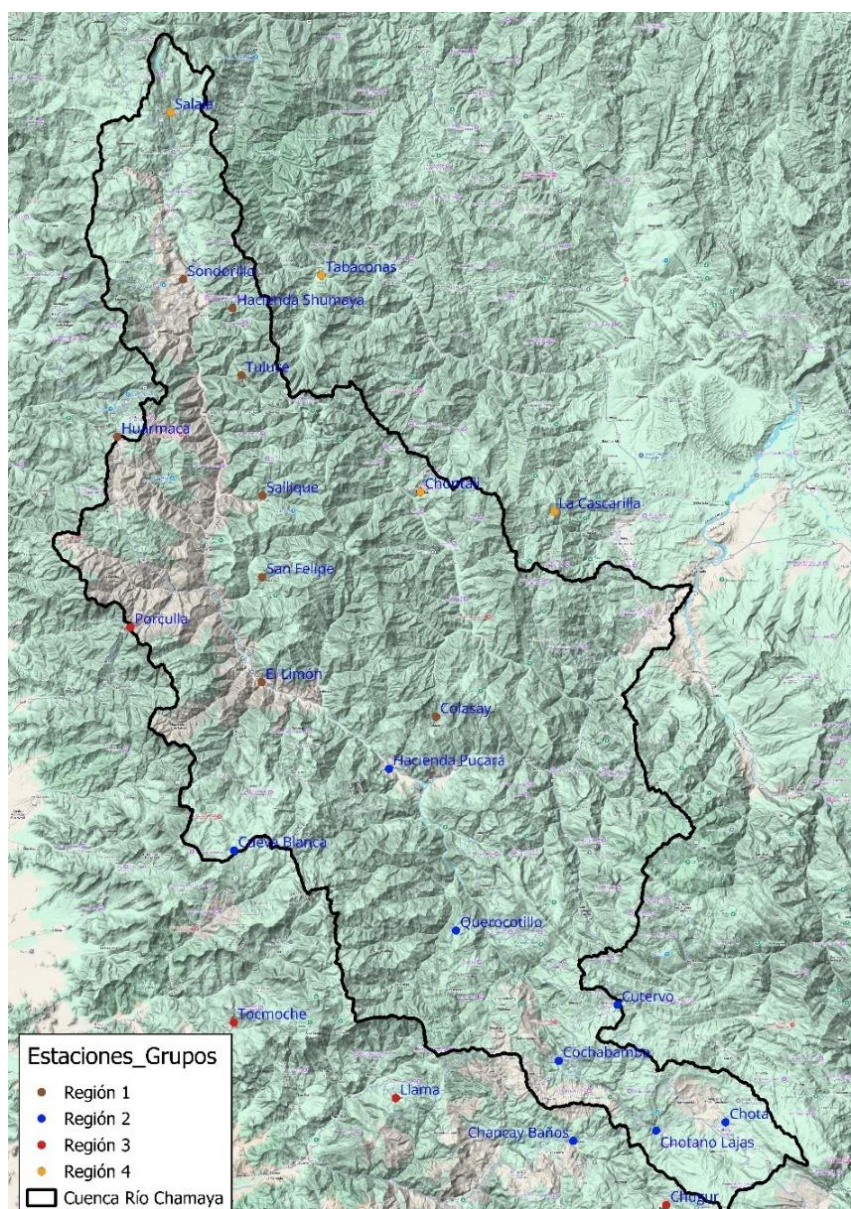
Regiones (4) y estaciones que las conforman

N°	Región 1	Región 2	Región 3	Región 4
1	Huarmaca	Hacienda Pucará	Porculla	Tabaconas
2	Tuluze	Querocotillo	Llama	Chontalí
3	Hacienda Shumaya	Cutervo	Tocmoche	La Cascarilla
4	Sondorillo	Cochabamba	Chugur	Salala
5	El Limón	Chancay Baños		
6	Sallique	Chotano Lajas		
7	Colasay	Cueva Blanca		
8	San Felipe	Chota		

El mapa verifica que la regionalización definitiva mantiene coherencia espacial y facilita la interpretación hidrológica de cada región. La separación espacial, junto con el reacomodo de estaciones respecto a los grupos preliminares, evidencia que el ajuste respondió a criterios estadísticos (D y H) y no únicamente a proximidad geográfica.

Figura 51

Regiones hidrológicas (4) en la zona de estudio



La Tabla 29 describe la variabilidad altitudinal dentro de cada región y permite contextualizar la homogeneidad alcanzada: aunque existe rango altitudinal, las regiones definidas reducen la mezcla de condiciones extremas dentro de un mismo grupo, lo cual contribuye a disminuir H1 y a consolidar regiones regionalmente consistentes.

Tabla 29

Rango de altitudes de regiones hidrológicas conformadas

Región	N (número de estaciones)	Altitud mínima (m.s.n.m.)	Altitud máxima (m.s.n.m.)
R-1	8	1110	2232
R-2	8	1012	3300
R-3	4	1399	2748
R-4	4	1627	2974

Los valores Dmax son < 3 en todas las regiones, confirmando ausencia de estaciones discordantes. Además, los valores de H1 son < 2 en todas las regiones, clasificándose como aceptablemente homogéneas (R-1, R-2, R-3) y definitivamente homogénea (R-4). Esto valida formalmente la regionalización final para proceder con la selección de distribución y la estimación de cuantiles.

Tabla 30

Máximos valores de discordancia en las regiones formadas

R-1	R-2	R-3	R-4
1.94	1.85	1.00	1.00

Tabla 31

Valores de heterogeneidad en las regiones formadas

H/Región	R-1	R-2	R-3	R-4
H1	1.74	1.15	1.43	0.92
Resultado	Aceptablemente Homogénea	Aceptablemente Homogénea	Aceptablemente Homogénea	Definitivamente Homogénea

d) Verificación gráfica con ratios L-momentos

Los diagramas muestran que los puntos (estaciones) se distribuyen en nubes relativamente compactas, sin valores aislados que sugieran discordancias extremas. Esta verificación visual es consistente con $D_{\max} < 3$ y respalda que la variabilidad de forma (asimetría y curtosis) se encuentra contenida en cada región, lo cual es favorable para el ARF.

Figura 52

Gráfico de dispersión de las ratios de Momentos L (relación entre el coeficiente de variación y el coeficiente de asimetría)

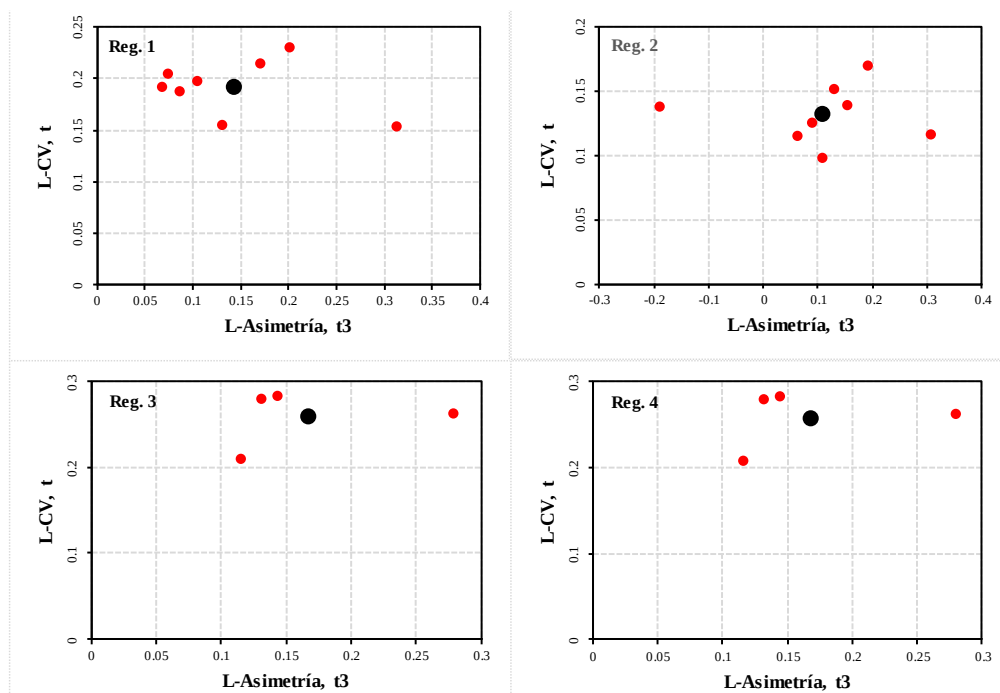
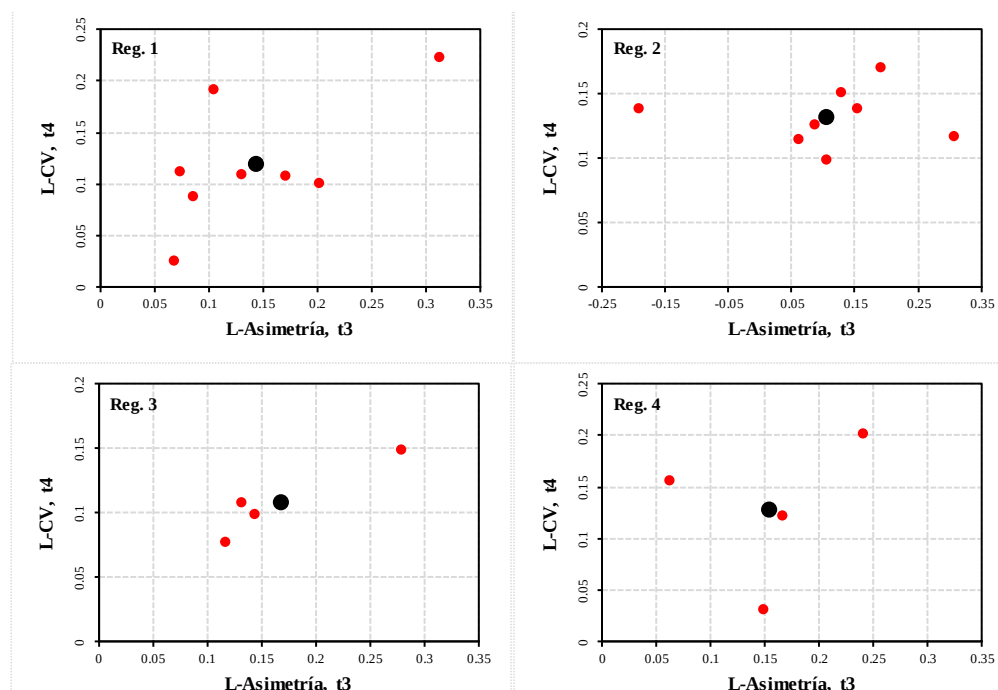


Figura 53

Gráfico de dispersión de las ratios de Momentos L (relación entre el coeficiente de curtosis y el coeficiente de asimetría)



e) Selección de distribución regional

La Tabla 32 evidencia que las distribuciones LNO, P3 y GEV cumplen mayoritariamente el criterio $|ZDIST| \leq 1.64$, mientras que GLO excede el umbral en todas las regiones, por lo que se descarta. En particular, GEV no resulta aceptable en R-3 ($|Z| = 1.94$), lo que refuerza la elección de una distribución alternativa más consistente entre regiones.

Tabla 32

Valores de Z por región según las distribuciones analizadas

Distribución\Región	R-1	R-2	R-3	R-4
LNO	0.26	0.32	0.63	0.25
P3	0.31	0.6	1.1	0.28
GEV	0.38	0.65	1.94	0.31
GLO	2.4	2.66	3.92	1.65

En la Tabla 33 con base en los menores valores de $|Z|$ y el cumplimiento del criterio, se selecciona LNO como distribución regional para las cuatro regiones, asegurando un marco uniforme para el cálculo de cuantiles regionales y la comparación posterior con ROI.

Tabla 33

Distribución de frecuencia óptima para cada región hidrológica

R-1	R-2	R-3	R-4
LNO	LNO	LNO	LNO

La distribución LNO fue seleccionada para las cuatro regiones debido a que presentó los menores valores absolutos del estadístico de bondad de ajuste: 0.26 en R-1, 0.32 en R-2, 0.63 en R-3 y 0.25 en R-4. La curva regional LNO produjo cuantiles adimensionales positivos en el rango evaluado. La revisión de la cola inferior de los ajustes probabilísticos locales identificó estimaciones negativas en algunas estaciones para periodos de retorno extremadamente próximos a 1 año. Tales valores no tienen interpretación física como precipitación y no fueron utilizados en el diseño hidrológico ni en la elaboración de productos cartográficos. Para las aplicaciones ingenieriles se consideraron cuantiles con $TR \geq 2$ años, mientras que los cálculos completos permanecieron disponibles en los archivos Excel de respaldo (Anexo 3).

f) Cuantiles regionales adimensionales y curvas regionales

La Tabla 34 presenta los cuantiles regionales adimensionales para LNO en función del TR. Se observa que el crecimiento con TR no es uniforme entre regiones: por ejemplo, para $TR = 1000$, la región 3 alcanza $q_R = 3.35$, superior a las demás (R-1: 2.63; R-2: 1.98; R-4: 2.22), lo que indica una cola más pesada y mayor amplificación de extremos en la región 3. Este comportamiento define diferencias regionales en la precipitación de diseño aun dentro de una misma familia de distribución.

Tabla 34

Cuantiles regionales adimensionales para la distribución con mejor ajuste de las regiones formadas

TR	PE (%)	PNE (%)	q	Región 1	Región 2	Región 3	Región 4
			$-\ln(-\ln(F))$	LNO	LNO	LNO	LNO
1.00	99.9	0.1	-1.93	0.29	0.44	0.08	0.53
1.00	99.8	0.2	-1.83	0.32	0.47	0.12	0.54
1.01	99.5	0.5	-1.67	0.37	0.51	0.17	0.57
1.01	99.0	1.0	-1.53	0.40	0.55	0.22	0.60
1.02	98.0	2.0	-1.36	0.45	0.59	0.27	0.62
1.05	95.0	5.0	-1.10	0.53	0.65	0.37	0.67
1.11	90.0	10.0	-0.83	0.60	0.71	0.47	0.72
1.25	80.0	20.0	-0.48	0.71	0.80	0.60	0.79
2	50.0	50.0	0.37	0.95	0.98	0.92	0.96
5	20.0	80.0	1.50	1.26	1.19	1.35	1.18
10	10.0	90.0	2.25	1.46	1.32	1.63	1.33
20	5.0	95.0	2.97	1.65	1.43	1.90	1.46
50	2.0	98.0	3.90	1.88	1.57	2.24	1.64
100	1.0	99.0	4.60	2.06	1.67	2.49	1.77
200	0.5	99.5	5.30	2.23	1.77	2.75	1.91
500	0.2	99.8	6.21	2.45	1.89	3.09	2.08
1000	0.1	99.9	6.91	2.63	1.98	3.35	2.22

Cada curva representa la relación q_R vs $y = -\ln(-\ln(F))$ y permite interpolar cuantiles para TR específicos. La separación entre curvas regionales refleja diferencias en crecimiento de extremos; regiones con curvas más empinadas incrementan más rápidamente sus cuantiles para TR altos, coherente con los valores de q_R observados en la tabla.

Figura 54

Curvas de distribución de frecuencias de cuantiles regionales, región 1

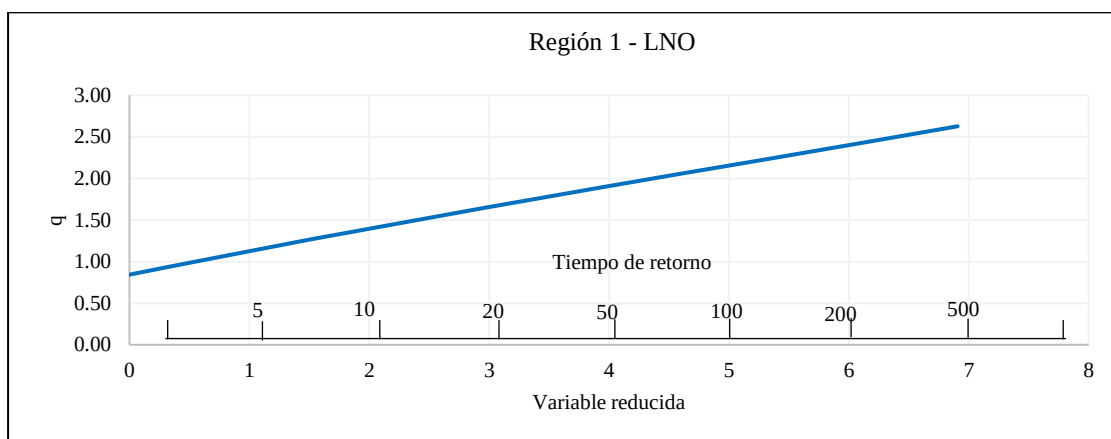


Figura 55

Curvas de distribución de frecuencias de cuantiles regionales, región 2

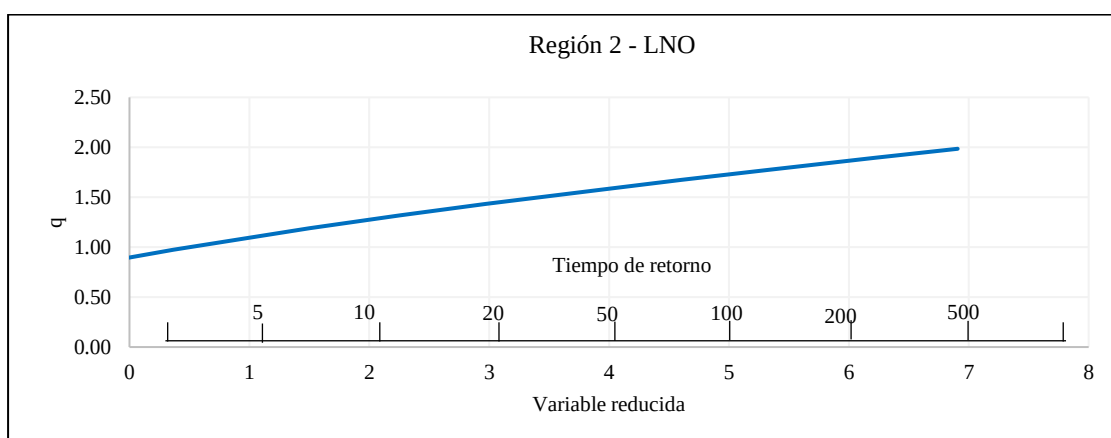


Figura 56

Curvas de distribución de frecuencias de cuantiles regionales, región 3

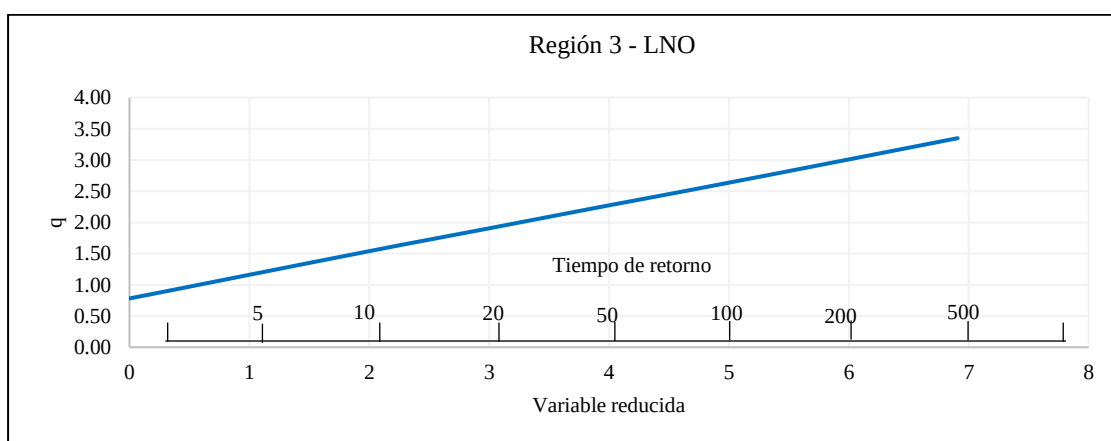
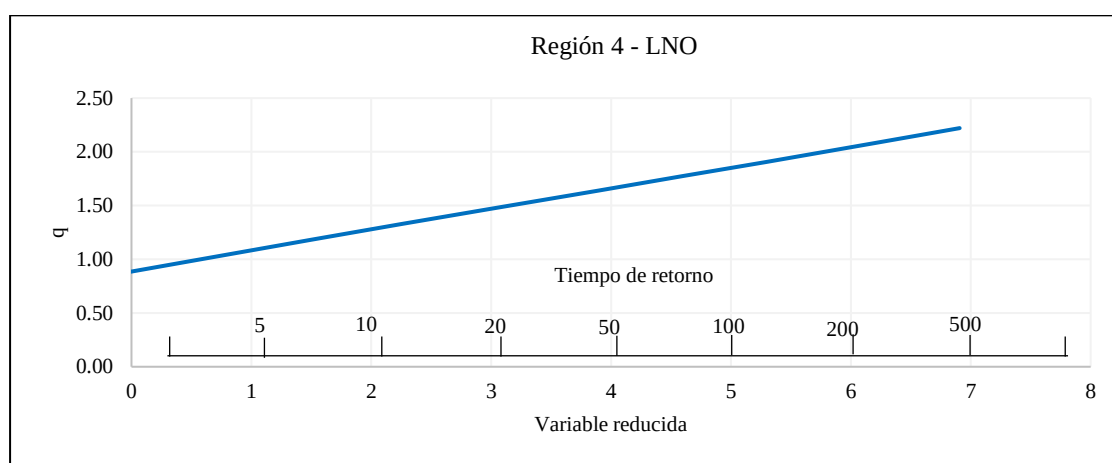


Figura 57

Curvas de distribución de frecuencias de cuantiles regionales, región 4



g) Cuantiles locales

Las Tablas 35 a 38 presentan los cuantiles locales de precipitaciones máximas diarias anuales (PMDA) estimados mediante el ajuste probabilístico seleccionado para cada estación. Estos valores corresponden al registro diario acumulado en intervalos de observación fijos. Cuando fue necesario expresar los resultados como precipitación máxima en una ventana continua de 24 horas para fines de diseño o para efectuar comparaciones con antecedentes que reportaron dicha magnitud, se aplicó el factor de corrección por intervalo fijo de observación igual a 1.13. La conversión se realizó una sola vez, sin modificar la identificación de las tablas de PMDA ni confundir ambas magnitudes.

$$P24h(TR) = 1.13 \times PMDA(TR) \quad (92)$$

Donde: P24h(TR): precipitación máxima estimada para una ventana continua de 24 horas y un periodo de retorno determinado, en mm. PMDA(TR): precipitación máxima diaria anual obtenida a partir de registros acumulados en intervalos fijos, en mm. 1.13: factor de corrección por intervalo fijo de observación. TR: periodo de retorno, expresado en años.

La Tabla 35 presenta los cuantiles locales de diseño de las estaciones que integran la Región 1 para periodos de retorno iguales o mayores que 2 años. Huarmaca concentra las mayores precipitaciones estimadas del grupo: alcanza 136 mm para un periodo de retorno de 100 años y 168 mm para 1 000 años. Colasay ocupa una posición intermedia-alta, con 113 mm y 149 mm para los mismos periodos. En contraste, Sondorillo y San Felipe registran magnitudes menores, aunque mantienen el crecimiento progresivo esperado al aumentar la recurrencia. El análisis diagnóstico de la cola inferior mostró valores no físicos en El Limón y Colasay cuando el periodo de retorno se aproxima a 1 año. Estos resultados no fueron utilizados para diseño ni para la elaboración de mapas regionalizados; se conservaron únicamente en el anexo técnico como evidencia de una limitación del ajuste LNO en probabilidades extremas inferiores.

Tabla 35

Cuantiles locales de PMDA estimados mediante el ajuste probabilístico seleccionado para cada estación de la Región 1

			2	3	4	5	7	8	24	25
TR	PE (%)	PNE (%)	Huarmaca	Tuluca	Hacienda Shumaya	Sondorillo	El Limón	Sallique	Colasay	San Felipe
2	50.00	50.00	63	38	40	28	42	39	55	31
5	20.00	80.00	85	52	53	35	57	52	71	42
10	10.00	90.00	99	60	63	39	64	61	80	48
20	5.00	95.00	111	68	72	44	71	69	90	54
50	2.00	98.00	125	77	83	50	78	79	103	61
100	1.00	99.00	136	84	91	56	82	86	113	65
200	0.50	99.50	146	90	99	61	87	94	123	70
500	0.20	99.80	158	99	109	70	92	104	138	76
1000	0.10	99.90	168	105	117	77	96	112	149	80

Nota. Los valores se expresan en mm. La tabla presenta únicamente cuantiles con $TR \geq 2$ años, ya que corresponden al rango aplicable para la estimación de precipitaciones de diseño. Los resultados obtenidos para periodos de retorno inferiores a 2 años se incorporaron en un anexo técnico como evidencia diagnóstica de la cola inferior del ajuste probabilístico.

La Tabla 36 resume los cuantiles locales de PMDA calculados para las estaciones que conforman la región 2 (Hacienda Pucará, Querocotillo, Cutervo, Cochabamba, Chancay Baños, Chotano Lajas, Cueva Blanca y Chota). Los resultados evidencian un crecimiento constante de la precipitación de diseño con el incremento del periodo de retorno, lo que indica estabilidad en el comportamiento de las frecuencias a escala local. Para TR = 100, los cuantiles se ubican aproximadamente entre 74 mm y 100 mm, destacando Cochabamba con valores cercanos a 100 mm, mientras que estaciones como Cutervo y Chota presentan cuantiles moderados en torno a 74–76 mm. Para TR = 1 000, se mantiene la misma tendencia, con máximos cercanos a 114 mm en Cochabamba y valores alrededor de 89–97 mm en el resto de estaciones, lo cual sugiere una respuesta regional con extremos relativamente moderados frente a otras regiones.

Tabla 36

Cuantiles locales de PMDA estimados mediante el ajuste probabilístico seleccionado para cada estación de la Región 2

			9	10	11	12	16	17	20	22
TR	PE (%)	PNE (%)	Hacienda Pucará	Querocotillo	Cutervo	Cochabamba	Chancay Baños	Chotano Lajas	Cueva Blanca	Chota
2	50.00	50.00	47	49	45	59	52	44	33	45
5	20.00	80.00	57	60	53	74	64	53	40	55
10	10.00	90.00	63	66	58	82	70	58	44	60
20	5.00	95.00	69	72	63	88	75	64	49	65
50	2.00	98.00	75	79	69	96	81	71	54	72
100	1.00	99.00	80	83	74	100	85	76	58	76
200	0.50	99.50	84	88	79	105	88	81	62	81
500	0.20	99.80	89	93	85	110	92	88	68	86
1000	0.10	99.90	93	97	89	114	95	94	72	91

Nota. Los valores se expresan en mm. La tabla presenta cuantiles locales de PMDA con TR ≥ 2 años, rango adoptado para aplicaciones hidrológicas. Los cálculos completos, incluidos los valores diagnósticos asociados con TR < 2 años, se entregan en los archivos Excel de respaldo y no fueron empleados para el dimensionamiento hidráulico ni para la elaboración de productos cartográficos.

La Tabla 37 reúne los cuantiles locales de diseño correspondientes a la Región 3. Tocmoche y Llama presentan las mayores precipitaciones del grupo para periodos de retorno altos. En Tocmoche se estimaron 137 mm para 100 años y 178 mm para 1 000 años; en Llama, los valores alcanzaron 130 mm y 173 mm, respectivamente. Porculla muestra las magnitudes más moderadas. La revisión de la cola inferior identificó cuantiles negativos en Tocmoche y Chugur para periodos de retorno muy próximos a 1 año. Al tratarse de valores incompatibles con la naturaleza no negativa de la precipitación, se interpretaron como artefactos matemáticos del ajuste LNO fuera de su rango de uso ingenieril y se excluyeron de las tablas destinadas al diseño.

Tabla 37

Cuantiles locales de PMDA estimados mediante el ajuste probabilístico seleccionado para cada estación de la Región 3

			1	13	18	21
TR	PE (%)	PNE (%)	Porculla	Llama	Tocmoche	Chugur
2	50.00	50.00	32	54	50	39
5	20.00	80.00	47	74	75	58
10	10.00	90.00	57	88	91	70
20	5.00	95.00	67	101	105	80
50	2.00	98.00	79	117	124	94
100	1.00	99.00	88	130	137	104
200	0.50	99.50	98	143	150	113
500	0.20	99.80	110	160	166	126
1000	0.10	99.90	119	173	178	135

Nota. Los valores se expresan en mm. La tabla presenta únicamente cuantiles con $TR \geq 2$ años, rango utilizado para fines de diseño. Los resultados correspondientes a la cola inferior del ajuste probabilístico se trasladaron a un anexo técnico, debido a que no constituyen precipitaciones de diseño.

La Tabla 38 presenta los cuantiles locales de PMDA para las estaciones de la región 4 (Tabaconas, Chontalí, La Cascarilla y Salala), mostrando el patrón esperado de incremento de la precipitación conforme aumenta el periodo de retorno. En esta región, La Cascarilla registra los mayores cuantiles en todos los TR, alcanzando 121

mm para TR = 100 y 146 mm para TR = 1 000, lo cual indica una mayor magnitud de eventos extremos respecto al resto del grupo. Por su parte, Tabaconas y Salala muestran cuantiles menores, con valores cercanos a 75 mm y 72 mm para TR = 100, y alrededor de 92 mm y 98 mm para TR = 1 000, respectivamente, mientras que Chontalí presenta valores intermedios (96 mm para TR = 100 y 118 mm para TR = 1 000). En conjunto, la región 4 presenta diferenciación entre estaciones dominantes en magnitud extrema (La Cascarilla) y estaciones con respuesta más moderada.

Tabla 38

Cuantiles locales de PMDA estimados mediante el ajuste probabilístico seleccionado para cada estación de la Región 4

			6	14	15	23
TR	PE (%)	PNE (%)	Tabaconas	Chontalí	La Cascarilla	Salala
2	50.00	50.00	41	52	73	41
5	20.00	80.00	51	65	86	48
10	10.00	90.00	57	73	94	53
20	5.00	95.00	63	81	102	58
50	2.00	98.00	70	90	113	66
100	1.00	99.00	75	96	121	72
200	0.50	99.50	81	103	128	79
500	0.20	99.80	87	111	139	89
1000	0.10	99.90	92	118	146	98

Nota. Los valores se expresan en mm. La tabla presenta cuantiles locales de PMDA con TR ≥ 2 años, rango adoptado para aplicaciones hidrológicas. Los cálculos completos, incluidos los valores diagnósticos asociados con TR < 2 años, se entregan en los archivos Excel de respaldo y no fueron empleados para el dimensionamiento hidráulico ni para la elaboración de productos cartográficos.

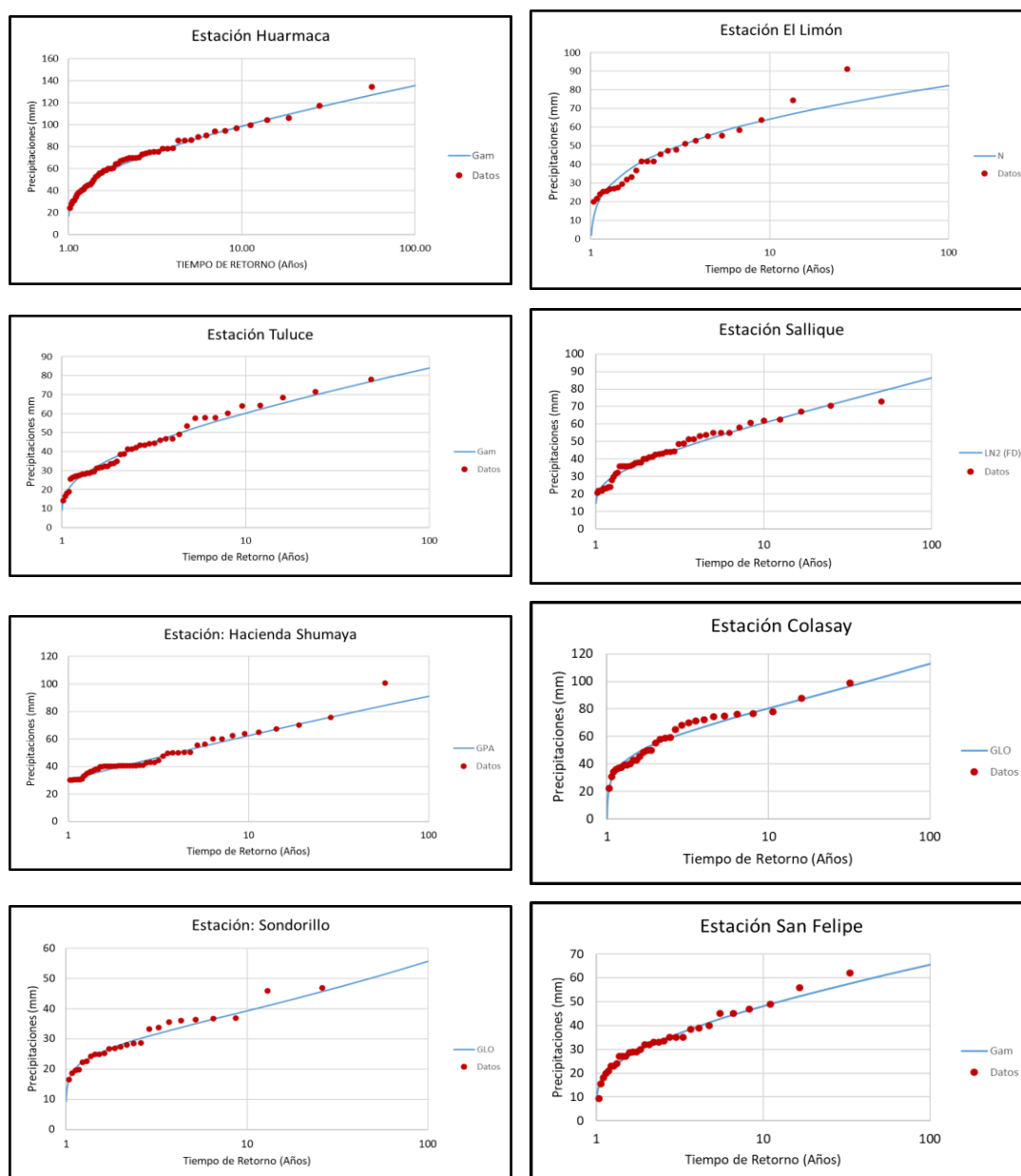
La revisión de los cuantiles locales evidenció valores negativos para periodos de retorno extremadamente próximos a 1 año en las estaciones El Limón, Colasay, Tocmoche y Chugur. Estas estimaciones no representan precipitaciones físicamente posibles, ya que la variable analizada solo admite valores iguales o superiores a cero. El comportamiento observado se debe a una limitación de los ajustes probabilísticos

locales al extrapolarse a la cola inferior de la distribución, fuera del rango de interés para el diseño hidrológico. En particular, los resultados negativos no provienen de una única familia probabilística: El Limón fue ajustada mediante distribución Normal, Colasay mediante Logística Generalizada, Tocmoche mediante Valor Extremo Generalizado y Chugur mediante Lognormal. Por esta razón, los valores asociados a periodos de retorno menores de 2 años se conservaron únicamente como evidencia diagnóstica y no fueron utilizados para el dimensionamiento hidráulico ni para la elaboración de productos cartográficos de diseño. Para aplicaciones ingenieriles se adoptaron cuantiles con $TR \geq 2$ años. Cuando se requiera profundizar en el comportamiento de la cola inferior, deberá efectuarse un análisis de sensibilidad con distribuciones cuyo soporte sea compatible con la naturaleza no negativa de la precipitación, como Gamma, Lognormal o Pareto Generalizada, verificando en cada caso la bondad de ajuste y la estabilidad de los cuantiles de la cola inferior.

Las Figura 58 a Figura 61 presentan, para cada región hidrológica, las curvas locales de distribución de frecuencias ajustadas a las estaciones y su contraste con los datos empíricos de precipitación máxima. En términos generales, se observa que las curvas ajustadas reproducen adecuadamente la tendencia creciente de la precipitación a medida que aumenta el tiempo de retorno, lo que evidencia la coherencia del ajuste y la estabilidad de la estimación de cuantiles locales para fines de diseño. Asimismo, las figuras permiten identificar diferencias de magnitud entre estaciones dentro de cada región, lo que refleja la variabilidad espacial intrarregional, sin afectar el patrón general de crecimiento ni la consistencia del comportamiento probabilístico esperado.

Figura 58

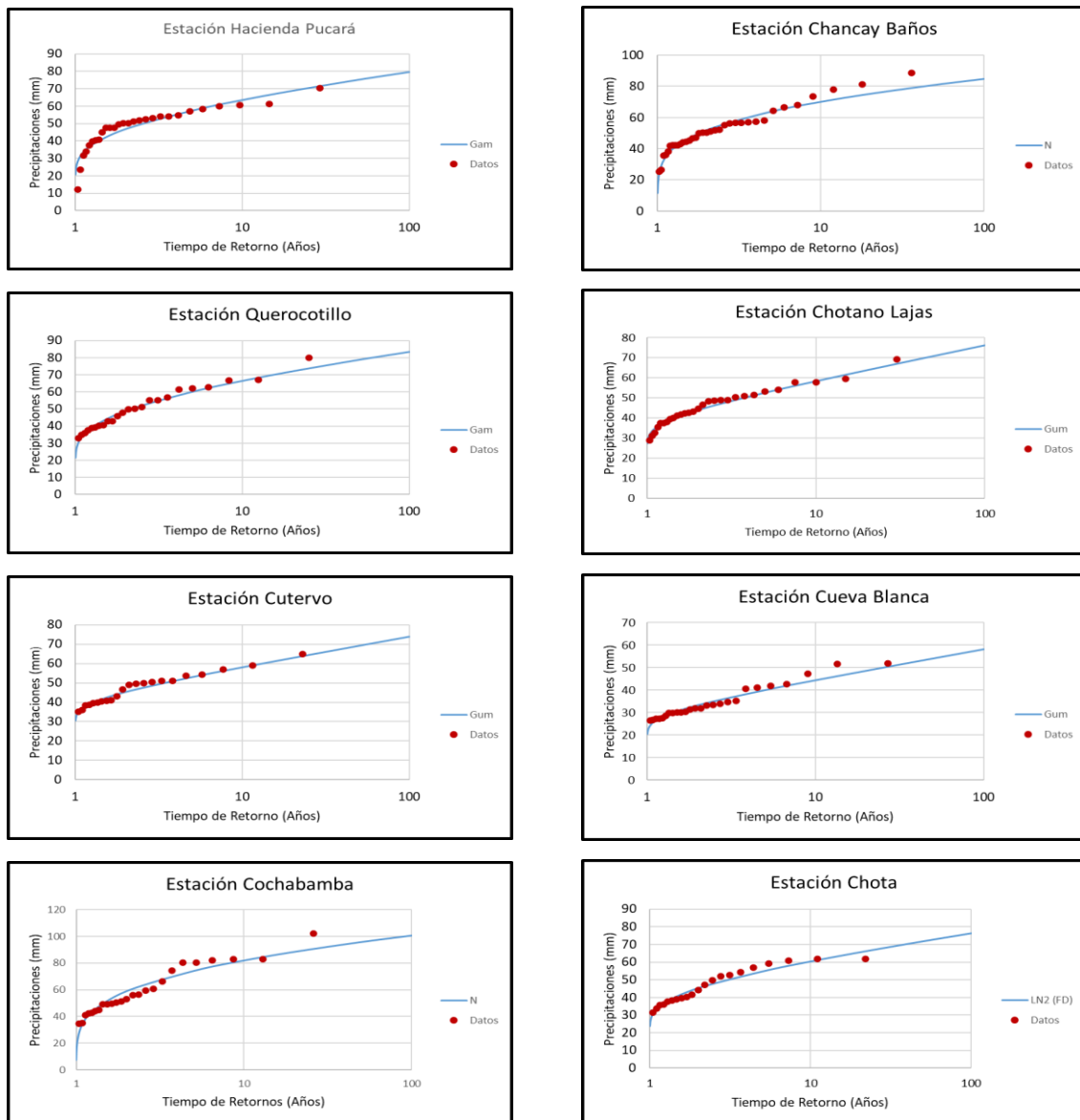
Curvas de distribución de frecuencias locales de cada estación de la región 1



Nota. Visualmente, se observa que la tendencia general es coherente, dado que la curva ajustada reproduce el crecimiento progresivo de la precipitación conforme aumenta el tiempo de retorno. No obstante, se identifican ligeras diferencias entre puntos y curva en los TR más altos, lo cual es esperable debido a la menor cantidad de observaciones en la cola superior y a la influencia de valores extremos. En estaciones como Huarmaca, la curva se extiende a cuantiles mayores, reflejando una mayor severidad de eventos extremos dentro de la región, mientras que estaciones como Sondorillo y San Felipe mantienen niveles menores en toda la escala de retorno, confirmando la variabilidad espacial intrarregional en la magnitud de la PMDA.

Figura 59

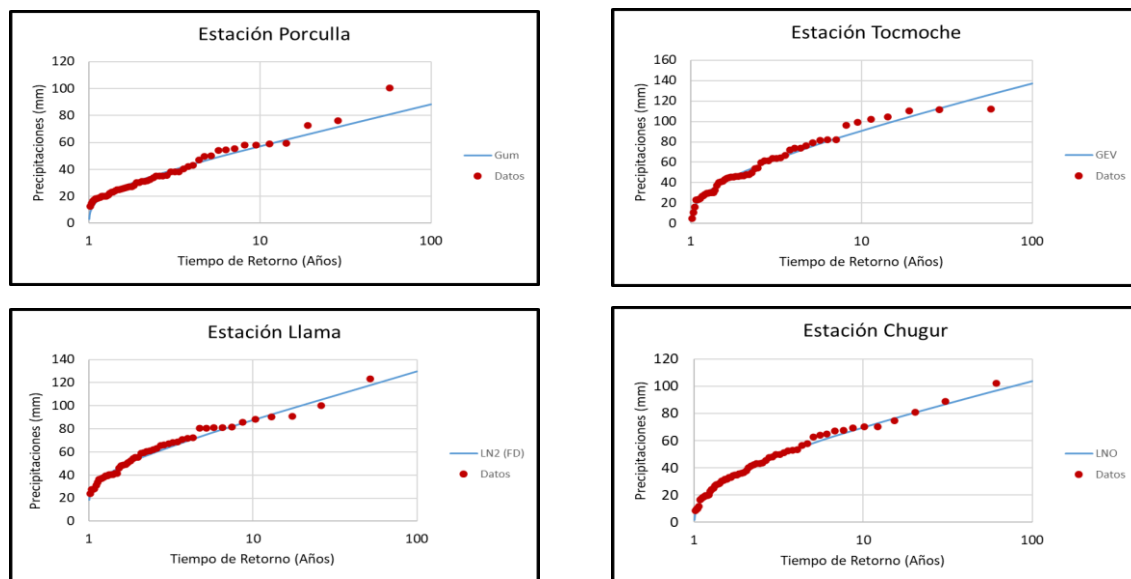
Curvas de distribución de frecuencias locales de cada estación de la región 2



Nota. En la mayoría de estaciones, la curva ajustada sigue adecuadamente la nube de puntos, lo que sugiere consistencia en el ajuste y estabilidad en la estimación de cuantiles locales. Las discrepancias se concentran principalmente en los TR altos, donde algunos puntos se separan de la curva por la incertidumbre asociada a la extrapolación. En términos comparativos dentro de la región, Cochabamba y Chancay Baños muestran trayectorias con mayor elevación en TR altos, mientras que estaciones como Cutervo y Chota presentan incrementos más controlados, lo que coincide con los cuantiles reportados para periodos de retorno de diseño.

Figura 60

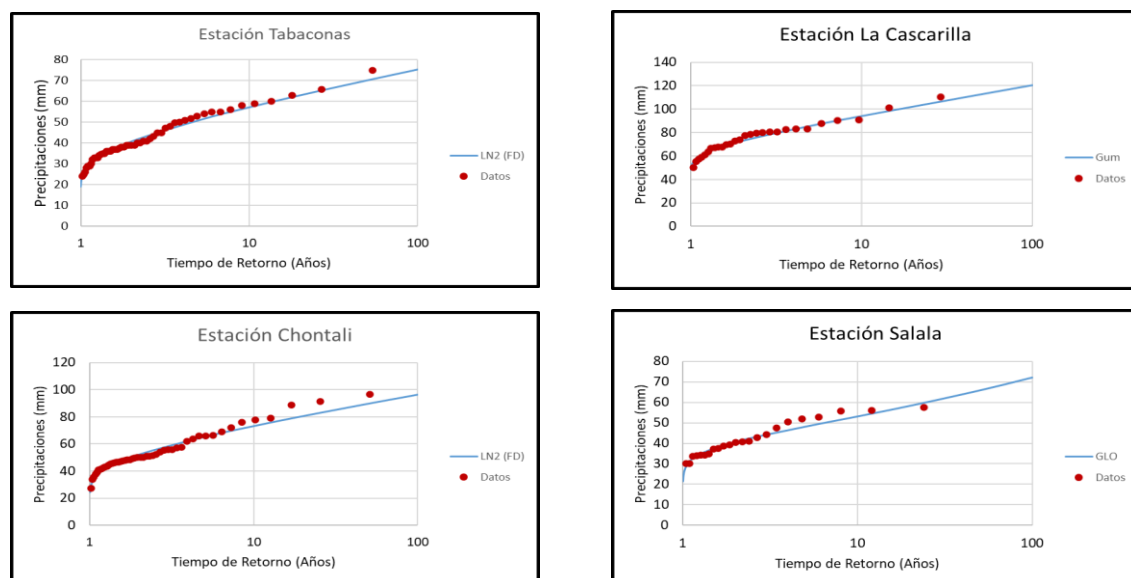
Curvas de distribución de frecuencias locales de cada estación de la región 3



Nota. En esta región, las diferencias entre estaciones se hacen más evidentes en la cola superior, confirmando que Llama y Toccoche concentran los mayores cuantiles de diseño, mientras que Porculla se mantiene en nivel inferior y Chugur presenta comportamiento intermedio.

Figura 61

Curvas de distribución de frecuencias locales de cada estación de la región 4



Nota. El comportamiento comparativo indica que Tabaconas y Salala presentan cuantiles menores y crecimiento más moderado, mientras que Chontalí mantiene valores intermedios, confirmando que la estación dominante en extremos dentro de esta región es La Cascarilla.

4.1.2.2.Región de influencia.

a) Conformación de la región de influencia por estación objetivo

La Tabla 39 presenta la conformación de seis grupos bajo el enfoque de Región de Influencia, donde cada grupo se define en función de una estación objetivo y su conjunto de estaciones “vecinas” con comportamiento estadístico similar. A diferencia del ARF, estos grupos no requieren una delimitación regional rígida, debido a que la selección se realiza de manera específica para cada punto de interés. En consecuencia, se observa que algunas estaciones se repiten en distintos grupos, lo cual evidencia la existencia de zonas de transición hidrometeorológica y la necesidad de que la información regional se adapte al entorno inmediato de la estación evaluada. De este modo, se sustenta que el ROI genera agrupamientos flexibles, orientados a estimar cuantiles de diseño representativos a escala local, pero soportados por información regional.

Tabla 39

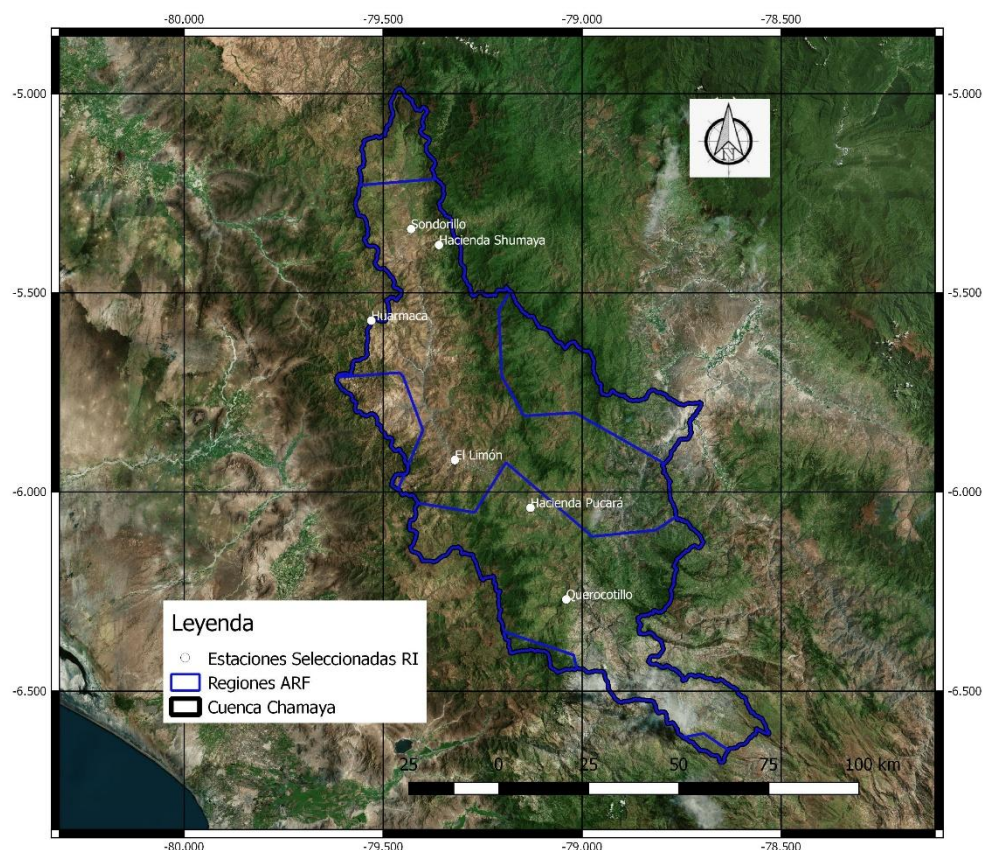
Estaciones agrupadas para diferentes periodos de retorno

N°	Región 1	Región 2	Región 3	Región 4	Región 5	Región 6
			Hacienda		Hacienda	
1	Huarmaca	Huarmaca	Pucará	El Limón	Shumaya	Huarmaca
2	Tuluze	Tuluze	Querocotillo	Sallique	Sondorillo	Tuluze
	Hacienda	Hacienda		Hacienda		Hacienda
3	Shumaya	Shumaya	Cutervo	Pucará	Tabaconas	Shumaya
4	Sondorillo	Sondorillo	Cochabamba	Querocotillo	Chontalí	Sondorillo
5	Tabaconas	El Limón	Chancay Baños	Colasay	La Cascarilla	El Limón
6	Sallique	Sallique	Chotano Lajas	San Felipe	Salala	Sallique
			Quebrada			
7	Chontalí	Chontalí	Shugar			San Felipe
8	San Felipe	Colasay	Chota			
9		San Felipe				

La Figura 62 muestra la distribución espacial de las estaciones seleccionadas mediante el enfoque de Región de Influencia, sobre el ámbito de la cuenca Chamaya y en referencia a las regiones definidas por ARF. Se aprecia que la selección ROI no se restringe estrictamente a los límites regionales del ARF, sino que incorpora estaciones cercanas ubicadas en regiones contiguas cuando presentan similitud estadística y potencial contribución a la estimación de cuantiles. Esta representación confirma el carácter adaptativo del ROI, dado que permite conformar conjuntos de datos que responden a la vecindad y a la continuidad espacial del régimen de precipitación, mejorando la robustez de la estimación para la estación objetivo, especialmente en el rango de periodos de retorno altos.

Figura 62

Estaciones agrupadas para diferentes periodos de retorno

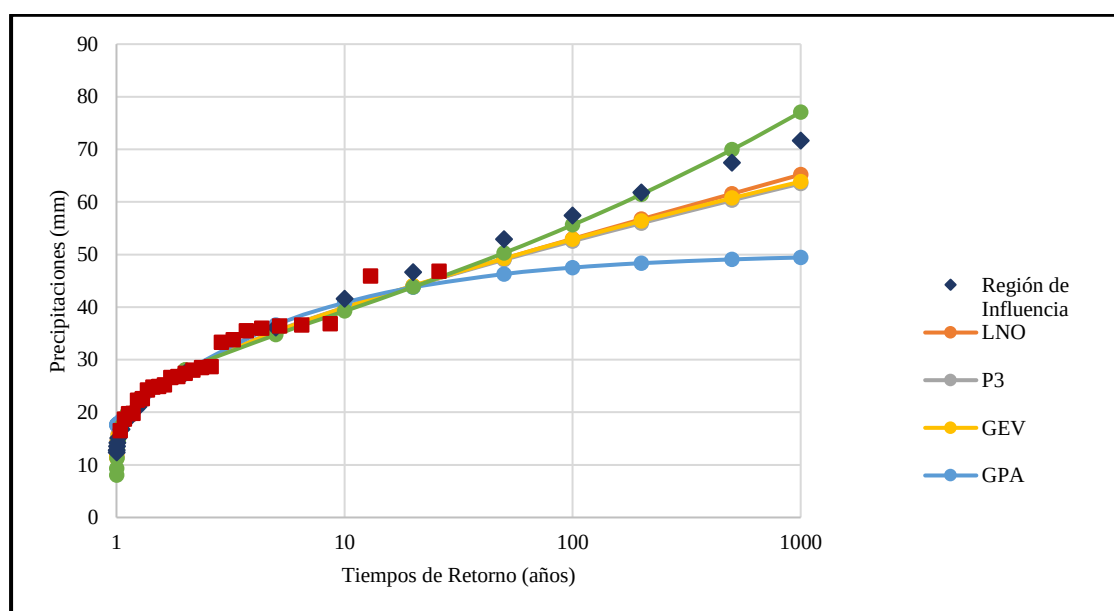


b) Estimación de precipitaciones de diseño y sensibilidad a la distribución

La Figura 63 compara el comportamiento de la precipitación en función del tiempo de retorno para la estación Sondorillo, incorporando tanto la información observada de la estación como la estimación regional derivada de su Región de Influencia. Se evidencia un ajuste satisfactorio en el rango de recurrencias frecuentes y medias, donde la tendencia regional reproduce la forma general del crecimiento de los cuantiles. Para recurrencias altas, las curvas asociadas a distintas distribuciones se separan, lo cual indica sensibilidad de la precipitación de diseño a la elección del modelo probabilístico. En este marco, el ROI permite disponer de un “abanico” de estimaciones para el extremo, y facilita seleccionar la distribución con mejor desempeño estadístico y coherencia hidrológica para el cuantil de diseño.

Figura 63

Precipitación de diseño para un periodo de retorno de 100 años y distribuciones de frecuencias para la estación Sondorillo (Grupo 1)

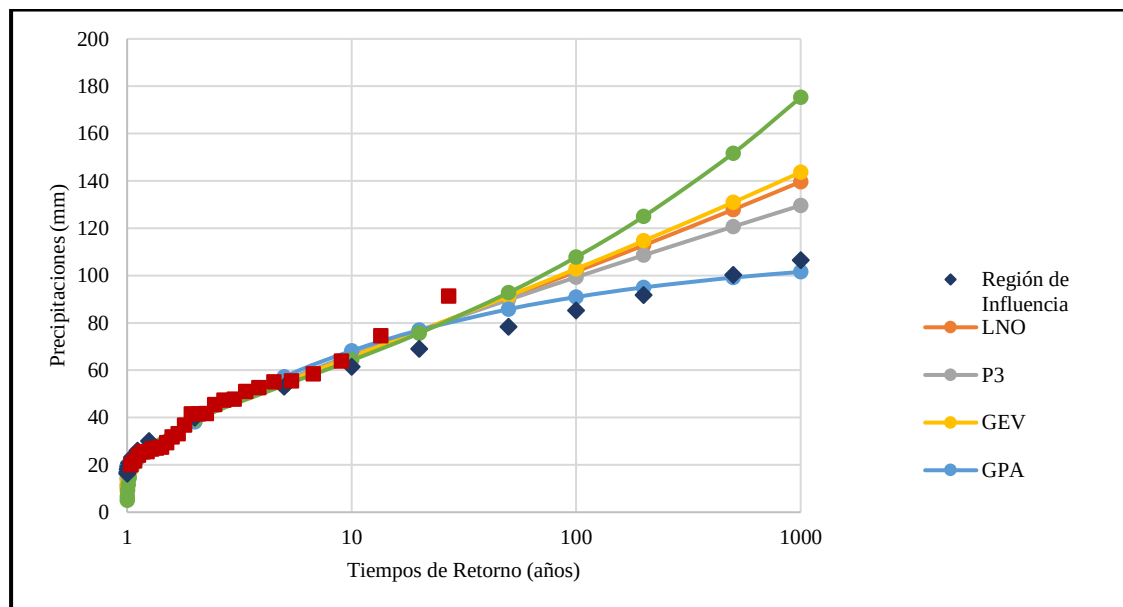


Nota. Los puntos representan la serie observada de la estación Sondorillo, mientras que las curvas corresponden a las distribuciones ajustadas con datos de la Región de Influencia; el cuantil de diseño para $TR = 100$ años se obtiene en la intersección con dicho periodo de retorno.

En la Figura 64 se observa que los cuantiles regionales obtenidos por la Región de Influencia mantienen consistencia con la tendencia de los datos de la estación El Limón en la mayor parte del dominio de análisis, lo cual respalda la representatividad del grupo seleccionado. A medida que aumenta el tiempo de retorno, se incrementa la divergencia entre distribuciones, reflejando diferencias en la forma de la cola y, por lo tanto, en la estimación de eventos raros. Este resultado es relevante para el diseño hidráulico, debido a que el ROI no solo transfiere información regional, sino que también permite evaluar la incertidumbre asociada a la selección de la distribución para TR elevados.

Figura 64

Precipitación de diseño para un periodo de retorno de 100 años y distribuciones de frecuencias para la estación El Limón (Grupo 2)

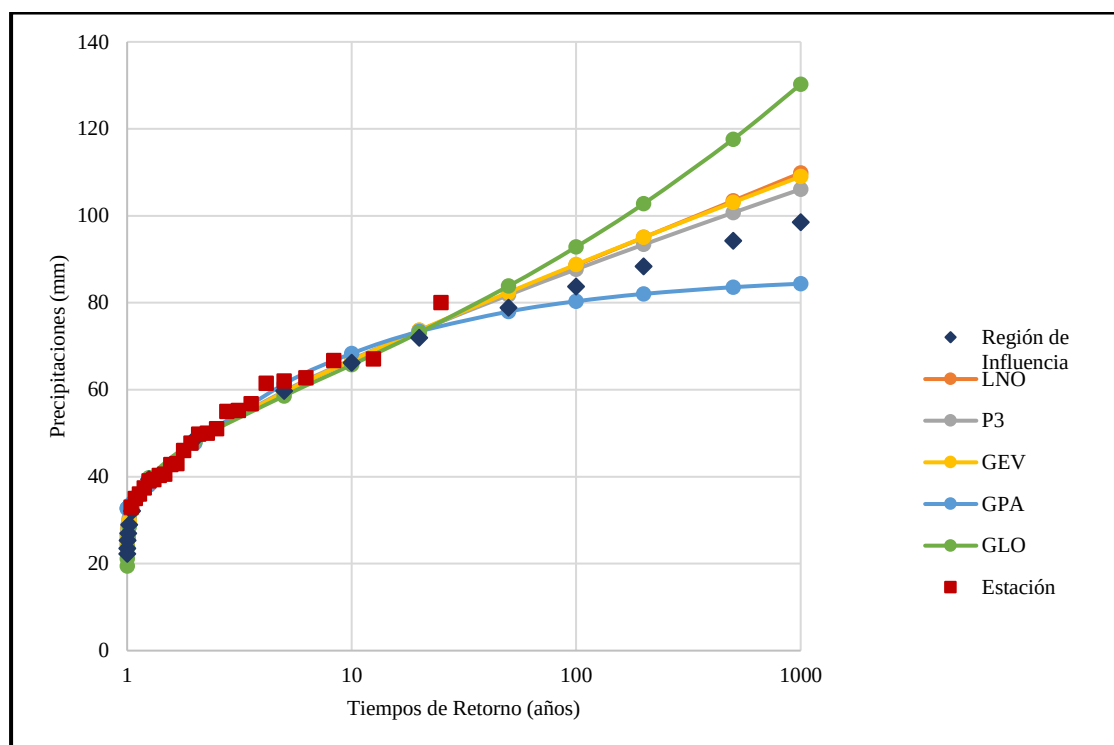


Nota. La estimación por Región de Influencia integra estaciones vecinas estadísticamente similares a El Limón; la separación entre curvas en recurrencias altas refleja la sensibilidad del cuantil extremo a la distribución adoptada.

La Figura 65 evidencia que, para la estación Querocotillo, las estimaciones basadas en la Región de Influencia siguen el patrón general de crecimiento de los cuantiles con el tiempo de retorno, reproduciendo adecuadamente la curvatura observada en los datos. En el tramo de mayores recurrencias, las diferencias entre distribuciones se hacen más notorias, lo cual sugiere que la caracterización de la cola es un factor dominante en la precipitación de diseño. En este contexto, el ROI aporta consistencia regional y permite contrastar modelos probabilísticos para seleccionar el más estable, evitando que el cuantil extremo dependa únicamente del tamaño muestral local.

Figura 65

Precipitación de diseño para un periodo de retorno de 100 años y distribuciones de frecuencias para la estación Querocotillo (Grupo 3)

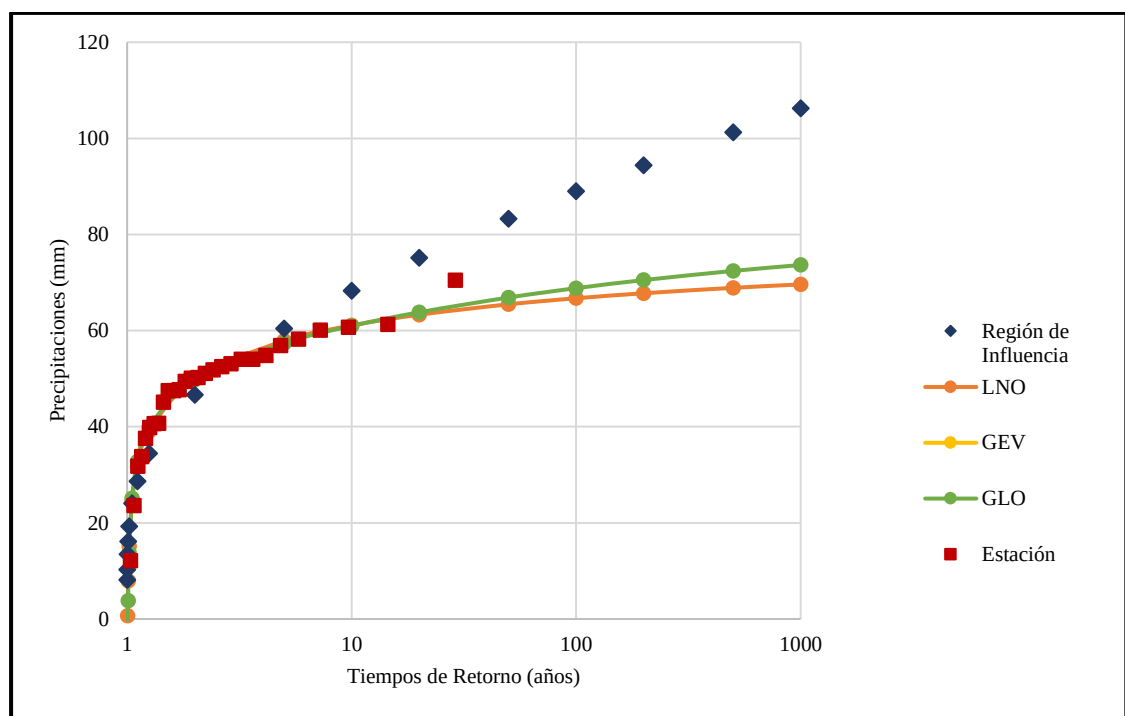


Nota. La figura compara el ajuste de distintas distribuciones para Querocotillo bajo el enfoque ROI; las diferencias en la cola superior controlan las estimaciones para periodos de retorno elevados.

En la Figura 66 se aprecia que las curvas ajustadas presentan un comportamiento diferenciado en recurrencias altas, mientras que en recurrencias bajas y medias muestran una tendencia compatible con la serie observada. Este patrón confirma que la estimación del evento de diseño para TR altos depende del grado de “pesadez” de la cola que impone cada distribución. En términos metodológicos, el ROI reduce la fragilidad del ajuste local al incorporar estaciones con respuesta similar, pero también exige justificar la distribución seleccionada, debido a que pequeñas variaciones en la cola pueden traducirse en cambios relevantes del cuantil de diseño.

Figura 66

Precipitación de diseño para un periodo de retorno de 100 años y distribuciones de frecuencias para la estación Hacienda Pucará (Grupo 4)

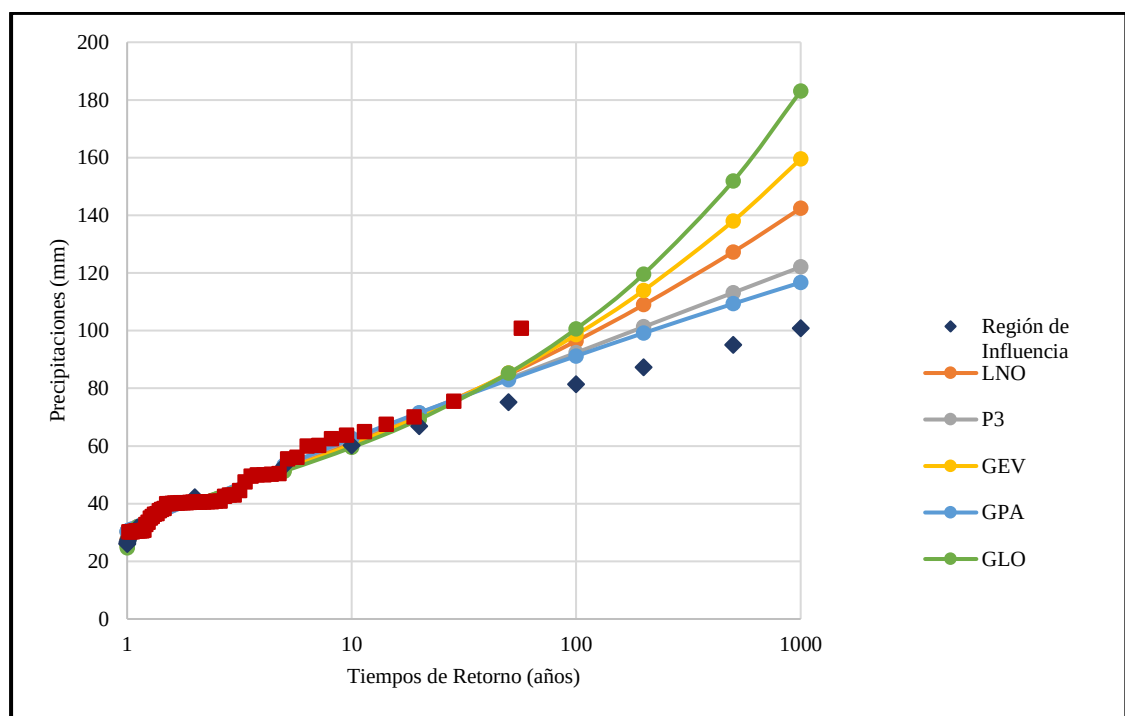


Nota. En Hacienda Pucará, las curvas muestran comportamientos distintos en tiempos de retorno altos, por lo que la selección de la distribución debe sustentarse con criterios de bondad de ajuste y consistencia hidrológica.

La Figura 67 muestra un incremento progresivo de la precipitación con el tiempo de retorno y una razonable coherencia entre la estimación regional de la Región de Influencia y la tendencia de la estación en la zona de frecuencias comunes. Para los tiempos de retorno mayores, se observa separación entre distribuciones, indicando incertidumbre asociada al modelo probabilístico en el extremo. En este sentido, la figura sustenta que el ROI es útil para estimar cuantiles de diseño en estaciones con limitada información extrema, siempre que la selección de estaciones vecinas y la distribución adoptada se sustenten con criterios estadísticos y consistencia física.

Figura 67

Precipitación de diseño para un periodo de retorno de 100 años y distribuciones de frecuencias para la estación Hacienda Shumaya (Grupo 5)

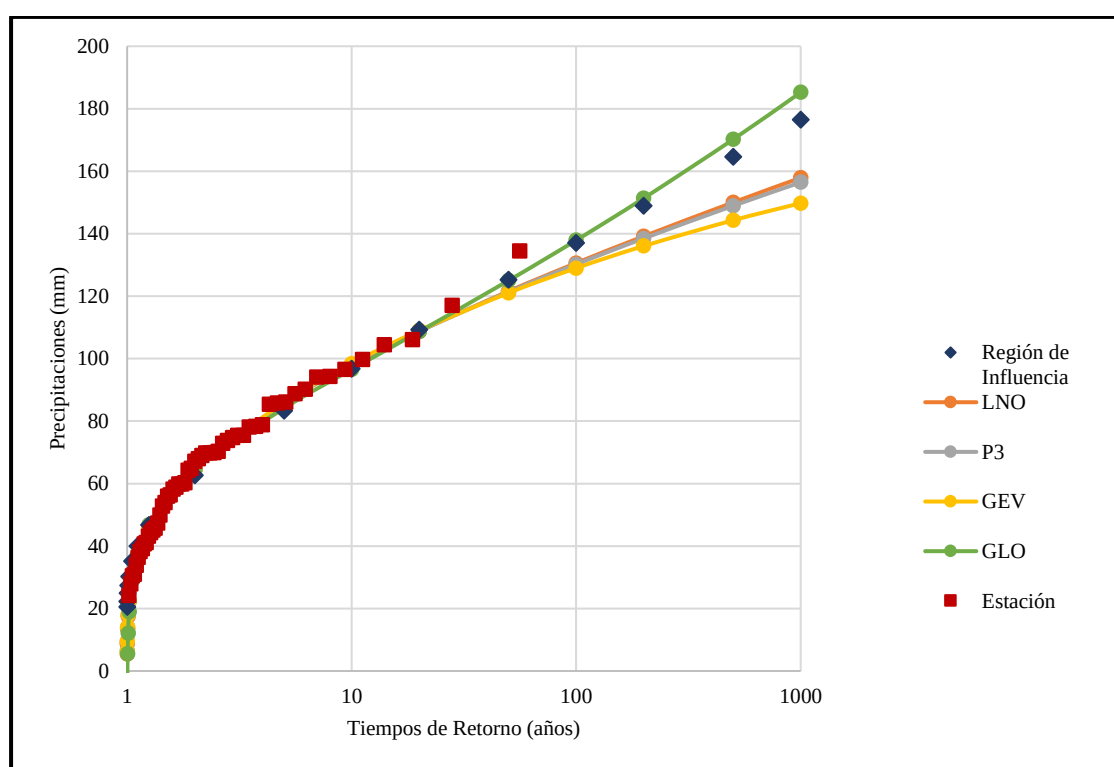


Nota. La Región de Influencia mejora la estabilidad de la estimación al incorporar información regional para Hacienda Shumaya; la incertidumbre en eventos raros se evidencia por la divergencia entre distribuciones en TR altos.

En la Figura 68 se aprecia que la Región de Influencia reproduce la tendencia general de la estación Huarmaca, particularmente en recurrencias bajas y medias, donde la información regional mejora la estabilidad del ajuste. Para recurrencias altas, la separación entre distribuciones vuelve a ser evidente, lo cual confirma que el cuantil extremo es sensible a la forma de la cola. De este modo, la figura refuerza el valor del ROI como enfoque local-regional, ya que combina coherencia espacial con capacidad de evaluación de incertidumbre por distribución, permitiendo definir una precipitación de diseño más sustentada para $TR = 100$ años.

Figura 68

Precipitación de diseño para un periodo de retorno de 100 años y distribuciones de frecuencias para la estación Huarmaca (Grupo 6)



Nota. Para Huarmaca, las curvas reflejan el comportamiento esperado de crecimiento con el tiempo de retorno; el cuantil de diseño se determina para $TR = 100$ años según la distribución seleccionada por desempeño estadístico en la Región de Influencia.

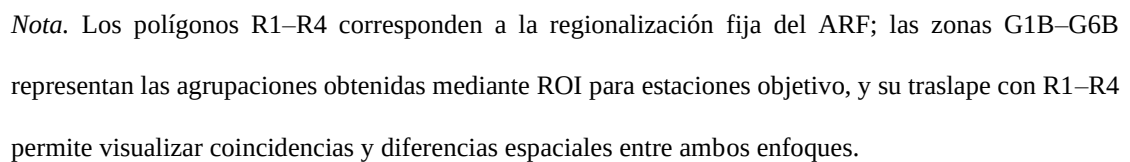
4.1.2.3. Análisis comparativo ARF vs ROI.

En este apartado se presenta la comparación de los cuantiles de precipitación de diseño estimados mediante el Análisis Regional de Frecuencia (ARF) y el enfoque de Región de Influencia (ROI/RI), contrastándolos con los cuantiles obtenidos a partir del ajuste local mediante L-momentos. La comparación se desarrolla en dos niveles: primero, una revisión espacial de la correspondencia entre las regiones homogéneas del ARF y las agrupaciones definidas por ROI; y, posteriormente, una comparación cuantitativa y gráfica de los valores estimados para distintos periodos de retorno en estaciones representativas.

a) Comparación espacial de la regionalización ARF y las regiones de influencia (ROI)

La Figura 69 muestra la superposición de dos esquemas de regionalización aplicados al análisis de frecuencia: (i) las regiones homogéneas del Análisis Regional de Frecuencia (ARF) delimitadas como R1, R2, R3 y R4, y (ii) los grupos definidos por el enfoque de Región de Influencia (ROI/RI) representados como G1B, G2B, G3B, G4B, G5B y G6B, además de la ubicación de las estaciones pluviométricas. En términos generales, la figura evidencia que el ARF conforma regiones fijas y continuas para toda el área de estudio, mientras que el ROI define agrupaciones específicas asociadas a estaciones objetivo, las cuales pueden traslaparse o intersectar más de una región ARF. Este contraste espacial respalda que ambos métodos comparten el objetivo de transferir información regional para estimar cuantiles de diseño, pero difieren en su lógica: el ARF prioriza la homogeneidad regional global, mientras que el ROI prioriza la similitud local respecto a la estación de interés, pudiendo integrar estaciones ubicadas en sectores contiguos o transicionales.

Comparación espacial entre regiones ARF y grupos ROI (RI)



En conjunto, el análisis comparativo evidencia tres comportamientos: (i) estaciones donde ARF es más conservador que ROI y local (Sondorillo, y en cola El Limón), (ii) estaciones donde ROI es más conservador que ARF y local (Hacienda Pucará), y (iii) estaciones donde ambos métodos son consistentes y las diferencias son marginales (Querocotillo y, en general, Huarmaca). En este marco, el ROI se interpreta como un enfoque sensible a la selección de estaciones influyentes, mientras que el ARF refleja la estructura regional fija definida por homogeneidad.

La Tabla 40 compara cuantiles de Sondorillo estimados por L-momentos locales, ARF y ROI. En todos los TR, el ARF tiende a entregar valores más altos que los otros métodos, especialmente en TR grandes (p. ej., TR 1 000: 76 mm en ARF frente a 70 mm en L-momentos y 72 mm en ROI). El ROI se mantiene como una alternativa intermedia, cercana a los valores locales, lo cual sugiere que incorpora información regional sin alejarse excesivamente del patrón de la estación.

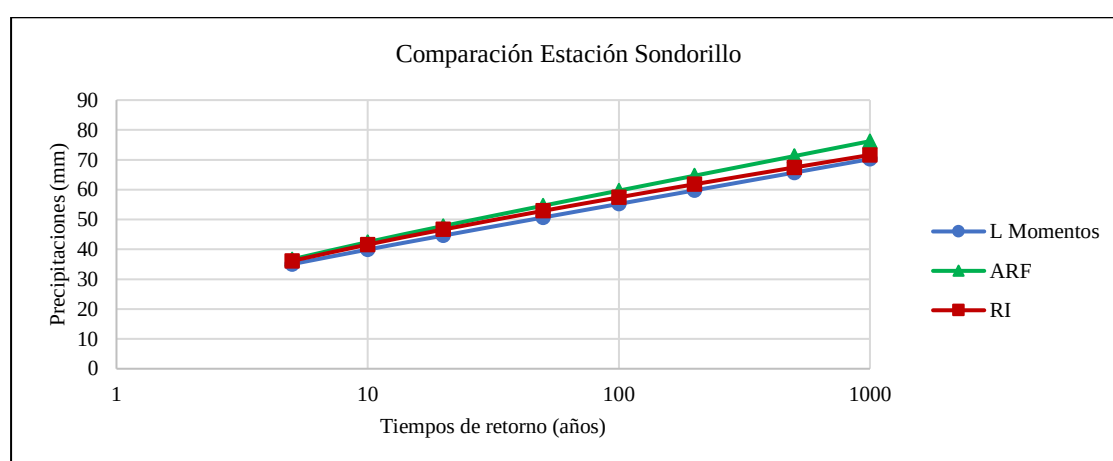
Tabla 40

Cuantiles registrados y estimados (mm) para diferentes periodos de retorno y métodos para la Estación Sondorillo

Tiempo de retorno (TR)	L Momentos	Análisis Regional de Frecuencias Hosking y Wallis	ARF Región de Influencia (RI)
5	35	37	36
10	40	42	42
20	45	48	47
50	51	55	53
100	55	60	57
200	60	65	62
500	66	71	67
1000	70	76	72

Figura 70

Comparación gráfica de tendencias – Estación Sondorillo



Nota. Se confirma gráficamente el patrón de la tabla: las tres curvas crecen con el TR, pero ARF presenta la mayor pendiente en la cola, mientras que ROI y L-momentos quedan más próximos entre sí. En consecuencia, para diseño con TR altos, el ARF resulta más conservador en esta estación.

La Tabla 41 evidencia que, para TR bajos y medios, los valores de ARF y ROI son menores que los estimados por L-momentos locales (p. ej., TR 50: 79–78 mm frente a 86 mm). Sin embargo, para TR altos el ARF supera a los otros métodos (TR 1 000: 111 mm en ARF frente a 102 mm local y 106 mm en ROI). Este comportamiento indica un cambio de dominancia en la cola, donde el ajuste regional proyecta una mayor severidad de extremos.

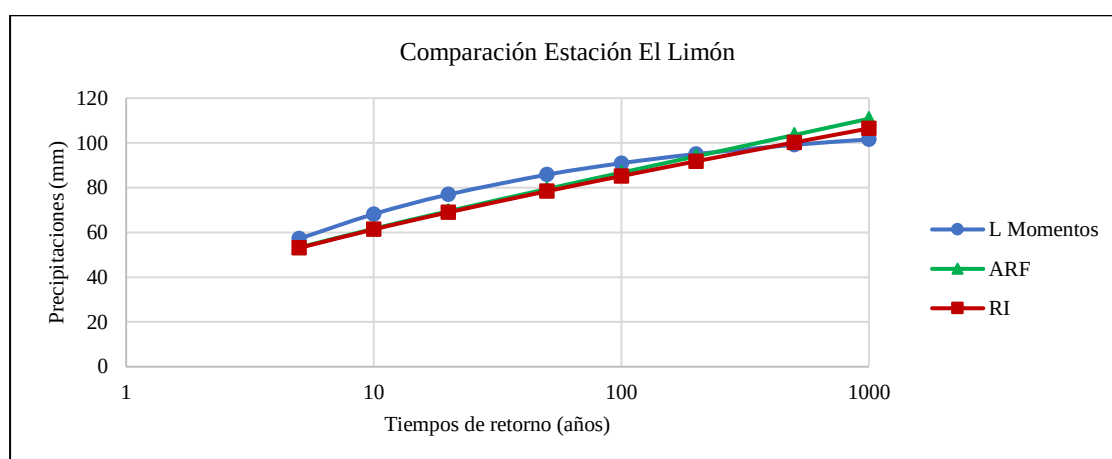
Tabla 41

Cuantiles registrados y estimados (mm) para diferentes periodos de retorno y métodos para la Estación El Limón

Tiempo de retorno (TR)	L Momentos	Análisis Regional de Frecuencias Hosking y Wallis	ARF Región de Influencia (RI)
5	57	53	53
10	68	62	61
20	77	69	69
50	86	79	78
100	91	87	85
200	95	94	92
500	99	104	100
1000	102	111	106

Figura 71

Comparación gráfica de tendencias – Estación El Limón



Nota. Muestra la convergencia relativa en TR bajos y la separación progresiva hacia TR altos, donde ARF se ubica por encima. El ROI acompaña la tendencia regional, pero mantiene una proyección más moderada que ARF, actuando como una solución de compromiso entre lo local y lo regional.

La Tabla 42 presenta alta consistencia entre métodos para todos los TR. Las diferencias son mínimas incluso en TR grandes (TR 1 000: 97 mm local, 99 mm ARF y 98 mm ROI), lo que sugiere estabilidad del ajuste y coherencia entre la estación y su información regional asociada.

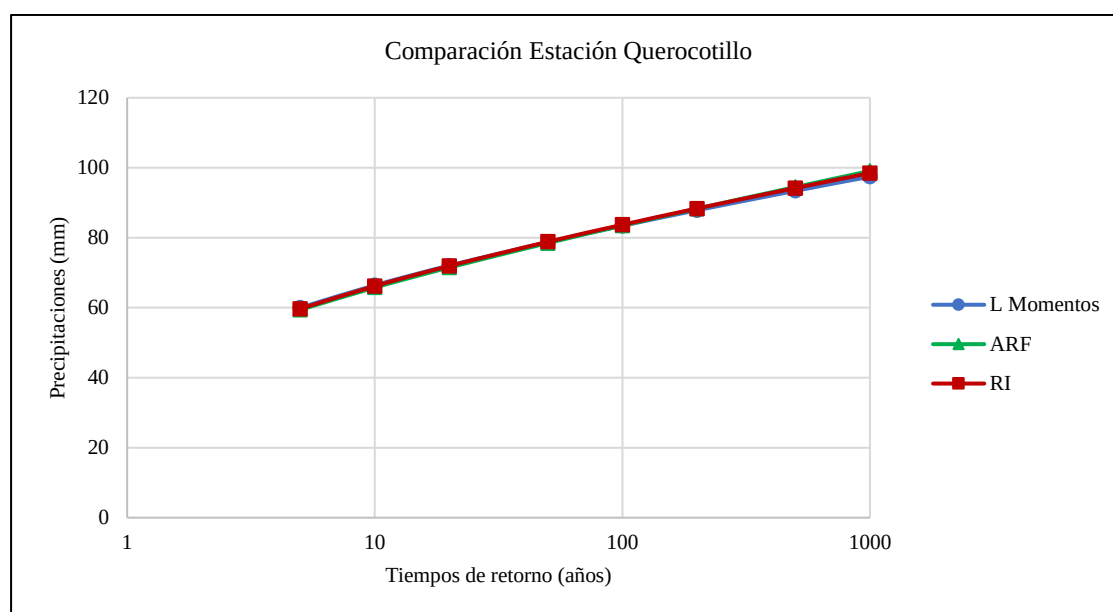
Tabla 42

Cuantiles registrados y estimados (mm) para diferentes periodos de retorno y métodos para la Estación Querocotillo

Tiempo de retorno (TR)	L Momentos	Análisis Regional de Frecuencias Hosking y Wallis	ARF Región de Influencia (RI)
5	60	59	60
10	66	66	66
20	72	71	72
50	79	78	79
100	83	83	84
200	88	88	88
500	93	94	94
1000	97	99	98

Figura 72

Comparación gráfica de tendencias – Estación Querocotillo



Nota. Ratifica que las curvas prácticamente se superponen en todo el rango de TR, por lo que la elección entre ARF y ROI no introduce cambios significativos en los cuantiles de diseño para esta estación.

La Tabla 43 muestra una diferencia marcada del ROI frente a los otros métodos, con cuantiles sistemáticamente mayores, sobre todo para TR altos (TR 1 000: 106 mm en ROI frente a 95 mm en ARF y 93 mm local). Este resultado indica que la región de influencia definida para esta estación induce una extrapolación más conservadora, lo cual debe discutirse como efecto del grupo seleccionado y de la sensibilidad a la distribución en la cola.

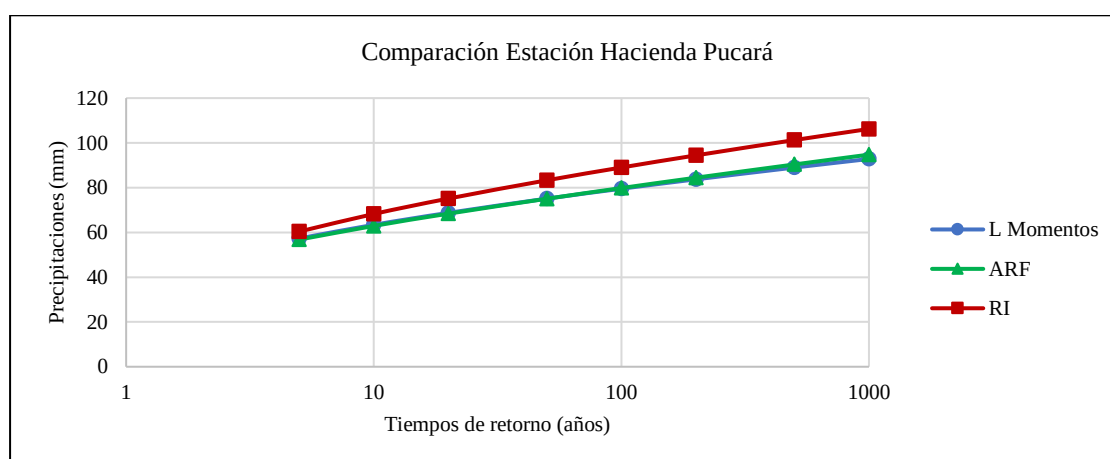
Tabla 43

Cuantiles registrados y estimados (mm) para diferentes periodos de retorno y métodos para la Estación Hacienda Pucará

Tiempo de retorno (TR)	L Momentos	Análisis Regional de Frecuencias Hosking y Wallis	ARF Región de Influencia (RI)
5	57	57	60
10	63	63	68
20	69	68	75
50	75	75	83
100	80	80	89
200	84	84	94
500	89	90	101
1000	93	95	106

Figura 73

Comparación gráfica de tendencias – Estación Hacienda Pucará



Nota. Evidencia una separación clara entre ROI y los demás métodos conforme aumenta el TR. En términos de interpretación, el ROI proyecta una cola más pesada o un crecimiento más acelerado de extremos, por lo que su uso en diseño implicaría márgenes de seguridad mayores para eventos raros.

La Tabla 44 muestra que ARF y L-momentos locales prácticamente coinciden en TR altos (TR 1 000: 117 mm en ambos), mientras que ROI entrega valores menores (TR 1 000: 101 mm). En esta estación, el ROI reduce la magnitud de cuantiles extremos respecto a los otros enfoques, lo cual sugiere que la región de influencia seleccionada conduce a una extrapolación más moderada.

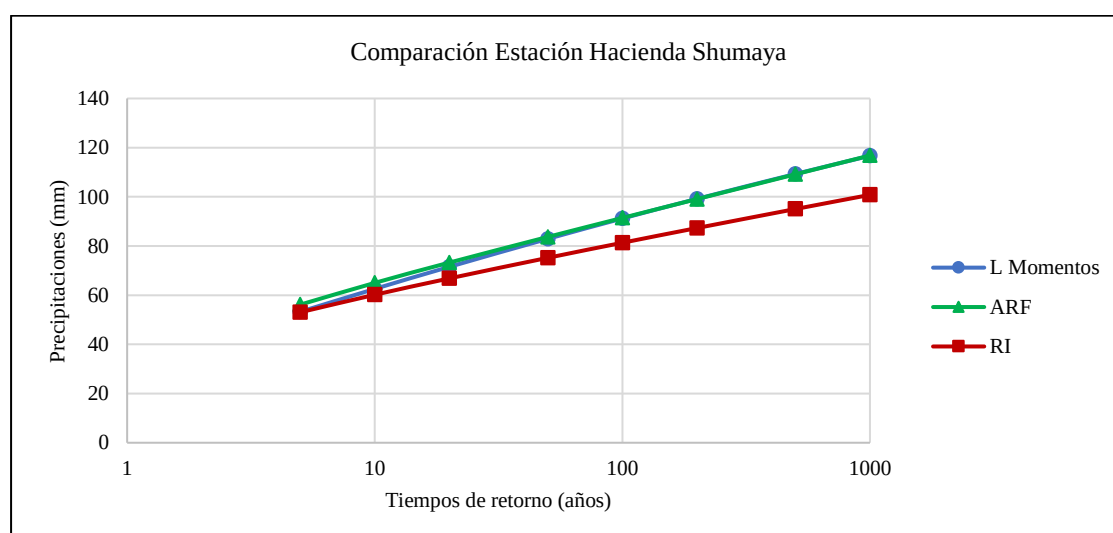
Tabla 44

Cuantiles registrados y estimados (mm) para diferentes periodos de retorno y métodos para la Estación Hacienda Shumaya

Tiempo de retorno (TR)	L Momentos	Análisis Regional de Frecuencias Hosking y Wallis	ARF Región de Influencia (RI)
5	53	56	53
10	63	65	60
20	72	73	67
50	83	84	75
100	91	91	81
200	99	99	87
500	109	109	95
1000	117	117	101

Figura 74

Comparación gráfica de tendencias – Estación Hacienda Shumaya



Nota. Confirma que ROI se separa hacia abajo en TR medios y altos, mientras ARF y el método local mantienen una trayectoria similar. Por tanto, para Hacienda Shumaya, el ROI implica un diseño menos conservador que el ARF.

La Tabla 45 evidencia cercanía entre los tres métodos, con diferencias pequeñas en TR bajos y una leve mayor proyección de ARF y ROI en TR altos (TR 1 000: 174–176 mm frente a 168 mm local). Esto sugiere que la componente regional incrementa ligeramente la estimación de extremos sin modificar sustancialmente la tendencia general.

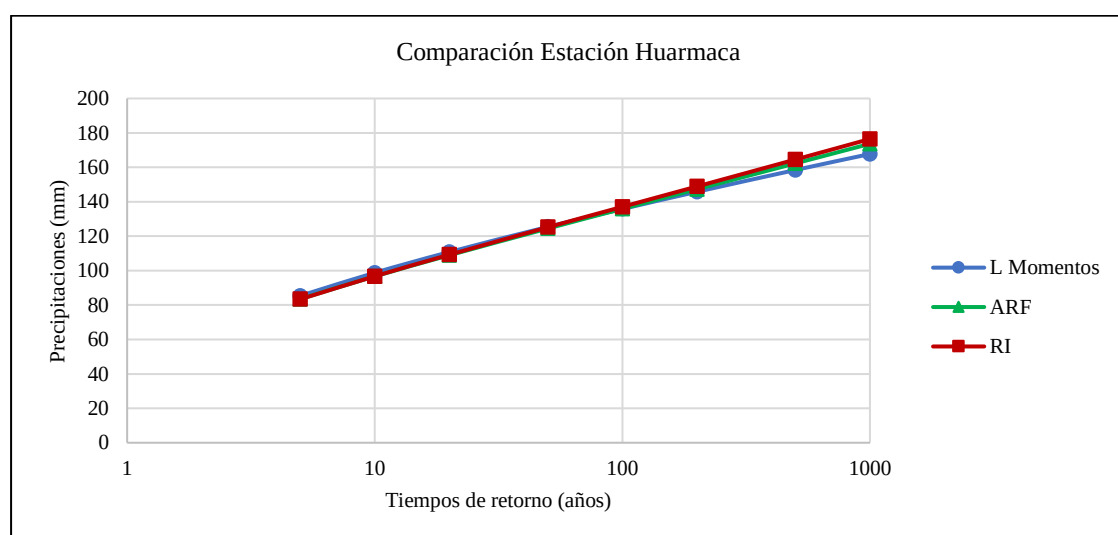
Tabla 45

Cuantiles registrados y estimados (mm) para diferentes periodos de retorno y métodos para la Estación Huarmaca

Tiempo de retorno (TR)	L Momentos	Análisis Regional de Frecuencias Hosking y Wallis	ARF Región de Influencia (RI)
5	85	84	83
10	99	97	97
20	111	109	109
50	125	124	125
100	136	136	137
200	146	147	149
500	158	162	165
1000	168	174	176

Figura 75

Comparación gráfica de tendencias – Estación Huarmaca



Nota. Muestra curvas con crecimiento similar, con una separación moderada en TR altos donde ARF y ROI quedan por encima del ajuste local. En consecuencia, para esta estación, ARF y ROI aportan un ajuste regional que refuerza la estimación de eventos raros sin introducir discrepancias grandes.

4.1.3. Mapas regionales de PMDA regionalizadas y caracterizadas espacialmente en regiones homogéneas para la cuenca del río Chamaya

Las tablas (Tabla 46 a Tabla 50) presentan la cobertura espacial de las regiones homogéneas obtenidas mediante ARF y su representación cartográfica en la cuenca del río Chamaya. Debe distinguirse entre la composición estadística del análisis regional y la cantidad de estaciones con área efectiva, tras el recorte de los polígonos de Thiessen. El ARF definitivo estuvo conformado por 24 estaciones: 8 en la Región 1, 8 en la Región 2, 4 en la Región 3 y 4 en la Región 4. Al interceptar los polígonos de influencia con el límite de la cuenca, 22 estaciones aportaron superficie efectiva. Tocmoche y Tabaconas se mantuvieron en sus respectivas regiones para el análisis estadístico, aunque sus polígonos no representaron área dentro de Chamaya. En términos territoriales, la Región 1 abarcó 3 571.59 km²; la Región 2, 2 837.21 km²; la Región 3, 420.44 km²; y la Región 4, 1 307.95 km². La suma alcanzó 8 137.19 km², superficie oficial adoptada para el estudio.

Tabla 46

Resumen de cobertura espacial y número de estaciones por región homogénea (ARF)

Región homogénea	N.º de estaciones en el ARF	N.º de estaciones con área efectiva dentro de la cuenca	Área total (km ²)	Participación del área (%)	Longitud (°) mín.–máx.	Latitud (°) mín.–máx.	Estaciones representativas
Región 1	8	8	3 571.59	43.89	–79.53 a –79.07	–5.97 a –5.34	Colasay, El Limón, Sondorillo
Región 2	8	8	2 837.21	34.87	–79.36 a –78.65	–6.58 a –6.04	Querocotillo, Hacienda Pucará, Cutervo
Región 3	4	3	420.44	5.17	–79.51 a –78.73	–6.67 a –5.84	Porculla, Llama, Chugur
Región 4	4	3	1 307.95	16.07	–79.45 a –78.90	–5.67 a –5.10	Chontalí, La Cascarilla, Salala
Total	24	22	8 137.19	100.00	—	—	—

Nota. El ARF final comprendió 24 estaciones distribuidas en cuatro regiones homogéneas. La cobertura cartográfica se calculó con 22 estaciones, cuyos polígonos de Thiessen presentaron un área efectiva dentro de la cuenca tras el recorte espacial. Tocmoche y Tabaconas permanecieron en el análisis regional,

pero no aportaron superficie al polígono de Chamaya. Las áreas regionales se calcularon con geometrías del SIG sin redondear; por ello, pueden presentarse diferencias de 0.01 km² al sumar los valores individuales mostrados con dos decimales.

Tabla 47

Estaciones y área de influencia por Región homogénea 1 definida mediante ARF

Región	Id	Estación	Longitud	Latitud	Área efectiva dentro de la cuenca (km ²)
1	2	Huarmaca	-79.53	-5.57	237.03
1	5	Sondorillo	-79.43	-5.34	437.34
1	3	Tuluze	-79.35	-5.48	319.53
1	7	El Limón	-79.32	-5.92	532.18
1	4	Hacienda Shumaya	-79.36	-5.38	132.19
1	8	Sallique	-79.32	-5.65	421.15
1	25	San Felipe	-79.32	-5.77	422.94
1	24	Colasay	-79.07	-5.97	1069.23
Total regional					3 571.59

Tabla 48

Estaciones y área de influencia por Región homogénea 2 definida mediante ARF

Región	Id	Estación	Longitud	Latitud	Área efectiva dentro de la cuenca (km ²)
2	20	Cueva Blanca	-79.36	-6.16	312.11
2	9	Hacienda Pucará	-79.13	-6.04	492.58
2	10	Querocotillo	-79.04	-6.27	939.20
2	11	Cutervo	-78.8	-6.38	408.00
2	12	Cochabamba	-78.89	-6.46	273.56
2	16	Chancay Baños	-78.87	-6.58	7.99
2	22	Chota	-78.65	-6.55	235.67
2	17	Chotano Lajas	-78.75	-6.56	168.09
Total regional					2 837.21

Nota. El área total de la Región 2 asciende a 2 837.21 km². La suma de las áreas individuales mostradas, con dos decimales, es de 2 837.20 km² debido al redondeo de los valores. El total regional se calculó a partir de las geometrías del SIG, sin redondear.

Tabla 49*Estaciones y área de influencia por Región homogénea 3 definida mediante ARF*

Región	Id	Estación	Longitud	Latitud	Área dentro de la cuenca (km ²)	Observación
3	1	Porculla	-79.51	-5.84	305.67	Polígono con área efectiva dentro de Chamaya
3	13	Llama	-79.12	-6.51	74.23	Polígono con área efectiva dentro de Chamaya
3	18	Tocmoche	-79.36	-6.41	0.00	Estación empleada en el ARF; sin área efectiva después del recorte
3	21	Chugur	-78.73	-6.67	40.54	Polígono con área efectiva dentro de Chamaya
Total regional					420.44	

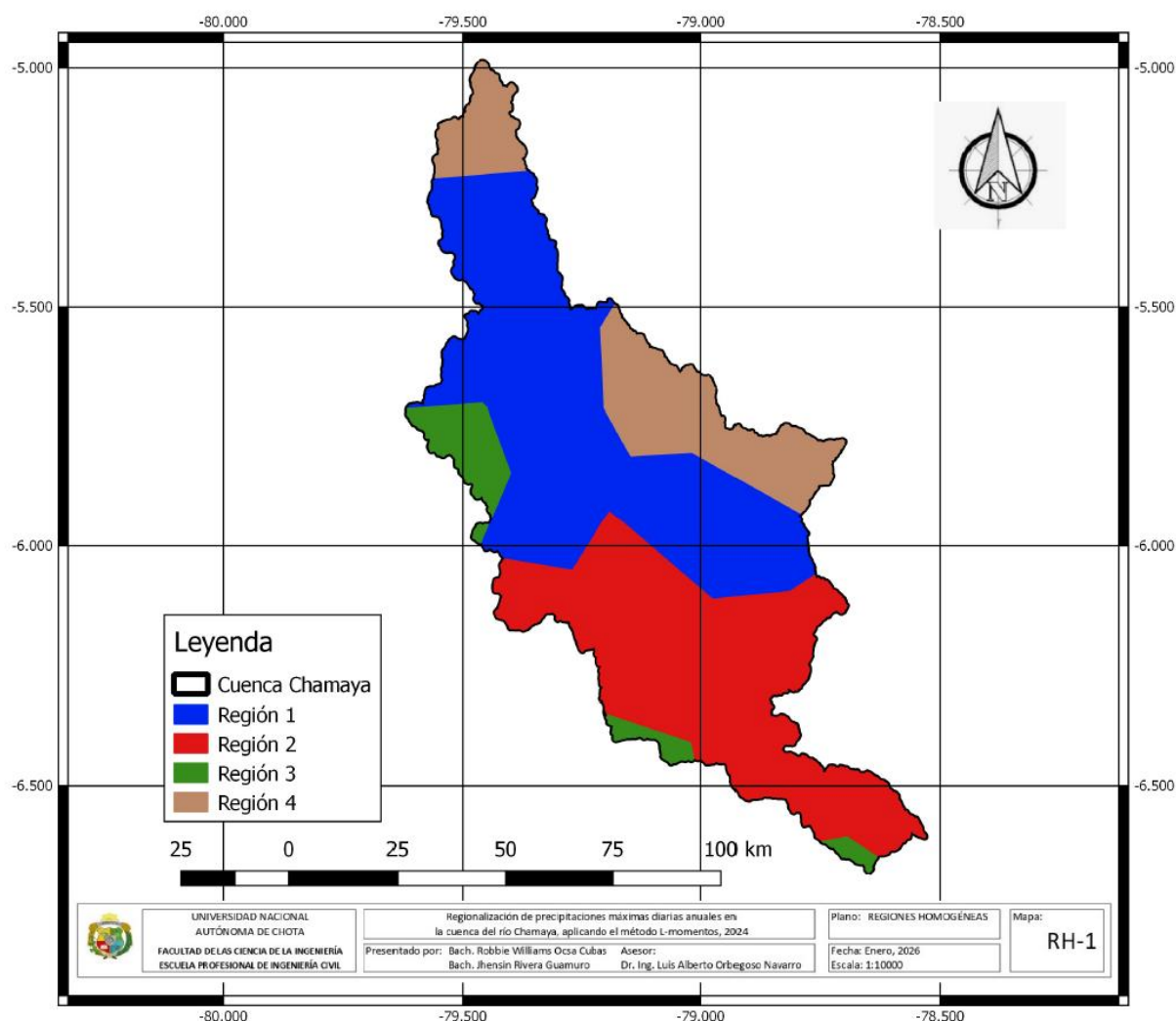
Tabla 50*Estaciones y área de influencia por Región homogénea 4 definida mediante ARF*

Región	Id	Estación	Longitud	Latitud	Área efectiva dentro de la cuenca (km ²)	Observación
4	23	Salala	-79.45	-5.10	340.39	Polígono con área efectiva dentro de Chamaya
4	15	La Cascarilla	-78.90	-5.67	444.10	Polígono con área efectiva dentro de Chamaya
4	14	Chontalí	-79.09	-5.64	523.46	Polígono con área efectiva dentro de Chamaya
4	6	Tabaconas	-79.23	-5.33	0.00	Estación empleada en el ARF; sin área efectiva después del recorte
Total regional					1 307.95	

La Figura 76 muestra la zonificación de la cuenca del río Chamaya en cuatro regiones homogéneas obtenidas mediante el Análisis Regional de Frecuencias (ARF). En términos de cobertura, se observa que las regiones 1 y 2 ocupan la mayor parte del territorio de la cuenca, mientras que las regiones 3 y 4 corresponden a áreas de menor extensión relativa, asociadas a sectores específicos del ámbito de estudio. Esta regionalización constituye la base para transferir información hidrológica y construir cuantiles de precipitación consistentes a escala de cuenca.

Figura 76

Regiones homogéneas identificadas mediante ARF en la cuenca del río Chamaya



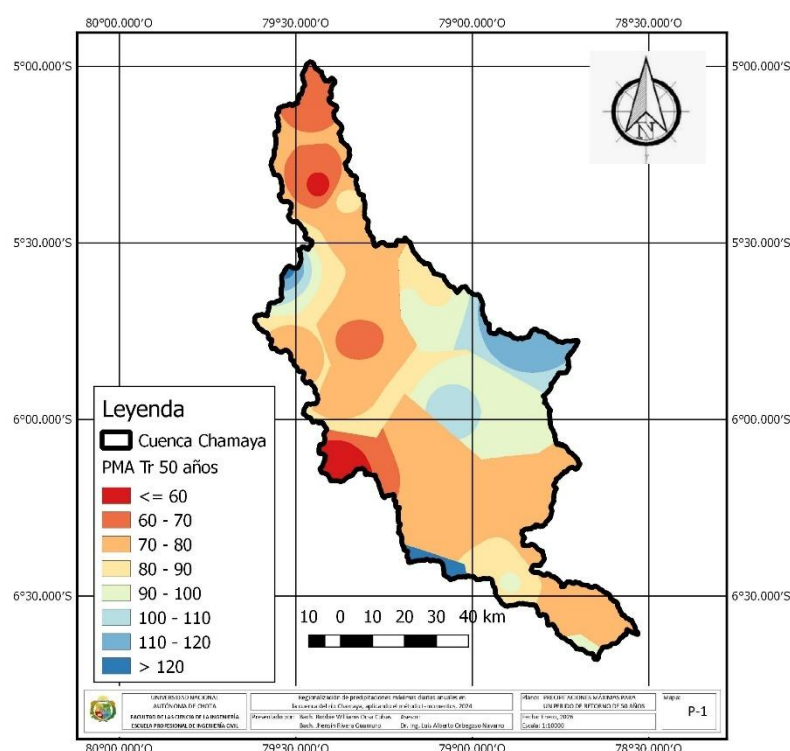
Nota. La delimitación cartográfica se obtuvo al interceptar los polígonos de Thiessen con el límite de Chamaya. Las estaciones Tocmoche y Tabaconas se conservaron en el ARF por su aporte estadístico regional, aunque no generaron área efectiva en la cuenca tras el recorte espacial.

La Figura 77 presenta la distribución espacial de la precipitación máxima diaria anual (PMDA) regionalizada para un periodo de retorno de 50 años, obtenida a partir del ARF. Los valores se agrupan en clases que van de ≤ 60 mm a > 120 mm, lo que evidencia un marcado gradiente espacial dentro de la cuenca. En general, se distinguen sectores con valores moderados en la mayor parte del territorio y núcleos con valores más altos concentrados en zonas específicas, lo que refleja la influencia del patrón

regional definido por las regiones homogéneas y la variabilidad espacial capturada por la regionalización. Este mapa permite identificar áreas con mayor exigencia hidrológica para el diseño bajo TR intermedios.

Figura 77

Mapa de PMDA regionalizado para $T=50$ años (mm).



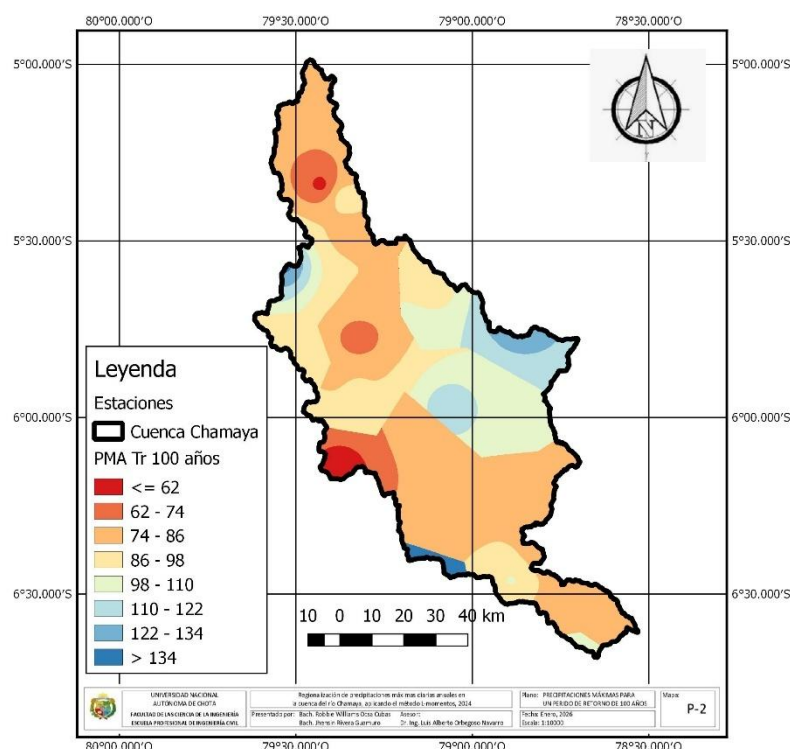
Nota. Intervalos en mm; TR en años. La superficie representa la PMDA regionalizada obtenida con ARF y posteriormente espacializada en SIG.

La Figura 78 muestra la PMDA regionalizada para un periodo de retorno de 100 años, con rangos que van desde ≤ 62 mm hasta > 134 mm. En comparación con TR = 50 años, se aprecia un incremento generalizado de la magnitud de los cuantiles y una mayor diferenciación entre zonas de valores bajos, medios y altos. La distribución espacial mantiene el patrón general observado en TR menores, pero con una intensificación de las clases superiores, lo cual es coherente con el aumento del evento

extremo asociado a mayores periodos de retorno. Este resultado respalda el uso del ARF para generar estimaciones espacialmente consistentes en escenarios de diseño más exigentes.

Figura 78

Mapa de PMDA regionalizada para $T=100$ años (mm).



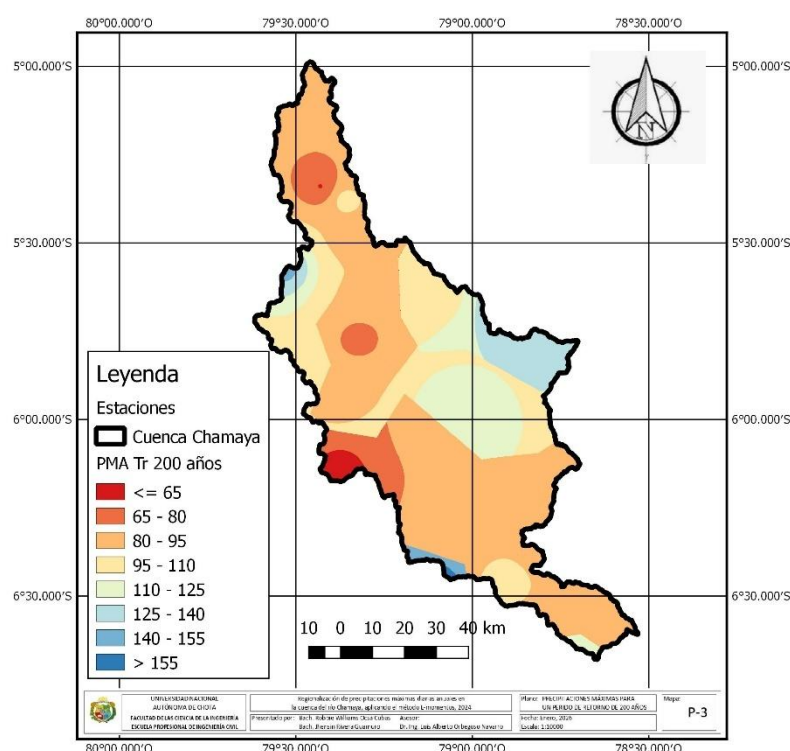
Nota. La ampliación de rangos respecto a TR = 50 años refleja el crecimiento de la PMDA para eventos menos frecuentes.

En la Figura 79 se presenta la PMDA regionalizada para TR = 200 años, con clases que varían desde ≤ 65 mm hasta > 155 mm. El mapa evidencia un aumento adicional de los cuantiles respecto a TR = 100 años y una expansión de las áreas asociadas a categorías altas, lo que incrementa el contraste espacial dentro de la cuenca. Este comportamiento es consistente con el carácter extremo del evento, donde las diferencias regionales se vuelven más relevantes para el diseño hidráulico e

hidrológico. En este marco, el ARF permite preservar la coherencia regional, reduciendo la inestabilidad que suele aparecer en ajustes estrictamente locales para periodos de retorno altos.

Figura 79

Mapa de PMDA regionalizada para $T=200$ años (mm).



Nota. TR altos tienden a amplificar la separación entre zonas de baja y alta PMDA; el ARF estabiliza la estimación al incorporar información regional.

La Figura 80 corresponde a la PMDA regionalizada para $TR = 500$ años, con rangos desde ≤ 75 mm hasta > 165 mm. Se observa el mayor nivel de exigencia hidrológica dentro de la serie de mapas, con un incremento notable de las clases superiores y una diferenciación espacial más marcada. Este patrón confirma que, a medida que aumenta el periodo de retorno, los cuantiles regionalizados crecen y las zonas con mayor propensión a eventos extremos se hacen más evidentes dentro de la

Figura 81

Mapa final de regiones homogéneas por ROI

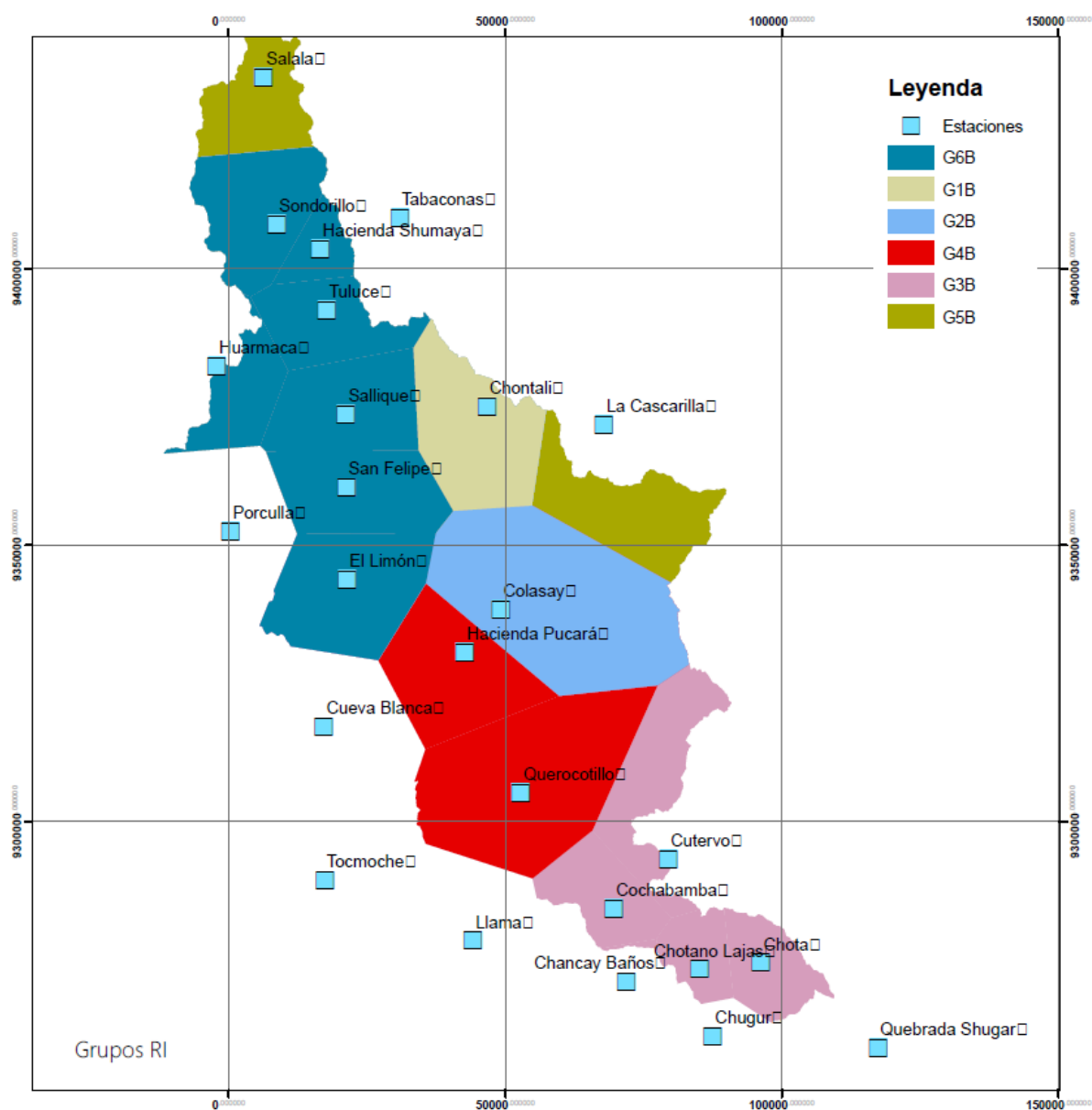
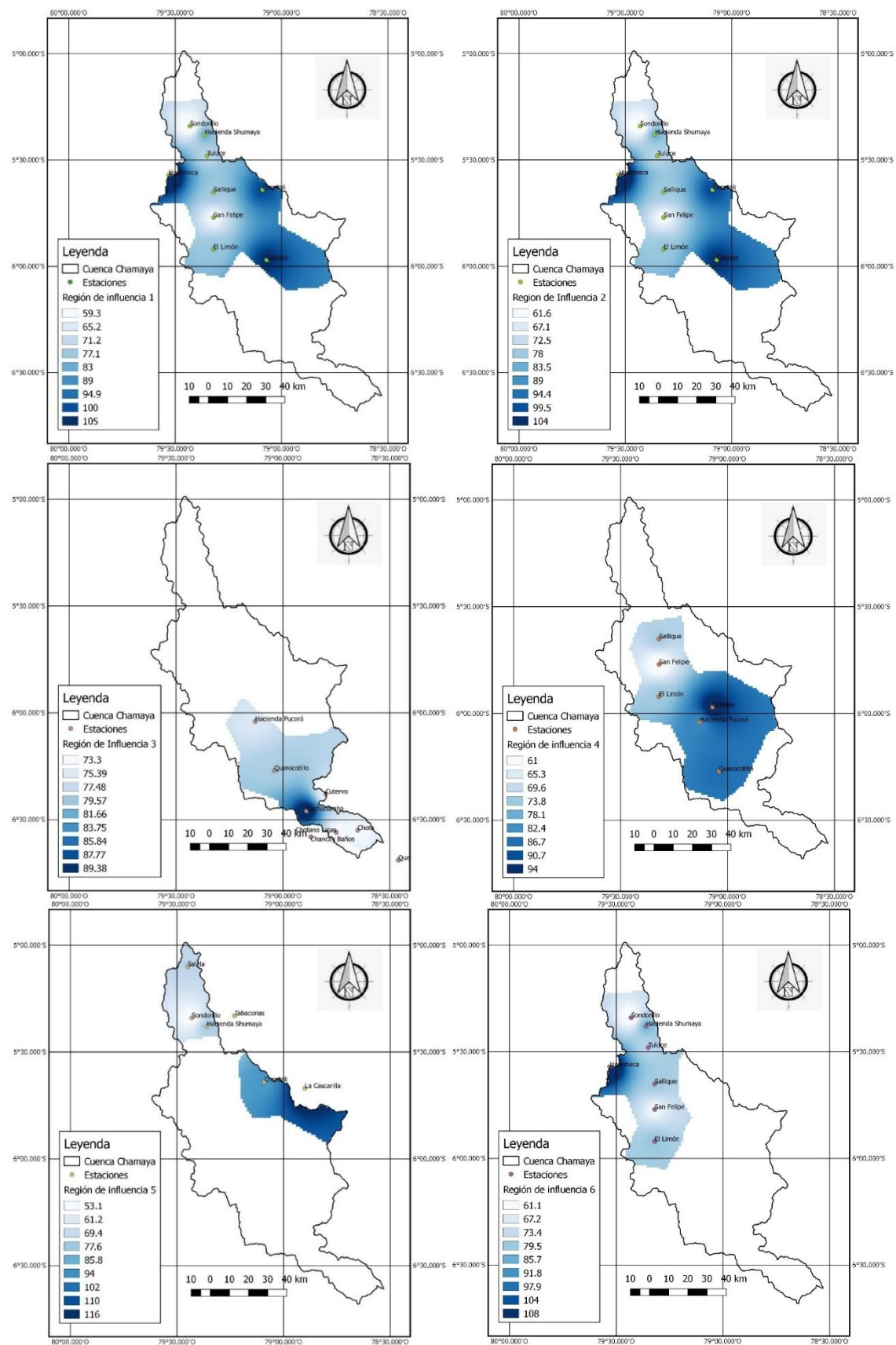


Figura 82

Región de influencia 1, 2, 3, 4, 5 y 6



4.1.4. Estrategias para mitigar la insuficiencia de datos hidrológicos en zonas poco instrumentadas, empleando la regionalización con L-momentos

La Tabla 51 evidenció una cobertura espacial heterogénea en la cuenca del río Chamaya al evaluar el número de estaciones y el área representada por cada región homogénea. La Región 1 concentró el 43.89% del área total, con 8 estaciones que aportaron superficie efectiva, y alcanzó una densidad de 2.24 est./1 000 km². La Región 2 representó el 34.87% del área y presentó la mayor densidad relativa entre las regiones extensas: 2.82 est./1 000 km². La Región 3 alcanzó la mayor densidad efectiva, de 7.14 est./1 000 km², debido a su menor extensión territorial. La Región 4 mantuvo una densidad media de 2.29 est./1 000 km² y un área de 435.98 km² por estación efectiva, lo que incrementa la incertidumbre espacial en las zonas de transición. En conjunto, los indicadores confirmaron la utilidad de aplicar estrategias de transferencia regional mediante ARF y ROI en áreas extensas con instrumentación limitada.

La Tabla 51 describe la cobertura espacial efectiva de las regiones homogéneas. La densidad se calculó a partir de las 22 estaciones cuyos polígonos de Thiessen aportaron superficie tras el recorte por el límite de Chamaya. Este indicador no reemplaza la composición estadística del ARF, que se basó en 24 estaciones. La distinción resulta necesaria en las regiones 3 y 4: ambas fueron verificadas con 4 estaciones, aunque solo 3 por región representaron áreas efectivas dentro del polígono de estudio.

Tabla 51

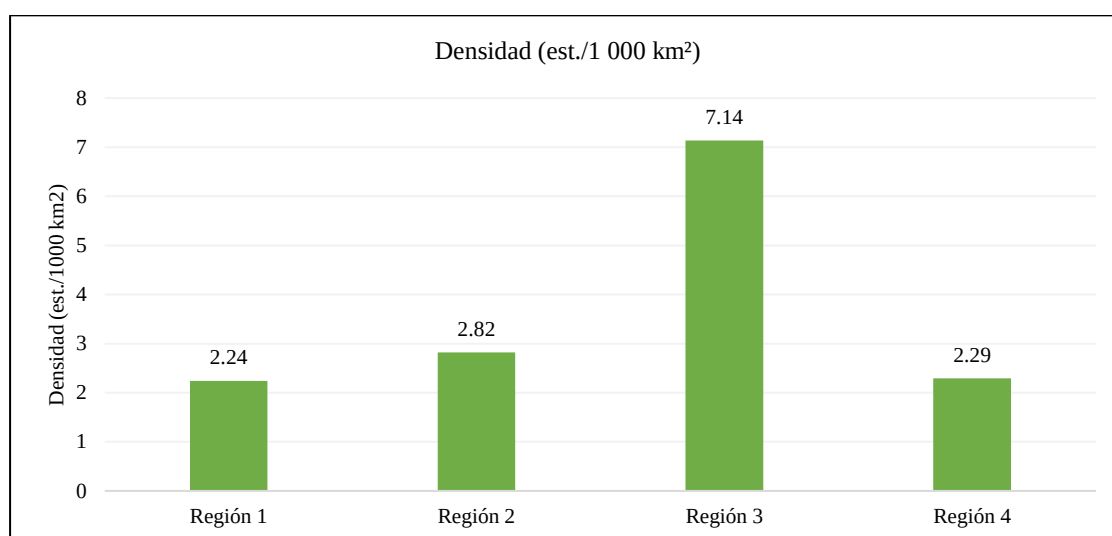
Indicadores de cobertura espacial e instrumentación por región homogénea (ARF)

Región homogénea	Área (km ²)	% del área total	N.º de estaciones en el ARF	N.º de estaciones con área efectiva	% de estaciones con área efectiva	Densidad espacial efectiva (est./1 000 km ²)	Área por estación efectiva (km ² /est.)	Diagnóstico de cobertura
Región 1	3 571.59	43.89	8	8	36.36	2.24	446.45	Media
Región 2	2 837.21	34.87	8	8	36.36	2.82	354.65	Media
Región 3	420.44	5.17	4	3	13.64	7.14	140.15	Alta
Región 4	1 307.95	16.07	4	3	13.64	2.29	435.98	Media
Total	8 137.19	100.00	24	22	100.00	2.70	369.87	—

Nota. La densidad espacial efectiva se calculó con las estaciones cuyos polígonos de Thiessen representaron superficie dentro de la cuenca. El número de estaciones del ARF se presenta por separado para no confundir la composición estadística de las regiones homogéneas con su cobertura cartográfica.

Figura 83

Densidad de estaciones por región homogénea (ARF)



Nota. La densidad de estaciones se concentró en la Región 3, mientras que las Regiones 1, 2 y 4 presentaron densidades similares, aunque significativamente menores. Este patrón reforzó que, aun con un número comparable de estaciones, las regiones de mayor extensión requieren mecanismos de regionalización para estimar PMDA en sectores no instrumentados y controlar la variabilidad espacial que no puede describirse solo mediante extrapolación local.

La Tabla 52 consolidó las estrategias que se obtuvieron a partir del enfoque de regionalización por L-momentos para enfrentar la escasez de datos en zonas poco instrumentadas. En específico, se establecieron dos rutas principales de transferencia de información: una basada en regiones homogéneas (ARF) y otra basada en regiones de influencia (ROI). Ambas permitieron transformar información puntual (estaciones) en estimaciones regionales consistentes, generando productos verificables como mapas PMDA por periodo de retorno, grupos de influencia y cuantiles transferibles. Complementariamente, se estructuraron estrategias de refuerzo de red y validación, con el fin de reducir incertidumbre en zonas extensas o de transición donde la instrumentación existente resulta limitada.

Tabla 52

Estrategias derivadas de la regionalización L-momentos para mitigar insuficiencia de datos (síntesis operativa)

Estrategia (resultado)	En qué se apoya	Qué problema mitiga	Producto verificable en resultados	Aplicación directa
Transferencia regional ARF	Regiones homogéneas + curva regional de crecimiento	Zonas sin estación dentro de una región	Mapas regionales PMDA por TR (ARF) + cuantiles regionales	Diseño en puntos no instrumentados dentro de región
Transferencia local ROI	Región “a medida” por estación/punto objetivo	Bordes entre regiones y transiciones espaciales	Mapas ROI por grupo + cuantiles por región de influencia	Estimación más estable en zonas de transición
Priorización de refuerzo de red	Densidad, área por estación y gradientes espaciales	Incertidumbre alta por baja representatividad	Ranking de zonas candidatas a instrumentar	Definición de ubicaciones objetivo
Validación cruzada regional	Comparación local vs ARF vs ROI	Sobre/subestimación regional	Métricas de error relativo por TR y por estación	Control de calidad de la regionalización
Integración de fuentes complementarias	Registros cortos + información regional	Series cortas o con vacíos	Ajuste regional con soporte de ARF/ROI	Mejoras para estaciones recientes

La Tabla 53 estableció una priorización técnica de refuerzo de la red pluviométrica considerando el área representada por estación y la densidad regional. Bajo estos criterios, las Regiones 1 y 4 mostraron prioridad alta, debido a que presentan valores superiores a 435 km² por estación, condición asociada a mayor incertidumbre espacial y a potenciales subrepresentaciones en zonas internas o de transición. La Región 2 presentó prioridad media–alta por su extensión y por la necesidad de capturar mejor los gradientes espaciales, mientras que la Región 3 mostró menor urgencia debido a su mayor densidad relativa.

Tabla 53

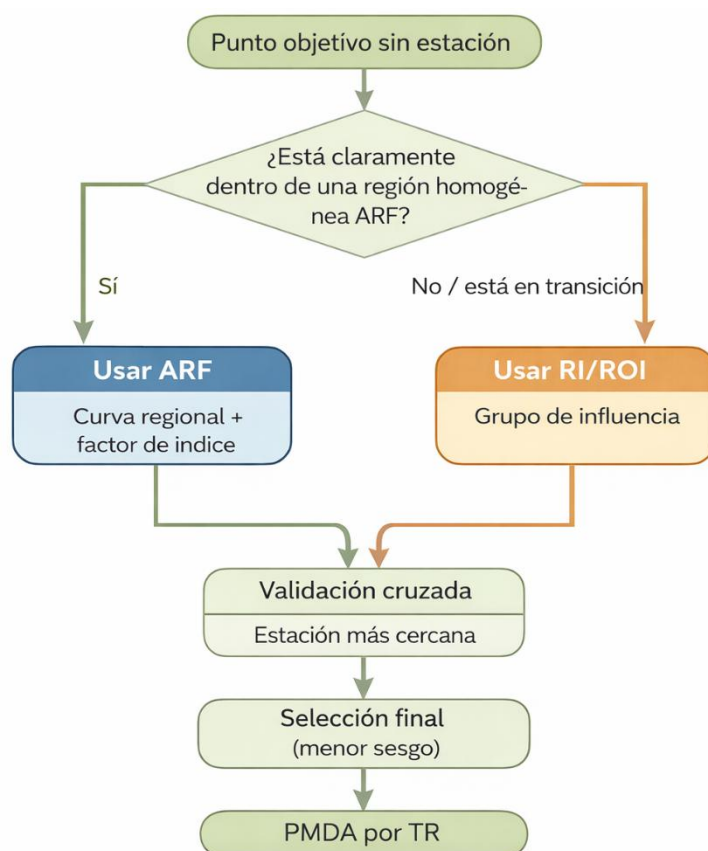
Priorización técnica de zonas para reforzar instrumentación (orientada a mitigación de vacíos)

Región (ARF)	Condición observada	Evidencia (indicador)	Prioridad de refuerzo	Criterio de ubicación sugerido
Región 1	Alta extensión con densidad media	446.45 km ² /est.	Alta	Sectores sin estación interna y bordes con Región 2 y 4
Región 2	Extensión alta con densidad media	354.65 km ² /est.	Media–alta	Zonas con gradiente espacial marcado y transición hacia Región 3
Región 3	Cobertura densa en área reducida	140.15 km ² /est.	Baja	Solo refuerzo puntual si existen vacíos altitudinales
Región 4	Densidad media y posible transición espacial	435.98 km ² /est.	Alta	Sectores de transición y áreas con mayor contraste espacial

La Figura 84 sintetizó el criterio operativo para seleccionar entre ARF y ROI en la estimación de PMDA en puntos no instrumentados. Cuando el punto objetivo se ubicó en el interior de una región homogénea, el ARF resultó adecuado por su estabilidad regional. En cambio, en zonas de transición o cercanas a límites regionales, el enfoque ROI permitió construir regiones adaptativas que reducen discontinuidades y aumentan coherencia espacial. El cierre mediante validación cruzada aseguró que la estrategia seleccionada mantenga consistencia con la evidencia observacional disponible.

Figura 84

Esquema de decisión ARF vs ROI para estimación en puntos no instrumentados



Nota. La selección ARF/ROI se realizó con criterio espacial (interior vs transición) y control posterior mediante validación cruzada por estación vecina.

4.2. Contrastación de hipótesis

Para la contrastación de hipótesis se utilizó el software Minitab 22, debido a que permitió organizar los valores regionales obtenidos en el ARF y ejecutar pruebas inferenciales con nivel de confianza del 95% ($\alpha = 0.05$). En este marco, se trabajó con los cuatro valores regionales de Dmax (máximo de discordancia por región) y H1 (heterogeneidad por región), los cuales se ingresaron como dos conjuntos de datos ($n = 4$). Luego, se verificó el supuesto de normalidad mediante la prueba de Anderson–Darling y, de este modo, se aplicó la prueba t de una muestra para evaluar si los promedios cumplían los umbrales establecidos por el ARF ($D_{\max} < 3$ y $H1 < 2$).

a) Hipótesis general

La aplicación del método de L-momentos permite regionalizar las precipitaciones máximas diarias anuales en la cuenca del río Chamaya mediante la delimitación de regiones homogéneas estadísticamente consistentes y la estimación de cuantiles de diseño para diferentes periodos de retorno.

b) Datos utilizados

Se utilizaron los valores máximos de discordancia por región (Dmax) y los valores de heterogeneidad (H1) de las regiones finales seleccionadas para la distribución regional y estimación de cuantiles.

Tabla 54

Máximos valores de discordancia y heterogeneidad en las regiones formadas

	R-1	R-2	R-3	R-4
Dmax	1.94	1.85	1.00	1.00
H1	1.74	1.15	1.43	0.92

c) Criterio de aceptación o rechazo

Se adoptó $\alpha = 0.05$. Por lo tanto, no se rechazó H_0 cuando $p > 0.05$, y se rechazó H_0 cuando $p \leq 0.05$. Para verificar el cumplimiento de los umbrales, se plantearon contrastes unilaterales:

- Para discordancia: $H_0: \mu(D_{\max}) = 3$ vs $H_1: \mu(D_{\max}) < 3$.
- Para heterogeneidad: $H_0: \mu(H_1) = 2$ vs $H_1: \mu(H_1) < 2$.

d) Prueba de normalidad (Anderson-Darling)

Los resultados indicaron que ambos conjuntos de datos presentaron comportamiento compatible con normalidad, ya que Dmax obtuvo $p = 0.080$ y H1 obtuvo $p = 0.831$ (ambos mayores que 0.05). Por consiguiente, se aplicó la prueba t de

una muestra como contraste inferencial complementario, interpretando sus resultados con cautela debido al reducido número de regiones evaluadas.

Figura 85

Prueba de normalidad de Dmax

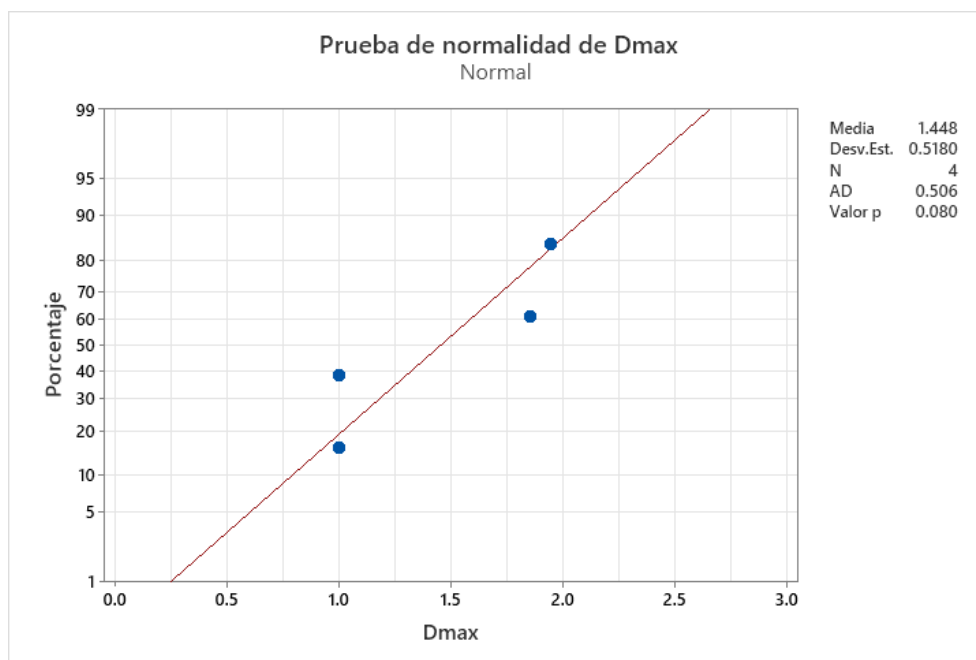
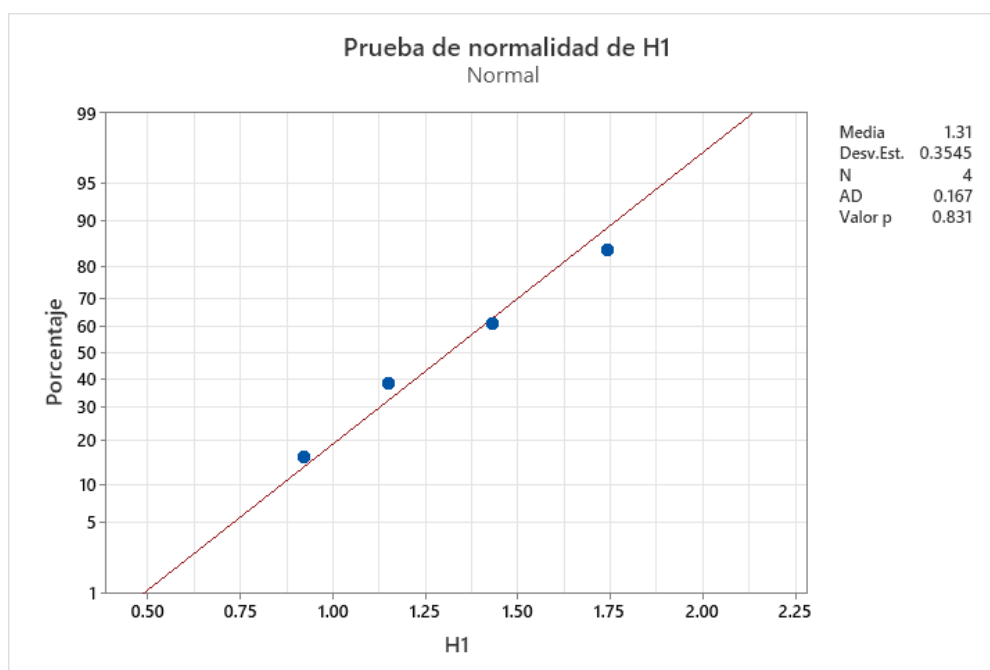


Figura 86

Prueba de normalidad de H1



e) Prueba t de una muestra (t-Student)

En la prueba t para Dmax, se obtuvo $t = -5.99$ con $p = 0.005$, por lo que se rechazó H_0 y se aceptó la alternativa, concluyéndose que el promedio poblacional de Dmax fue significativamente menor que 3. En la prueba t para H1, se obtuvo $t = -3.89$ con $p = 0.015$, por lo que se rechazó H_0 y se aceptó la alternativa, concluyéndose que el promedio poblacional de H1 fue significativamente menor que 2. De este modo, los resultados estadísticos respaldaron que la regionalización final cumplió simultáneamente los criterios de discordancia ($D_{max} < 3$) y heterogeneidad ($H_1 < 2$), aceptándose la hipótesis general planteada.

Tabla 55

Estadísticas descriptivas para Dmax y H1

	N	Media	Desv.Est.	Error estándar de media	Límite superior de 95% para μ
Dmax	4	1.448	0.518	0.259	2.057
H1	4	1.310	0.354	0.177	1.727

Tabla 56

Prueba t-Student para Dmax

Hipótesis nula (H_0): $\mu = 3$

Hipótesis alterna (H_1): $\mu < 3$

Valor T	Valor p
-5.99	0.005

Tabla 57

Prueba t de Student para H1

Hipótesis nula (H_0): $\mu = 2$

Hipótesis alterna (H_1): $\mu < 2$

Valor T	Valor p
-3.89	0.015

f) Limitaciones del análisis estadístico inferencial

La contrastación inferencial se efectuó con cuatro observaciones correspondientes a los valores regionales de Dmax y H1 obtenidos en las regiones homogéneas R-1, R-2, R-3 y R-4. Este tamaño muestral se ajustó a la estructura territorial definida mediante el Análisis Regional de Frecuencia (ARF) y comprendió la totalidad de las regiones validadas; por tanto, no se trató de una selección parcial de datos. Sin embargo, trabajar con $n=4$ implica una capacidad inferencial limitada. La potencia estadística de la prueba t de una muestra depende del tamaño del efecto, de la variabilidad observada y del nivel de significancia adoptado. Cuando el número de observaciones es reducido, la probabilidad de detectar diferencias moderadas disminuye y los resultados deben interpretarse con cautela (Cohen, 1998).

La prueba de normalidad de Anderson–Darling también se aplicó con fines exploratorios. Los valores obtenidos fueron: $p=0.080$ para Dmax, y: $p=0.831$ para H1. Ambos resultados fueron mayores que el nivel de significancia adoptado: $\alpha=0.05$

En consecuencia, no se rechazó el supuesto de normalidad. No obstante, este resultado no demuestra que los datos sigan necesariamente una distribución normal, ya que una muestra de cuatro observaciones presenta escasa capacidad discriminatoria ante posibles desviaciones de dicho supuesto. Por esta razón, la prueba t de una muestra se mantuvo como un análisis inferencial complementario y no como la única evidencia para aceptar la hipótesis.

El respaldo principal de la regionalización se basó en la verificación técnica individual de cada una de las cuatro regiones. En todos los casos se cumplieron simultáneamente los criterios establecidos para discordancia y heterogeneidad: $D_{max}<3$, $H1<2$

La significancia obtenida mediante la prueba t reforzó esta lectura; sin embargo, su interpretación quedó restringida al conjunto regional evaluado. Los resultados no deben generalizarse automáticamente a otras cuencas ni a configuraciones regionales distintas sin una validación adicional.

4.3. Discusión

El análisis de PMDA se basó en una red relativamente densa para el contexto altoandino, compuesta por 25 estaciones convencionales con amplia cobertura espacial (-79.53 a -78.46 ; -6.69 a -5.10) y un marcado gradiente altitudinal (1012 a 3300 msnm), lo que permitió estimar los máximos diarios anuales. Esta configuración supera la escala de redes empleadas en antecedentes que trabajaron con menos estaciones para máximos de 24 h, tal como: 7 estaciones por Corrales y Quezada (2025), 5 estaciones en Cadenillas (2024), 12 estaciones en Flores-Quispe et al. (2022), y se alinea con la recomendación de fortalecer la representatividad espacial en áreas con instrumentación heterogénea, donde la regionalización resulta útil cuando la red es dispersa de acuerdo con Yazawa y Shoji (2023). Además, el patrón de proximidad estimado (vecinos típicos de 10 a 30 km, pero con incrementos a 40–50 km y distancias máximas cercanas a 208.2 km) confirma una distribución no uniforme del muestreo, condición esperable en cuencas extensas y de relieve complejo, y que justifica realizar modelación regional.

De forma complementaria, el diagnóstico estadístico de las series confirma que la base de PMDA presenta disponibilidad heterogénea (21 a 60 años; completitud del 72.7% al 98.3%) y variabilidad marcada entre estaciones, con medias de 29.0 mm a 75.3 mm y eventos extremos puntuales que alcanzan 145.2 mm a 130.8 mm, lo cual es consistente con Flores-Quispe et al. (2022) y Cadenillas (2024). A diferencia de criterios más restrictivos que seleccionan series con fallas bajas (Pinheiro et al., 2024) o reducen significativamente el universo de estaciones para asegurar la consistencia

(Roca, 2021), en esta investigación se priorizó conservar la cobertura espacial mediante correcciones focalizadas: se identificó no estacionariedad por tendencia en 4 estaciones y no homogeneidad por saltos en 3 estaciones, tras lo cual se consolidó el uso de las 25 series corregidas.

La delimitación de regiones homogéneas no puede sostenerse únicamente en la agrupación preliminar, aun cuando exista coherencia geográfica. En este estudio, la combinación Ward-K-means permitió proponer 4 conglomerados iniciales con ausencia de outliers regionales ($D_{\max} < 3$; 2.04, 0.41, 2.43 y 1.31), pero la verificación formal con Hosking y Wallis confirmó heterogeneidad marcada ($H1 = 5.82, 5.87$ y 9.8), lo que se explica por la variabilidad apreciable de las ratios L-momentos entre estaciones (t 0.10 a 0.28, con contrastes en $t3$ y $t4$). Este comportamiento es consistente con Vásquez (2025) y Flores-Quispe et al. (2022), quienes muestran que los clústeres estadísticos requieren validación hidrológica y control geográfico para consolidar regiones estables; asimismo, coincide con Ticona (2024) y Cadenillas (2024) en que la homogeneidad debe sustentarse con pruebas de heterogeneidad aplicadas a L-momentos. En consecuencia, el ajuste iterativo aplicado (traslados/ exclusiones/ reorganización) permitió consolidar 4 regiones definitivas con tamaños más balanceados (8, 8, 4 y 4), manteniendo $D_{\max} < 3$ (1.94, 1.85, 1.00, 1.00) y alcanzando homogeneidad aceptable o definitiva ($H1 = 1.74, 1.15, 1.43$ y 0.92), incluso con rangos altitudinales internos (R-1: 1110–2232 m; R-2: 1012–3300 m; R-3: 1399–2748 m; R-4: 1627–2974 m). A diferencia de antecedentes donde coexisten distribuciones regionales distintas dentro del mismo ámbito (Cadenillas, 2024), aquí la selección por Z-estadístico descartó GLO por exceder el umbral en todas las regiones y llevó a una solución uniforme LNO ($|Z| \leq 1.64$ en las cuatro regiones), lo que favorece comparabilidad ARF-ROI; sin embargo, la estructura de colas diferenció regiones,

destacando R-3 con mayor amplificación de extremos ($qR(TR = 1000) = 3.35$ frente a 2.63, 1.98 y 2.22), coherente con cuantiles locales elevados en estaciones como Tocmoche (178 mm a $TR = 1000$) y Llama (173 mm), y con la posterior conversión a P24h mediante el factor 1.13 para su uso en diseño.

La comparación ARF vs ROI confirma que ambos transfieren información regional, pero con lógicas distintas y efectos diferentes en la cola: el ARF impone regiones fijas validadas por homogeneidad global, mientras que el ROI construye vecindades por estación objetivo, permitiendo traslapes y capturando zonas de transición, tal como advirtieron Roca (2021). En los resultados, ese contraste se materializa en tres patrones: (i) ARF más conservador que ROI y el ajuste local, como en Sondorillo ($TR = 1000$: 76 mm ARF vs 72 mm ROI vs 70 mm local) y en El Limón para TR altos ($TR = 1000$: 111 mm ARF vs 106 mm ROI vs 102 mm local), lo que sugiere que la regionalización fija intensifica la severidad de extremos al promediar colas regionales; (ii) ROI más conservador que ARF y local, como en Hacienda Pucará ($TR = 1000$: 106 mm ROI vs 95 mm ARF vs 93 mm local), evidenciando sensibilidad a la selección de estaciones influyentes y a la forma de la cola; y (iii) consistencia casi total entre métodos, como en Querocotillo ($TR = 1000$: 97–99 mm), donde la estación representa bien su entorno regional y el ROI no altera la extrapolación. Este comportamiento coincide con Martínez (2022) y Corrales y Quezada (2025), pero también con la advertencia dada por Eraso et al. (2024) quien indica que, en recurrencias altas, la estimación es sensible al modelo probabilístico y a la representatividad espacial del grupo. Por tanto, el ARF aporta estabilidad y trazabilidad regional (regiones homogéneas con $H1 < 2$), mientras que el ROI aporta adaptabilidad local y diagnóstico de incertidumbre por distribución en TR altos; por ello, la comparación sugiere usar ARF como marco base regional y ROI como contraste local.

La comparación cuantitativa más próxima geográficamente corresponde a Cadenillas Campos (2024), quien regionalizó precipitaciones máximas de 24 horas en la cuenca parcial del río Chotano. Se identificaron cuatro estaciones coincidentes con la base empleada en Chamaya: Chota, Chotano Lajas, Cutervo y Chugur. La estación Huambos no formó parte del inventario de la presente investigación. El contraste requiere cautela porque los estudios no construyeron sus series extremas con la misma escala temporal. En Chamaya se analizaron precipitaciones máximas diarias anuales (PMDA), obtenidas al seleccionar un único máximo diario por año, y posteriormente se aplicó el factor 1.13 para expresar los resultados como precipitaciones máximas acumuladas durante 24 horas continuas. Cadenillas Campos (2024), en cambio, trabajó con registros mensuales de precipitación máxima de 24 horas. Esta diferencia modifica la estructura de la muestra y condiciona el ajuste probabilístico. Por ello, la Tabla 58 presenta los valores de ambos estudios de manera paralela, como referencia contextual, sin calcular diferencias porcentuales ni asumir una equivalencia estadística directa.

Tabla 58

Contraste cuantitativo referencial con Cadenillas Campos (2024), considerando la diferente estructura temporal de las series

Estación	TR (años)	PMDA local de Chamaya (mm)	P24h de diseño en Chamaya: $1.13 \times$ PMDA (mm)	P24h reportada por Cadenillas Campos (2024) a partir de registros mensuales (mm)
Chota	50	71.71	81.03	47.34
Chota	100	76.25	86.16	52.70
Chota	500	86.32	97.55	65.71
Chotano Lajas	50	70.85	80.06	48.36
Chotano Lajas	100	76.16	86.06	54.68
Chotano Lajas	500	88.43	99.92	67.80
Cutervo	50	69.14	78.12	45.38
Cutervo	100	73.83	83.42	48.51
Cutervo	500	84.67	95.67	54.27
Chugur	50	93.85	106.05	45.38
Chugur	100	103.69	117.17	48.51
Chugur	500	125.95	142.33	54.27

Nota. TR: periodo de retorno; PMDA: precipitación máxima diaria anual; P24h: precipitación máxima estimada para una ventana continua de 24 horas. La tabla constituye un contraste descriptivo referencial.

Los valores no deben restarse ni interpretarse como estimaciones estrictamente equivalentes, debido a que Chamaya utilizó series de máximos diarios anuales, mientras que Cadenillas Campos (2024) trabajó con registros mensuales de precipitación máxima de 24 horas.

La Tabla 58 muestra que los valores estimados para Chamaya fueron superiores a los reportados por Cadenillas Campos (2024) en las cuatro estaciones coincidentes y en los tres periodos de retorno examinados. La separación fue más notoria en Chugur; no obstante, su trazabilidad ya fue verificada mediante la revisión del código de estación, la serie empleada y la aplicación única del factor 1.13. Estos contrastes no deben atribuirse automáticamente a una mayor severidad hidrológica ni interpretarse como evidencia de sobreestimación en alguno de los estudios. La diferencia puede responder a la escala temporal de las series, la longitud y depuración de los registros, el número de estaciones de soporte, la distribución probabilística seleccionada y el procedimiento de regionalización. En consecuencia, el antecedente permite contextualizar el orden de magnitud de los cuantiles, pero no constituye una validación numérica directa de los resultados de Chamaya.

La comparación con Roca Calderón (2021) cumplió una finalidad metodológica distinta. Su investigación se desarrolló en la zona suroccidental del Perú, bajo condiciones hidroclimáticas diferentes a las de Chamaya. Por esa razón, no correspondió contrastar valores absolutos estación por estación. Se examinó, en cambio, la sensibilidad de los cuantiles regionalizados frente al ajuste local. La Tabla 59 resume el rango de cuantiles obtenido mediante ARF y las variaciones porcentuales registradas en ambos estudios.

Tabla 59

Contraste metodológico de la sensibilidad del ARF respecto al ajuste local con Roca Calderón (2021)

Estudio y ámbito	TR (años)	Rango de cuantiles ARF (mm)	Menor diferencia ARF–local (%)	Mayor diferencia ARF–local (%)
Presente investigación: cuenca Chamaya	100	59.70–135.90	–4.6	8.1
Presente investigación: cuenca Chamaya	500	71.20–162.30	–0.1	8.4
Presente investigación: cuenca Chamaya	1 000	76.20–173.70	0.1	9.1
Roca Calderón (2021): zona suroccidental del Perú	100	10.00–80.00	–58.0	111.0
Roca Calderón (2021): zona suroccidental del Perú	500	13.00–103.00	–60.0	104.0
Roca Calderón (2021): zona suroccidental del Perú	1 000	15.00–115.00	–59.0	104.0

Nota. En Chamaya se consideraron las seis estaciones seleccionadas para la verificación comparativa entre ARF y ROI. En Roca Calderón (2021), los valores sintetizan las ocho estaciones utilizadas para evaluar el desempeño del ARF. La comparación examina la sensibilidad metodológica de la regionalización; no supone equivalencia climática entre ambos ámbitos.

La Tabla 59 evidencia que el ARF generó una variación más acotada que el ajuste local en la cuenca Chamaya. Para TR = 500 años, las diferencias oscilaron entre –0.1% y 8.4%; en Roca Calderón (2021), el rango se situó entre –60.0% y 104.0%. Para TR = 1 000 años se mantuvo el mismo patrón: Chamaya presentó variaciones entre 0.1% y 9.1%, mientras que el antecedente suroccidental alcanzó valores entre –59.0% y 104.0%. Esta menor dispersión respalda la estabilidad interna de las regiones homogéneas delimitadas en Chamaya, aunque no constituye una validación externa definitiva. El ROI mostró mayor sensibilidad al vecindario hidrológico seleccionado, con diferencias respecto del ajuste local de entre –13.6% y 14.5% para TR = 1 000 años. En términos operativos, el ARF se mantiene como procedimiento base para puntos interiores de regiones homogéneas verificadas; el ROI debe emplearse como contraste complementario en zonas de borde y sectores de transición.

La regionalización espacial de la PMDA en la cuenca del río Chamaya evidenció cuatro regiones homogéneas, con distintas extensiones territoriales y distintos soportes cartográficos. La Región 1 (3 571.59 km²; 43.89%) y la Región 2 (2 837.21 km²; 34.87%) concentraron conjuntamente el 78.76% del área y reunieron 16 estaciones empleadas en el ARF, todas con área efectiva dentro de la cuenca. La Región 3 (420.44 km²; 5.17%) y la Región 4 (1 307.95 km²; 16.07%) representaron el 21.24% restante. Ambas fueron verificadas estadísticamente con 4 estaciones cada una; sin embargo, solo 3 estaciones por región aportaron superficie tras el recorte de los polígonos de Thiessen. Esta diferenciación permitió aprovechar información externa para fortalecer la regionalización sin atribuir áreas de influencia inexistentes dentro del polígono de Chamaya. Este patrón es consistente con la evidencia de que, a escala de cuenca, es posible identificar un número acotado de unidades homogéneas cuando se emplean técnicas multivariadas y criterios regionales (Ramírez et al., 2023), y con estudios que, mediante L-momentos y pruebas de homogeneidad, delimitan regiones para la estimación de extremos, tales como, Ticona (2024) y Roca (2021). Pero, a diferencia de contextos con redes más dispersas, donde la densidad de estaciones condiciona la robustez espacial (Roca, 2021), en Chamaya, la regionalización estadística final se sustentó en 24 estaciones agrupadas en cuatro regiones homogéneas, mientras que la cartografía interna de la cuenca se construyó con 22 estaciones cuyos polígonos de Thiessen presentaron superficie efectiva tras el recorte espacial. Esta diferenciación permitió aprovechar estaciones externas para fortalecer el ARF sin atribuirles áreas que no representaban dentro del polígono de estudio.

Los mapas regionalizados mostraron crecimiento sistemático de los cuantiles al incrementar el periodo de retorno y una intensificación del contraste espacial, manteniendo el patrón general de gradientes dentro de la cuenca: para TR = 50 años los

rangos se ubicaron entre ≤ 60 mm y > 120 mm; para TR = 100 años entre ≤ 62 mm y > 134 mm; para TR = 200 años entre ≤ 65 mm y > 155 mm; y para TR = 500 años entre ≤ 75 mm y > 165 mm. Este comportamiento coincide con la tendencia reportada en regionalización de tormentas de diseño de Martínez (2022) y Escamilla (2023). Asimismo, la preservación del patrón espacial con aumento de magnitudes es concordante con regionalizaciones por clúster para PMD/24 h que diferencian dominios pluviométricos y producen valores crecientes en TR altos, tales como, Flores-Quispe et al. (2022) y Cadenillas (2024), y refuerza lo indicado por Yazawa y Shoji (2023) quienes argumentan que el ARF permite obtener superficies de diseño operativamente útiles, especialmente para TR elevados donde los ajustes estrictamente locales tienden a ser inestables y la regionalización resulta ventajosa para redes no uniformes.

En la cuenca del río Chamaya, las estrategias planteadas para mitigar la insuficiencia de datos hidrológicos se sustentaron en convertir una red pluviométrica espacialmente desigual en un sistema de estimación transferible y controlable mediante regionalización por L-momentos. En primer lugar, la transferencia regional mediante ARF se definió como estrategia base para estimar PMDA en puntos no instrumentados ubicados dentro de una región homogénea, generando cuantiles consistentes por TR y productos verificables (mapas y cuantiles regionales), en línea con Corrales y Quezada (2025) y con Martínez (2022). En segundo lugar, se incorporó la transferencia local ROI como estrategia complementaria para zonas de borde y transición, donde la continuidad espacial suele ser crítica, lo que es consistente con Roca (2021), quién destacaron que ARF y ROI son claves cuando la pluviometría es limitada y existen discrepancias en áreas particulares, y con Yazawa y Shoji (2023), al enfatizar que la gestión del riesgo debe integrar diferencias regionales de extremos donde la red es dispersa. En tercer lugar, la priorización de refuerzo de red basada en densidad y área

por estación (prioridad alta en regiones con $> 435 \text{ km}^2/\text{est.}$) operativizó una estrategia de reducción de incertidumbre a través de instrumentación focalizada, coherente con Escamilla (2023) y con Kupzig et al. (2024). Finalmente, la validación cruzada regional (local vs ARF vs ROI) y la integración de fuentes complementarias para series cortas se alinearon con Pinheiro et al. (2024), quienes demostraron que bases alternativas pueden reproducir patrones espaciales similares y aportar ventajas operativas, y con Ticona (2024) y Cadenillas (2024), al confirmar que L-momentos permite sostener regiones homogéneas y estimaciones confiables de extremos aun con redes reducidas, siempre que se incorporen verificaciones y criterios de homogeneidad.

CAPÍTULO V: CONCLUSIONES Y RECOMENDACIONES

5.1. Conclusiones

Al regionalizar las precipitaciones máximas diarias anuales (PMDA) en la cuenca del río Chamaya aplicando el método L-momentos, se llegó a las siguientes conclusiones:

- 1) Las series de PMDA de 25 estaciones meteorológicas, con cobertura espacial entre -79.53 a -78.46 de longitud y -6.69 a -5.10 de latitud, y un rango altitudinal de 1012 a 3300 msnm, integrando 19 estaciones dentro de la cuenca y 6 de apoyo por cuencas colindantes, presentaron longitudes heterogéneas de 21 a 60 años y completitud de 72.7% a 98.3%, con promedios entre 29.0 mm (Sondorillo) y 75.3 mm (La Cascarilla) y máximos anuales destacados de 145.2 mm (Huarmaca) y 130.8 mm (Tocmoche). La evaluación de calidad temporal identificó 4 estaciones con tendencia (Huarmaca, Tuluca, Chontalí y Tocmoche) y 3 no homogéneas (Porculla, Cueva Blanca y Chugur) que requirieron corrección, lo que permitió consolidar finalmente una base depurada de 25 estaciones utilizables para el análisis regional posterior.
- 2) La delimitación preliminar generó cuatro conglomerados sin estaciones discordantes críticas, con valores de D_{max} entre 0.41 y 2.43; sin embargo, la heterogeneidad regional alcanzó valores de H_1 entre 5.82 y 9.80, lo que evidenció la necesidad de reajustar la composición de los grupos. El procedimiento iterativo permitió validar cuatro regiones homogéneas conformadas por 8, 8, 4 y 4 estaciones, respectivamente, con $D_{max} \leq 1.94$ y valores de H_1 entre 0.92 y 1.74. La distribución LNO presentó el mejor ajuste regional en los cuatro grupos, con valores de $|ZDIST|$ entre 0.25 y 0.63, y generó cuantiles regionales adimensionales para $TR = 1\ 000$ años comprendidos entre $q_R = 1.98$ en R-2 y $q_R = 3.35$ en R-3. La comparación entre el Análisis Regional de Frecuencia (ARF) y el enfoque de Región de Influencia (ROI) demostró que ninguno de los métodos resulta

universalmente más conservador, ya que las diferencias dependen de la estación objetivo y de las estaciones que integran el grupo influyente. Para $TR = 1\,000$ años, el ARF estimó valores superiores en Sondorillo (76 mm frente a 72 mm mediante ROI), El Limón (111 mm frente a 106 mm), Hacienda Shumaya (117 mm frente a 101 mm) y, con una diferencia mínima, en Querocotillo (99 mm frente a 98 mm). En contraste, el ROI produjo cuantiles mayores en Hacienda Pucará (106 mm frente a 95 mm mediante ARF) y en Huarmaca (176 mm frente a 174 mm). En términos de aplicación, el ARF se recomienda como procedimiento base para estimar precipitaciones de diseño en puntos interiores de regiones homogéneas validadas, ya que proporciona una estructura regional estable y trazable. El ROI debe emplearse como contraste complementario en zonas de borde, sectores de transición o ubicaciones donde la similitud hidrológica local no coincida plenamente con los límites regionales.

- 3) Se elaboraron mapas regionales de PMDA a partir de cuatro regiones homogéneas definidas mediante ARF, cuya cobertura total alcanzó 8 137.19 km². La Región 1 abarcó 3 571.59 km², equivalente al 43.89% del área, mientras que la Región 2 comprendió 2 837.21 km², correspondiente al 34.87%. La Región 3 representó 420.44 km² (5.17%) y la Región 4 alcanzó 1 307.95 km² (16.07%). Las Regiones 3 y 4 fueron verificadas estadísticamente con 4 estaciones cada una, aunque solo 3 por región aportaron área efectiva para la espacialización cartográfica dentro de la cuenca. Los mapas evidenciaron un incremento progresivo y un mayor contraste entre los extremos al aumentar el periodo de retorno: las clases variaron de ≤ 60 mm a >120 mm para $TR = 50$ años; de ≤ 62 mm a >134 mm para $TR = 100$ años; de ≤ 65 mm a >155 mm para $TR = 200$ años; y de ≤ 75 mm a >165 mm para $TR = 500$ años. Estos resultados permitieron delimitar los sectores con mayor exigencia hidrológica dentro de la cuenca del río Chamaya.

- 4) La regionalización de PMDA mediante L-momentos permitió establecer estrategias operativas de transferencia de información para zonas poco instrumentadas de la cuenca del río Chamaya, sustentadas en 4 regiones homogéneas con coberturas desiguales. Los indicadores confirmaron brechas de representatividad en regiones extensas con densidad media (R-1: 2.24 est./1 000 km² y 446.45 km²/est.; R-4: 2.29 est./1 000 km² y 435.98 km²/est.); en consecuencia, se definieron dos rutas complementarias: ARF para estimar PMDA en puntos interiores de cada región y ROI para zonas de transición, generando productos verificables (cuantiles transferibles y mapas PMDA por TR) y una priorización técnica de refuerzo de la red con prioridad alta en R-1 y R-4, media–alta en R-2 (354.65 km²/est.) y baja en R-3.

5.2. Recomendaciones

Se sugiere la instauración de un protocolo consolidado de control de calidad y de actualización de series (verificación de la completitud, homogeneización y detección de tendencia) como premisa categórica de cualquier estimación de diseño, comenzando por la revisión cíclica de las estaciones que han mostrado presentar tendencia y no homogeneidad, complementando los huecos mediante la gestión de series de datos históricos y homogeneizando su forma de registro para mantenerse sustentadas en una única base fiable para las futuras regionalizaciones.

Se aconseja aplicar una forma operativa de selección metodológica: usar ARF cuando el punto de diseño se localiza en el interior de una región homogénea, o usar ROI si el punto de diseño se localiza en límites de transición; además, se recomienda incorporar validación cruzada (local–ARF–ROI) en estaciones representativas para detectar eventuales sobreestimaciones o subestimaciones en TR altos.

Se recomienda usar los mapas regionalizados a partir de TR para la planificación y el diseño hidráulico, destacando la localización de “zonas críticas con

predominancia de extremos”, y añadiendo que los mapas regionalizados deben actualizarse a medida que se incorporen nuevas estaciones o se extiendan las series, a fin de contar con una correlación espacial coherente de los cuantiles.

Se recomienda institucionalizar la estrategia dual ARF + ROI como esquema de transferencia de información en áreas sin estaciones y acompañarla con un plan de refuerzo progresivo de la red en áreas grandes de transición, a la vez que se sugiere plantear el control sucesivo de las estimaciones (contrastación por TR y coherencia espacial) para generar certidumbre de que los cuantiles transferidos sean técnicamente consistentes para aplicaciones de diseño y gestión del riesgo.

Se recomienda al Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI), a la Autoridad Nacional del Agua (ANA) y a las entidades responsables de formular proyectos hidráulicos que reevalúen periódicamente el supuesto de estacionariedad de las series de precipitaciones máximas diarias anuales en la cuenca del río Chamaya. Esta revisión deberá priorizar las estaciones Huarmaca, Tuluze, Chontalí y Tocmoche, ya que presentaron tendencias estadísticamente significativas en el análisis temporal. Como criterio operativo, se sugiere actualizar las pruebas de tendencia y de homogeneidad cada cinco años, ya sea al incorporarse nuevos registros o tras eventos hidrometeorológicos extremos que puedan modificar el comportamiento de las series. Si las tendencias persisten, será necesario complementar el análisis estacionario convencional con modelos no estacionarios y evaluar la influencia de la variabilidad climática y del cambio climático sobre los cuantiles de diseño. La actualización deberá extenderse a las regiones homogéneas, a los resultados de ARF y ROI y a los mapas regionalizados.

REFERENCIAS

- Achite, M., Caloiero, T., Jehanzaib, M., Wałęga, A., Kuriqi, A., & Pellicone, G. (2024). Geo-Statistical Characterization of Annual Maximum Daily Rainfall Variability in Semi-Arid Regions. *Atmosphere*, 15(12), 1-10. <https://doi.org/10.3390/atmos15121519>
- Ahmed, A., Rahman, A., Rafi, R. S., Khan, Z., & Mannan, H. (2025). Statistical and Physical Significance of Homogeneous Regions in Regional Flood Frequency Analysis. *Water*, 17(12), 1-34. <https://doi.org/10.3390/w17121799>
- Alejo-Sanchez, L. E., Márquez-Grajales, A., Salas-Martínez, F., Franco-Arcega, A., López-Morales, V., Acevedo-Sandoval, O. A., & Villegas-Vega, R. (2025). Missing data imputation of climate time series: A review. *MethodsX*, 15(1). <https://doi.org/10.1016/j.mex.2025.103455>
- Alvarez, I. J. (2022). *Modelo hidrológico para pronosticar caudales recesivos y estimar el coeficiente de agotamiento: caso río Jequetepeque aguas arriba de la presa "Gallito Ciego" - Contumazá, 2020*. [Tesis de grado, Universidad Nacional de Cajamarca]. <https://repositorio.unc.edu.pe/handle/20.500.14074/5261>
- ANA. (2011). *Diagnóstico de problemas y conflictos en la gestión de los recursos hídricos en la cuenca Chinchipe - Chamaya*. Autoridad Nacional del Agua (ANA). https://www.ana.gob.pe/sites/default/files/publication/files/estudio_final_2011-chinchipe-chamaya_0.pdf
- Bedient, P. E., Huber, W., & Vieux, B. (2008). *Hydrology and Floodplain Analysis*, Prentice Hall. *Prentice Hall*.
- Bustamante, F. (2019). *Regionalización morfométrica adimensional de cuenca hidrográfica Chamaya, con fines de transferencia de información hidrológica*. [Título para optar el título de ingeniero Hidráulico, Universidad Nacional de Cajamarca]. <http://hdl.handle.net/20.500.14074/3395>
- Cadenillas, L. Y. (2024). *Regionalización de precipitaciones máximas de 24 horas utilizando L-momentos en la cuenca del río Chotano, 2023*. [Tesis para optar el título profesional de ingeniero civil, Universidad Nacional Autónoma de Chota]. <http://hdl.handle.net/20.500.14142/546>
- Cao, C., Bai, Y., Peng, K., Han, B., & Zhang, Y. (2025). Regionalization of hydrological cycle changes in 31 source catchments of Yellow River Basin considering multiple

- hydrological variables. *Journal of Hydrology: Regional Studies*, 59, 102340. <https://doi.org/10.1016/j.ejrh.2025.102340>
- Castellarin, A., Vogel, R. M., & Matalas, N. C. (2007). Multivariate probabilistic regional envelopes of extreme floods. *Revista de Hidrología*, 336((3-4)), 376-390. <https://doi.org/https://doi.org/10.1016/j.jhydrol.2007.01.007>
- Cifuentes, D. (2020). *Automatización de las estimaciones de caudales de crecidas para diferentes periodos de retorno*. [Tesis para optar al grado de Magíster en Ciencias de la Ingeniería con Mención en Ingeniería Civil, Universidad de Concepción]. <https://repositorio.udec.cl/handle/11594/556>
- Cohen, J. (1998). *Statistical power analysis for the behavioral sciences (2nd ed.)*. Lawrence Erlbaum Associates.
- Corrales, I. D., & Quezada, J. P. (2025). *Regionalización de la curva IDF para la cuenca hidrográfica del Río Puca, Provincia del Guayas*. [Tesis de grado, Universidad Politécnica Salesiana de Ecuador]. <http://dspace.ups.edu.ec/handle/123456789/30277>
- Cutillas-Lozano, L. G., López, M. S., Velasco, A. P., Andrés-Doménech, I., & Olcina-Cantos, J. (2023). Local-scale regionalisation of climate change effects on rainfall pattern: application to Alicante City (Spain). *Theoretical and Applied Climatology*, 154(1), 377-402. <https://doi.org/https://doi.org/10.1007/s00704-023-04565-3>
- De-La-Hoz, E. J., De-La-Hoz, E. J., & Fontalvo, T. J. (2019). Metodología de Aprendizaje Automático para la Clasificación y Predicción de Usuarios en Ambientes Virtuales de Educación. *Información tecnológica*, 30(1), 247-254. <http://dx.doi.org/10.4067/S0718-07642019000100247>
- Departamento General de Irrigación. (2016). *Partes de una Cuenca*. Argentina: Aquabook.
- Eraaso, N. C., Gilon, C., & Pabón, S. M. (2024). *Regionalización de lluvias de diseño en el departamento de Nariño*. [Tesis de grado, Universidad Marina]. <https://repositorio.umariana.edu.co/server/api/core/bitstreams/d680a48e-f317-433a-861c-c4380d116c22/content>
- Escamilla, L. M. (2023). *Actualización de factores de regionalización de la cuenca Teapa y aplicación en MPE y HEC-HMS*. [Tesis de grado, Universidad Autónoma de México]. <https://ru.dgb.unam.mx/server/api/core/bitstreams/2185fa73-671a-47a4-896f-ce6197078f1a/content>
- Esteban, N. T. (2018). *Tipos de investigación*. Universidad Santo Domingo de Guzmán. <http://repositorio.usdg.edu.pe/handle/USDG/34>

- Estrany-Planas, P., Blanco-Gómez, P., Ortiz-Vallespí, J. I., Orihuela-Martínez, J., & Vilarrosa, V. (2025). Regional Frequency Analysis Using L-Moments for Determining Daily Rainfall Probability Distribution Function and Estimating the Annual Wastewater Discharges. *Hydrology*, 12(6), 152. <https://doi.org/10.3390/hydrology12060152>
- Ferguglia, O., Palazzi, E., & Arnone, E. (2024). Elevation dependent change in ERA5 precipitation and its extremes. *Climate Dynamics*, 62(8), 8137-8153. <https://doi.org/10.1007/s00382-024-07328-6>
- Ferreira, T. R., Liska, G. R., & Beijo, L. A. (2024). Assessment of alternative methods for analysing maximum rainfall spatial data based on generalized extreme value distribution. *Discover Applied Sciences*, 6(2), 34. <https://doi.org/10.1007/s42452-024-05685-9>
- Fitriana, F. S., Fischer, S., Singh, S. K., & Pahlow, M. (2025). Regional flood frequency analysis for hydrologically diverse regions in New Zealand. *Australasian Journal of Water Resources*, 29(2), 135-150. <https://doi.org/10.1080/13241583.2025.2530280>
- Flores, E. L., Flores, E., & Huaquisto, S. (2022). Regionalización de precipitación máxima diaria en Moquegua Perú. *Revista científica de la UCSA*, 9(2), 43-57. <https://doi.org/10.18004/ucsa/2409-8752/2022.009.02.043>
- Gamez, W. R. (2010). *Texto básico de hidrología*. Universidad Nacional Agraria La Molina. <https://doi.org/https://repositorio.una.edu.ni/id/eprint/2464>
- Govindaraju, R. S., & Das, B. S. (2007). *Moment analysis for subsurface hydrologic applications*. Springer Science & Business Media. <https://doi.org/https://books.google.es/books?hl=es&lr=&id=dCwLcyk8or8C&oi=fnd&pg=PR9&dq=Moment+analysis+for+subsurface+hydrologic+applications&ots=3SzNg2HofN&sig=M3BDcoEJJF39bdLEzTBzKrf-Zmo#v=onepage&q=Moment%20analysis%20for%20subsurface%20hydrologic%20applications&f=f>
- Grajales, T. (2000). *Tipos de investigación*. Efaind. <chrome-extension://efaidnbmninnibpcajpcglclefindmkaj/https://cmapspublic2.ihmc.us/rid=1RM1F0L42-VZ46F4-319H/871.pdf>
- Greenwood, J. A., Landwehr, J. M., Matalas, C. N., & Wallis, J. R. (1979). Probability weighted amoments: Definition and relation yo parameters of several. *Avances en las ciencias de la Tierra y el espacio*, 15(5), 1049-1054. <https://doi.org/https://doi.org/10.1029/WR015i005p01049>

- Grender, J. M. (1990). *heory and derivation for Weibull parameter probability weighted moment estimators* (Vol. 260). US Department of Agriculture, Forest Service, Southern Forest Experiment Station.
<https://books.google.es/books?id=o2EoknTD5cwC&ots=uTufUID1ep&dq=Theory%20and%20Derivation%20for%20Weibull%20Parameter%20Probability%20Weighted%20Moment%20Estimators%2C%20United%20States%20Department%20of%20Agriculture.&lr&hl=es&pg=PP1#v=onepage&q=Theory%2>
- Guevara, E. (2015). *Métodos para el análisis de variables hidrológicas y ambientales*. Autoridad Nacional del Agua.
<https://repositorio.ana.gob.pe/bitstream/handle/20.500.12543/3940/ANA0002522.pdf?sequence=1&isAllowed=y>
- Hernández-Sampieri, R., & Mendoza, C. (2020). *Metodología de la investigación: las rutas cuantitativa, cualitativa y mixta*. Mcgraw-hill.
https://d1wqtxts1xzle7.cloudfront.net/64591365/Metodolog%C3%ADa_de_la_investigaci%C3%B3n._Rutas_cuantitativa__cualitativa_y_mixta-libre.pdf?1601784484=&response-content-disposition=inline%3B+filename%3DM
- Hosking, J. R. (1990). Momentos L: Análisis y estimación de distribuciones mediante combinaciones lineales de estadísticas de orden. *Revista de la Royal Statistical Society: Serie B (Metodológica)*, 52(1), 105-124. <https://doi.org/10.1111/j.2517-6161.1990.tb01775.x>
- Hosking, J. R., & Wallis, J. R. (1997). *Análisis de frecuencia regional*. Cambridge University Press.
- Huang, Z., & Chen, Y. (2024). Impact of Rain Gauge Density on Flood Forecasting Performance: A PBDHM's Perspective. *Water*, 17(1), 18. <https://doi.org/10.3390/w17010018>
- Idrogo, D. L. (2024). *Regionalización de caudales para determinar umbrales de inundación en la cuenca Jequetepeque bajo condiciones normales, secas y húmedas*. [Tesis de grado, Universidad Nacional Pedro Ruiz Gallo]. <https://hdl.handle.net/20.500.12893/13216>
- Ilbay-Yupa, M., Lavado-Casimiro, W., Rau, P., Zubieta, R., & Castellón, F. (2021). Updating regionalization of precipitation in Ecuador. *Theoretical and Applied Climatology*, 143(1), 1513-1528. <https://doi.org/10.1007/s00704-020-03476-x>

- Ishaq, F. (2025). Application of Bayesian Techniques and L-Moments I Regional Frequency Analysis of Rainfall Extremes: Insight from Northern Pakistan. *Eksplorium-Buletin Pusat Teknologi Bahan Galian Nuklir*, 46(1), 1755-1770.
- Kavwenje, S., Zhao, L., Chen, L., Ngongondo, C., Chaima, E., Aborisade, M. A., & Kumambala, P. (2024). Integrated statistical and graphical non-parametric trend analysis of annual and seasonal rainfall in the Shire River Basin, Malawi. *Theoretical and Applied Climatology*, 155(3), 2053-2069. <https://doi.org/10.1007/s00704-023-04743-3>
- Kupzig, J., Kupzig, N., & Flörke, M. (2024). Regionalization in global hydrological models and its impact on runoff simulations: a case study using WaterGAP3 (v 1.0. 0). *Geoscientific Model Development*, 17(17), 6819-6846. <https://gmd.copernicus.org/articles/17/6819/2024/gmd-17-6819-2024.html>
- Llauca, H., Lavado-Casimiro, W., León, K., Jimenez, J., Traverso, K., & Rau, P. (2021). Assessing near real-time satellite precipitation products for flood simulations at sub-daily scales in a sparsely gauged watershed in Peruvian andes. *Remote Sensing*, 13(4), 826. <https://doi.org/https://doi.org/10.3390/rs13040826>
- Llmas, J. (1993). *Hidrología general: Principios y Aplicaciones*. Servicio. [Universidad del País Vasco España].
- MacQueen, J. (1967). Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, 5, 281-289.
- Markovič, L., & Faško, P. (2024). Regional frequency analysis for maximum 5-day precipitation totals using L-moments approach in Slovakia. *Theoretical and Applied Climatology*, 155(7), 5679-5693. <https://doi.org/10.1007/s00704-024-04970-2>
- Martínez, E. E. (2022). *Regionalización para obtener lluvias de diseño en la cuenca del río Santa Catarina del estado de Nuevo León*. [Tesis de grado, Universidad Autónoma de México]. <https://ru.dgb.unam.mx/server/api/core/bitstreams/6c170da8-2675-4884-b783-b2ad27c2d9ff/content>
- Mohammed, R., & Scholz, M. (2023). Quality Control and Homogeneity Analysis of Precipitation Time Series in the Climatic Region of Iraq. *Atmosphere*, 14(2), 197. <https://doi.org/10.3390/atmos14020197>
- Morales, F. (2012). Conozca 3 tipos de investigación: Descriptiva, Exploratoria y Explicativa. [online], 11(3), 1-5.

- MTC. (2018). *Manual de hidrología, hidráulica y drenaje*. Lima: Ministerio de Transportes y comunicaciones.
- Parizaca, Y. M. (2012). *Regionalización de precipitaciones máximas en la cuenca del río Ramis*. [Tesis de grado, Universidad Nacional del Altiplano]. <http://repositorio.unap.edu.pe/handle/20.500.14082/5494>
- Pettitt, A. N. (1979). A Non-Parametric Approach to the Change-Point Problem. *Royal Statistical Society*, 28(2), 126-135. <https://doi.org/https://doi.org/10.2307/2346729>
- Pinheiro, S. A., Reis, G. B., de Souza Fraga, M., da Silva, D. D., Abreu, M. C., & Parma, L. M. (2024). Flow regionalization using precipitation data from different bases as a predictive variable. *Physics and Chemistry of the Earth, Parts A/B/C*, 133, 103516. <https://doi.org/10.1016/j.pce.2023.103516>
- Ponce, V. M. (mayo de 2008). *Preguntas y respuestas sobre el período de retorno a ser usado para diseño*. 210620: https://ponce.sdsu.edu/periodos_de_retorno_articulo.html
- Prahadchai, T., Busababodhin, P., & Park, J. S. (2024). Regional flood frequency analysis of extreme rainfall in Thailand, based on L-moments. *Communications for Statistical Applications and Methods*, 31(1), 37-53. <http://www.csam.or.kr/journal/view.html?uid=2087&pn=lastest&vmd=Full>
- Puelles, J. (2015). *Estudio hidráulico e hidrológico de la Cuenca Alto Perú y el Porvenir en el asentamiento humano Las Mercedes Alto Perú, distrito de la Oroya, provincia de Yauli – Junín para la construcción futura de obras de arte ante amenazas de derrumbes*. Universidad Peruana de Ciencias Aplicadas. Lima: UPC. https://repositorioacademico.upc.edu.pe/bitstream/handle/10757/620953/PUELLES_MJ.pdf?sequence=1
- Ramírez, R. G., Wilson, M. G., Gabioud, E. A., Pighin, R. J., & Van Opsta, N. V. (2023). Regionalización Hidrológica de la Cuenca del Arroyo. *Argentina. Revista argentina de ingeniería*, 21, 115-122. <https://confedi.org.ar/wp-content/uploads/2023/05/Articulo12-RADI-21.pdf>
- Restrepo, J. D. (2005). *Los sedimentos del río Magdalena: Reflejo de la crisis ambiental*. Colciencias. <https://doi.org/https://books.google.es/books?hl=es&lr=&id=HKV1j3wKT2wC&oi=fnd&pg=PA21&dq=hidrolog%C3%ADa+fluvial&ots=Hd5MU5yUxM&sig=YseJp7MuVHsgKrXUk1QU-ji37fc#v=onepage&q=hidrolog%C3%ADa%20fluvial&f=false>
- Ricra, A. (2022). *Efecto de la precipitación pluvial y humedad relativa en el balance hídrico en la UNH Paturpampa Huancavelica - 2020*. [Tesis para optar el título de Ingeniero

- Ambiental y Sanitario, Universidad Nacional de Huancavelica].
<http://repositorio.unh.edu.pe/handle/UNH/4707>
- Roca, W. D. (2021). *Regionalización de precipitaciones máximas diarias anuales en la zona sur-occidental del Perú usando momentos-L*. [Tesis de grado, Universidad Nacional de Ingeniería]. <http://hdl.handle.net/20.500.14076/22235>
- Rojas, P. (9 de febrero de 2015). *¿Qué es una curva hipsométrica?* INGE CIV:
<https://ingeciv.com/que-es-una-curva-hipsometrica/>
- SENAMHI. (2020). *Mapa Climático del Perú*. Retrieved 18 de Octubre de 2025, from Servicio Nacional de Meteorología e Hidrología del Perú:
<https://www.senamhi.gob.pe/?p=mapa-climatico-del-peru>
- SENAMHI. (2023). *Boletín climático anual 2023*. Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI).
- SENAMHI. (2024). *Informe Técnico N°02-2024/SENAMHI-DMA-DHI-DAM Condiciones actuales de precipitación y perspectivas climáticas, hrológicas y agrometeorológicas en la región Andina*. Servicio Nacional de Meteorología e Hidrología del Perú (SENAMHI). https://cdn.www.gob.pe/uploads/document/file/6182521/5452866-informe-n-02-2024_senamhi-dma-dhi-dam-f-f-f-f-f.pdf
- Silva, I. D., Medeiros, D. M., Sakamoto, M. S., Leal Jr, J. B., Mendes, D., & Ambrizzi, T. (2022). Evaluating homogeneity and trends in extreme daily precipitation indices in a semiarid region of Brazil. *Frontiers in Earth Science*, 10(1).
<https://doi.org/10.3389/feart.2022.1071128>
- Skidelsky, D., Villaverde, Á. G., Mastaglia, M. I., Ramírez, R., Sato, R., Giménez, V., & Margasin, A. (2023). Regionalización de caudales máximos mediante técnicas estadísticas multivariadas en cuencas de la provincia de Entre Ríos. *Revista de geología aplicada a la ingeniería y al ambiente*(50), 9-21.
<https://ria.utn.edu.ar/xmlui/handle/20.500.12272/6242>
- Song, J. H., & Kim, H. (2025). Impact of precipitation station density and homogeneous regions on hydrological modeling performance. *Journal of Hydrology*, 661(B), 1-10.
<https://doi.org/10.1016/j.jhydrol.2025.133618>
- Stedinger, J. R., Vogel, R. M., & Foufoula, E. (1993). Frequency analysis of extreme events. *Handbook of Hydrology*.
- Ticona, R. (2024). *Estimación de precipitaciones máximas bajo la implementación del método de regionalización y estadístico aplicado en la cuenca cabanillas Departamento de*

- Puno. [Tesis de grado, Universidad Andina Néstor Cáceres Velásquez].
<https://repositorio.uancv.edu.pe/items/312f7fd8-a222-49df-83a8-2f1059e0dcce>
- Valenzuela-Morales, G. Y., Hernández-Téllez, M., Fonseca-Ortiz, C. R., Gómez-Albores, M. A., Esquivel-Ocadiz, A., Arévalo-Mejía, R., & Mastachi-Loza, C. A. (2023). Climatic and socioeconomic regionalization of the meteorological drought in Mexico using a predictive algorithm. *Natural Hazards*, 117(2), 1381-1403.
<https://doi.org/10.1007/s11069-023-05908-z>
- Vásquez, G. A. (2025). *Regionalización de precipitaciones promedio en la cuenca hidrográfica del Alto Marañón, mediante métodos de agrupación y el vector regional (RVM)*. [Tesis de grado, Universidad Nacional de Cajamarca].
<http://hdl.handle.net/20.500.14074/8179>
- Villón, M. (2004). *Hidrología*, 2ª ed. Instituto tecnológico de Costa Rica.
- Wallis, J. R., Matalas, C. N., & Slack, J. R. (1997). Apparent regional skew. *Water Resources Research*, 13(1), 159-182. <https://doi.org/10.1029/wr013i001p00159>
- Wang, H., Chen, W., Hu, Z., Xu, Y., & Shen, D. (2022). Optimal Rain Gauge Network Design Aided by Multi-Source Satellite Precipitation Observation. *Remote Sensing*, 14(23), 1-23. <https://doi.org/10.3390/rs14236142>
- Ward, J., & Joe, H. (1963). Hierarchical Grouping to Optimize an Objective Function. *Journal of the American statistical association*, 58(301), 236-244.
<https://www.tandfonline.com/doi/abs/10.1080/01621459.1963.10500845#>
- Weiss, L. L. (1964). Ratio of true to fixed-interval maximum rainfall. *Journal of the Hydraulics Division*, 90(HY1), 77-82. <https://doi.org/10.1061/JYCEAJ.0001008>
- Wilks, D. S. (2006). *Statistical methods in the atmospheric sciences*. Academic Press.
- Yazawa, T., & Shoji, A. (2023). Spatial analysis of historical extreme rainfall characteristics using regionalization in the Lake Biwa and Yodo River Basin, Japan. *Journal of Water and Climate Change*, 14(3), 916-936. <https://doi.org/10.2166/wcc.2023.465>
- Yue, S. P., Pilon, P., Phinney, B., & Cavadias, J. (2002). The influence of autocorrelation on the ability to detect trend in hydrological series. *Hydrological processes*, 16(9), 1807-1829. <https://doi.org/10.1002/hyp.1095>
- Zaghloul, M. A., Elshorbagy, A., & Papalexiou, S. M. (2024). Enhancing regional flood frequency analysis by integrating site-similarity measures with watershed modeling. *Journal of Hydrology*, 641(1). <https://doi.org/10.1016/j.jhydrol.2024.131754>
- Zamanillo, E., Larenze, G., Tito, M. J., Pérez, M., Garat, M. E., & Gómez, P. (2009). Aspectos metodológicos y resultados de la regionalización de precipitaciones máximas para la

provincia de Entre Ríos. *Cuadernos del CURIHAM*, 15, 13-29.

<https://doi.org/https://doi.org/10.35305/curiham.v15i0.66>

Zucarelli, G. (2017). *Regionalización hidrológica con métodos estadísticos multivariados*.

[Tesis de grado, Universidade da Coruña]. <http://hdl.handle.net/2183/19463>

ANEXOS

Anexo 1. Matriz de consistencia

Tesis: Regionalización de precipitaciones máximas diarias anuales en la cuenca del río Chamaya aplicando el método L – momentos, 2024

Tesista(s): Robbie Williams Ocsa Cubas; Jhensin Rivera Guamuro.

Formulación del problema	Objetivos	Hipótesis	Variable	Dimensión	Indicadores	Metodología	
¿Cómo se regionalizan las precipitaciones máximas diarias anuales en la cuenca del río Chamaya mediante la aplicación del método de L-momentos, 2024?	Objetivo general	La aplicación del método de L-momentos permite regionalizar las precipitaciones máximas diarias anuales en la cuenca del río Chamaya mediante la delimitación de regiones homogéneas estadísticamente consistentes y la estimación de cuantiles de diseño para diferentes periodos de retorno.	Precipitaciones máximas diarias anuales (PMDA) observadas	Inventario y cobertura espacial de estaciones	Ubicación de estaciones (altitud)		Enfoque: cuantitativo Tipo: básica Nivel: descriptivo Diseño: no experimental
	Proximidad espacial						
	Distribución espacial						
	Compleitud y consistencia del registro			Longitud del registro por estación (n años)			
				Años faltantes, % faltantes, % completos			
	Caracterización estadística de PMDA			Media / Mediana de PMDA			
				Desviación estándar y coeficiente de variación (CV)			
	Estacionariedad/ tendencia			Máximo, mínimo			
				Estaciones con tendencia y requerimiento de corrección			
	Homogeneidad (cambios en media/varianza)			Resultados Mann-Kendall, Spearman y tendencia en la media			
				Estaciones no homogéneas y corrección			
				Resumen cumplen/no cumplen por prueba			
	Agrupación preliminar	Resultados Mann-Whitney, t y F					
		Grupos preliminares Ward + K-means					
	Cálculo de L-momentos	Dmax por región preliminar					
		H1 por región preliminar					
	Ajuste iterativo y regiones definitivas (ARF)	l1, t (L-CV), t3 (L-Cs), t4 (L-Ck), n					
		Composición final de regiones homogéneas					
	Selección de distribución regional (ARF)	Dmax final por región y H1 final por región					
		Verificación gráfica ratios-L					
	Cuantiles regionales y locales (ARF)	Z por distribución y región (criterio $ Z \leq 1.64$)					
		Distribución óptima por región					
	Región de Influencia (ROI)	Cuantiles regionales adimensionales qR por TR					
		Cuantiles locales por estación y región					
	Comparación ARF vs ROI	Grupos ROI por estación objetivo					
		Sensibilidad a distribución en ROI (curvas por estación objetivo)					
	Cartografía y caracterización espacial	Comparación espacial ARF vs ROI					
Cuantiles comparados (Local vs ARF vs ROI) por estación							
Cobertura espacial por región ARF (área, % área, rangos lon/lat)							
Área de influencia por estación y región							
Mapa regiones homogéneas ARF							
	Mapas PMDA regionalizada por TR (50, 100, 200, 500)						
	Mapas finales ROI (grupos)						
						La muestra estuvo conformada por las series de PMDA (máximos diarios en 24 h por año) obtenidas de las estaciones meteorológicas seleccionadas para el estudio, tanto dentro como cercanas a la cuenca del río Chamaya. Estas series correspondieron a las estaciones que resultaron aptas para el análisis regional y para la conformación y verificación de regiones homogéneas.	

Anexo 2. Documentación

Anexo 2.1. Correo en el que SENAMHI proporcionó la data meteorológica

31/10/25, 3:17 p.m.

Gmail - TESISTA JHENSIN RIVERA GUAMURO



JHENSIN RIVERA GUAMURO <jhensinriveraguamuro@gmail.com>

TESISTA JHENSIN RIVERA GUAMURO

1 mensaje

Luis Felipe Gamarra Chávarry (UACGD) <lgamarra@senamhi.gob.pe>

24 de junio de 2025, 9:02

Para: "jhensinriveraguamuro@gmail.com" <jhensinriveraguamuro@gmail.com>, "Roberto Felipe Díaz Mendoza (UACGD)" <rdiaz@senamhi.gob.pe>, "Carmen Rosa Sulca Guzman (UACGD)" <csulca@senamhi.gob.pe>, "Milagros Del Pilar Martínez Tabra (UACGD)" <mmartinez@senamhi.gob.pe>

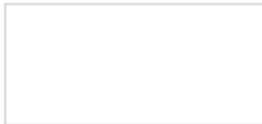
TESISTA JHENSIN RIVERA GUAMURO:

Reciba el cordial saludo del SERVICIO NACIONAL DE METEOROLOGÍA E HIDROLOGÍA DEL PERÚ - SENAMHI, a fin de remitirle los datos de las estaciones comprendidas en el expediente N° 06717 escala MENSUAL, periodo (1995 -2024); cabe indicar que bajo la modalidad de servicio gratuito es por ÚNICA VEZ. A la espera de su confirmación en cuanto a la recepción del presente y hacerle recordar que una vez concluida y aprobada su TESIS/INVESTIGACIÓN hacer llegar un ejemplar para nuestra biblioteca especializada, saludos.

Cabe indicar que de acuerdo a los procedimientos establecidos en la Directiva N° 003-2016-SENAMHI-SG-OPP-UM, para la atención a estudiantes, tesisistas e investigadores bajo la modalidad de servicio gratuito, dichos usuarios firman una DECLARACIÓN JURADA, que tiene un firme compromiso LEGAL de entregar un ejemplar de su tesis para nuestra biblioteca. En ese sentido la Srta. Carmen Sulca (csulca@senamhi.gob.pe), estará atenta al cumplimiento de dicho acuerdo, saludos.

Nota: Se envía solo las variables y datos disponibles. No se cuenta con datos de las estaciones HACIENDA SHUMAYA y TULLUCE

EXPEDIENTE: 2025-06717



Luis Felipe Gamarra Chavarry
DIRECTOR DE PLANEAMIENTO, COORDINACION Y
CONTROL
GERENCIA GENERAL
SENAMHI - PERÚ

D: Jr. Cahuide 785, Jesús María - Lima
T: 01 6141414 Anexo -
C: -
E: lgamarra@senamhi.gob.pe
W: www.senamhi.gob.pe

10 adjuntos

CHOTA_M_1995-2024.xlsx
13K

CHOTANO LAJAS_M_1995-2024.xlsx
14K

COCHABAMBA_M_1995-2024.xlsx
13K

CUEVA BLANCA_M_1995-2024.xlsx
14K

CUTERVO_M_1995-2024.xlsx
13K

EL LIMON.xlsx
13K

HACIENDA PUCARA_M_1995-2024.xlsx
14K

PORCULLA_M_1995-2024.xlsx
12K

QUEROCOTILLO_M_1995-2024.xlsx
13K

SONDORILLO_M_1995-2024.xlsx
13K

Anexo 3. Resultado del análisis de regionalización

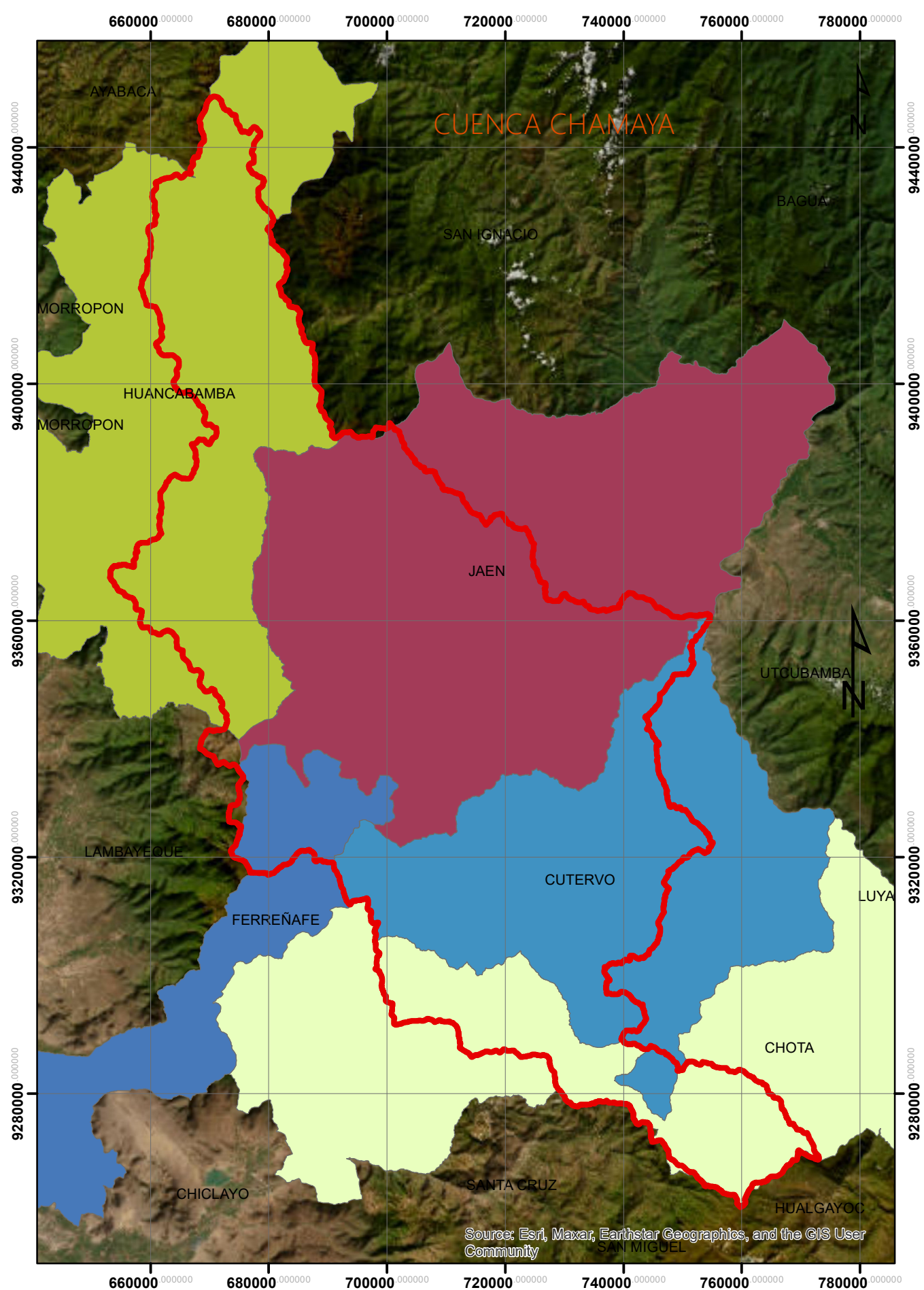
Este anexo compila los datos primarios y los archivos de cálculo utilizados para la regionalización de las precipitaciones máximas diarias anuales (PMDA) en la cuenca del río Chamaya mediante L-momentos. Incluye las series por estación (SENAMHI y estaciones de apoyo), metadatos de ubicación (coordenadas y altitud), control de calidad y depuración, y los resultados del proceso de agrupación y validación de regiones homogéneas (ARF y ROI), así como cuantiles por periodos de retorno e insumos para los mapas. La versión digital se organiza en cuatro carpetas: (1) Indicaciones, (2) Proceso, (3) Grupos iniciales y (4) Grupos finales (Regiones homogéneas). La versión electrónica puede descargarse en el siguiente enlace:

<https://drive.google.com/drive/folders/1OKOKkVrzrpEAydo2VLfDmzfKQth-lOrB>

También se encuentra disponible a través del código QR adjunto.



Anexo 4. Mapas



UNIVERSIDAD NACIONAL
AUTÓNOMA DE CHOTA

FACULTAD DE CIENCIAS DE LA INGENIERÍA
ESCUELA PROFESIONAL DE INGENIERÍA CIVIL

Regionalización de precipitaciones máximas diarias anuales en la cuenca del
río Chamaya, aplicando el método L-Momentos, 2024

Presentado por: Bach. Robbie Williams Ocsa Cubas
Bach. Jhensin Rivera Guamuro

Asesor:
Dr. Ing. Luis Alberto Orbegoso Navarro

Plano:

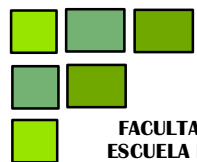
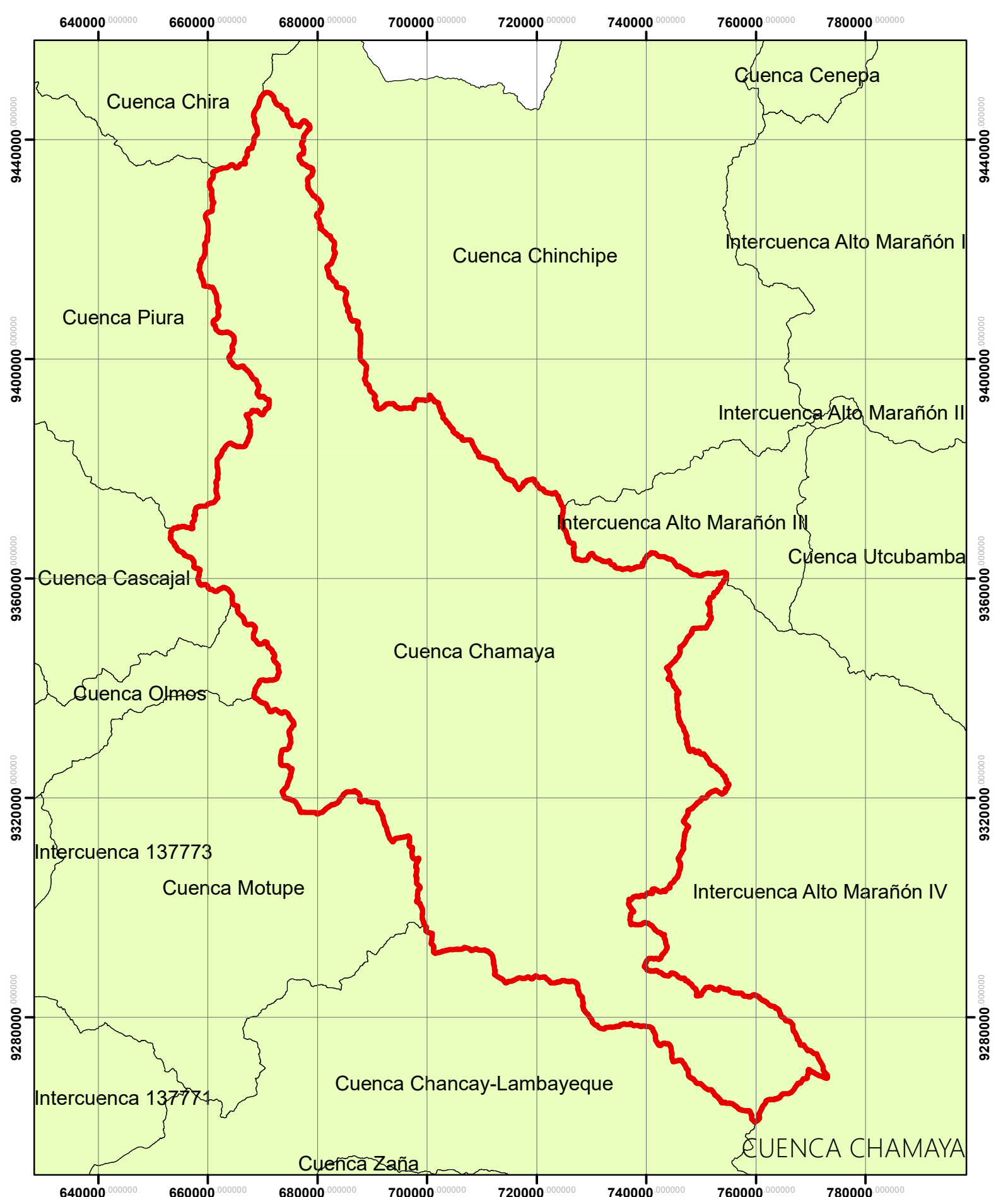
MAPA DE UBICACIÓN
DE LA CUENCA CHAMAYA

Fecha: Enero, 2026
Escala: 1:869,927

0 5 10 20 30 40
km

Mapa:

U-1



UNIVERSIDAD NACIONAL
AUTÓNOMA DE CHOTA

FACULTAD DE CIENCIAS DE LA INGENIERÍA
ESCUELA PROFESIONAL DE INGENIERÍA CIVIL

Regionalización de precipitaciones máximas diarias anuales en la cuenca del
río Chamaya, aplicando el método L-Momentos, 2024

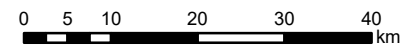
Presentado por: Bach. Robbie Williams Ocsa Cubas
Bach. Jhensin Rivera Guamuro

Asesor:
Dr. Ing. Luis Alberto Orbegoso Navarro

Plano:

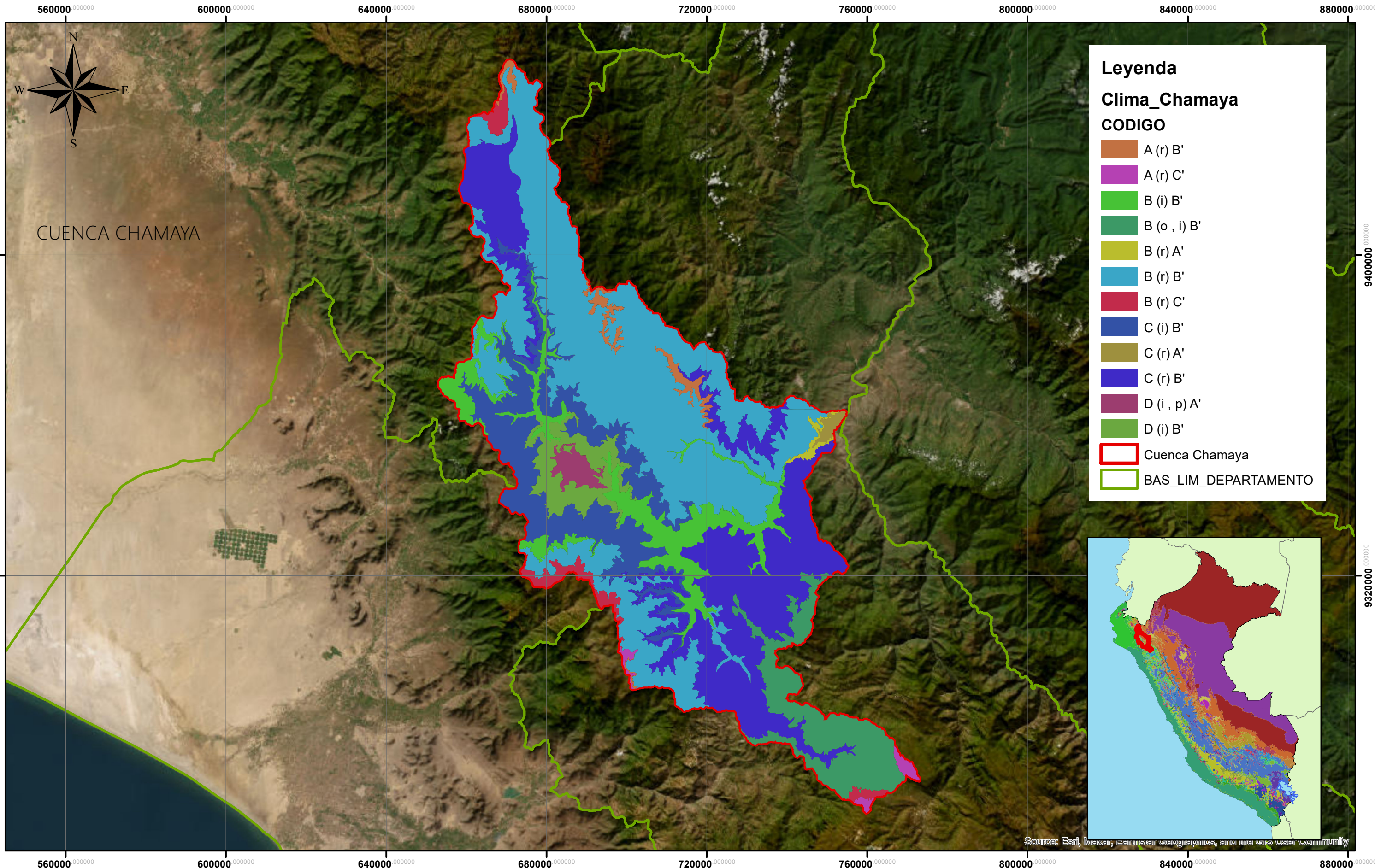
MAPA HIDROLÓGICO
DE LA CUENCA CHAMAYA

Fecha: Enero, 2026
Escala: 1:869,927



Mapa:

U-1



Legenda

Clima_Chamaya

CODIGO

A (r) B'

A (r) C'

B (i) B'

B (o , i) B'

B (r) A'

B (r) B'

B (r) C'

C (i) B'

C (r) A'

C (r) B'

D (i , p) A'

D (i) B'

Cuenca Chamaya

BAS_LIM_DEPARTAMENTO



UNIVERSIDAD NACIONAL
AUTÓNOMA DE CHOTA

FACULTAD DE CIENCIAS DE LA INGENIERÍA
ESCUELA PROFESIONAL DE INGENIERÍA CIVIL

Regionalización de precipitaciones máximas diarias anuales en la cuenca del
río Chamaya, aplicando el método L-Momentos, 2024

Presentado por:

Bach. Robbie Williams Ocsa Cubas
Bach. Jhensin Rivera Guamuro

Asesor:

Dr. Ing. Luis Alberto Orbegoso Navarro

Plano:

**MAPA CLIMATOLÓGICO
DE LA CUENCA CHAMAYA**

Fecha: Enero, 2026
Escala: 1:869,590

0

5

10

20

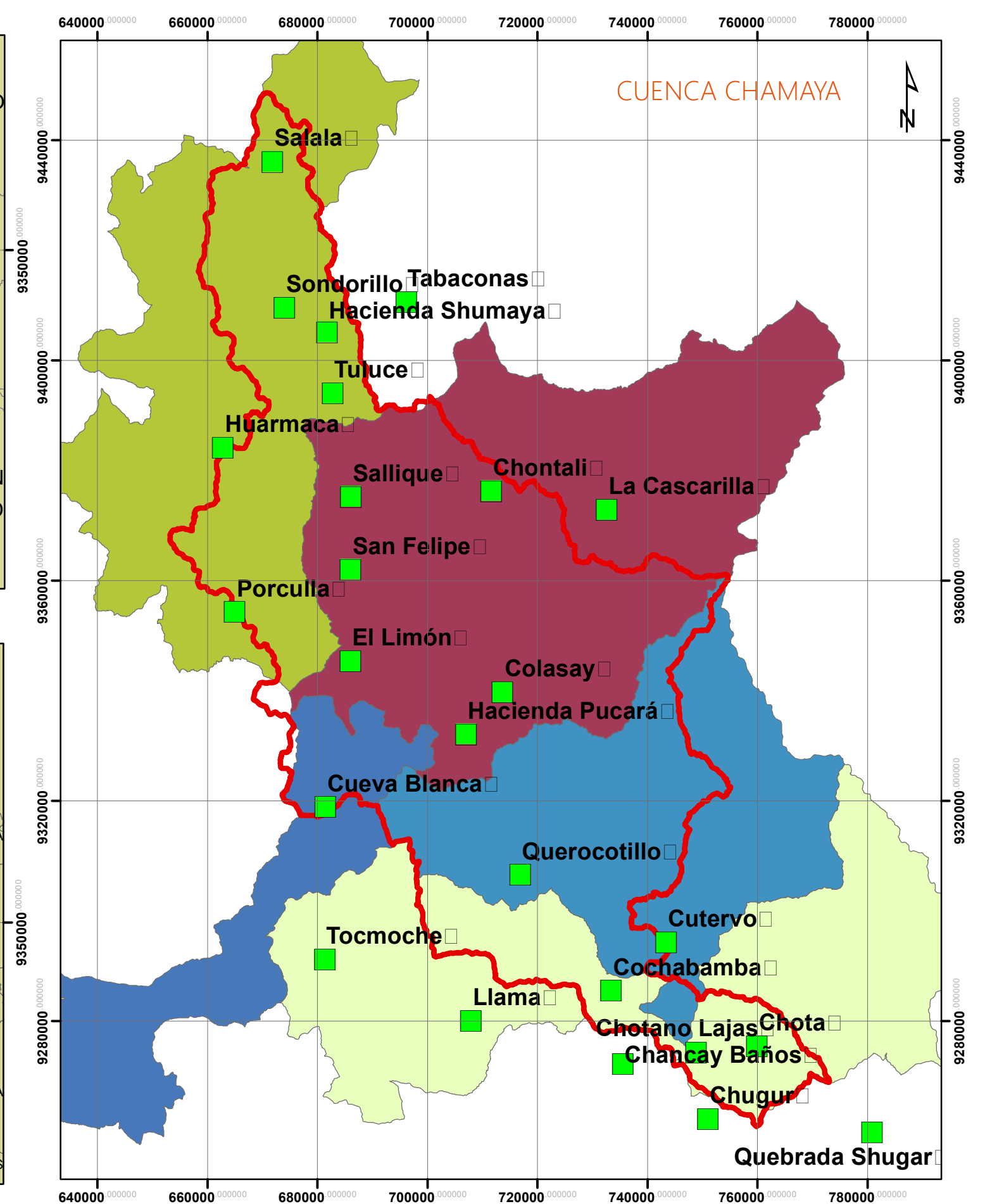
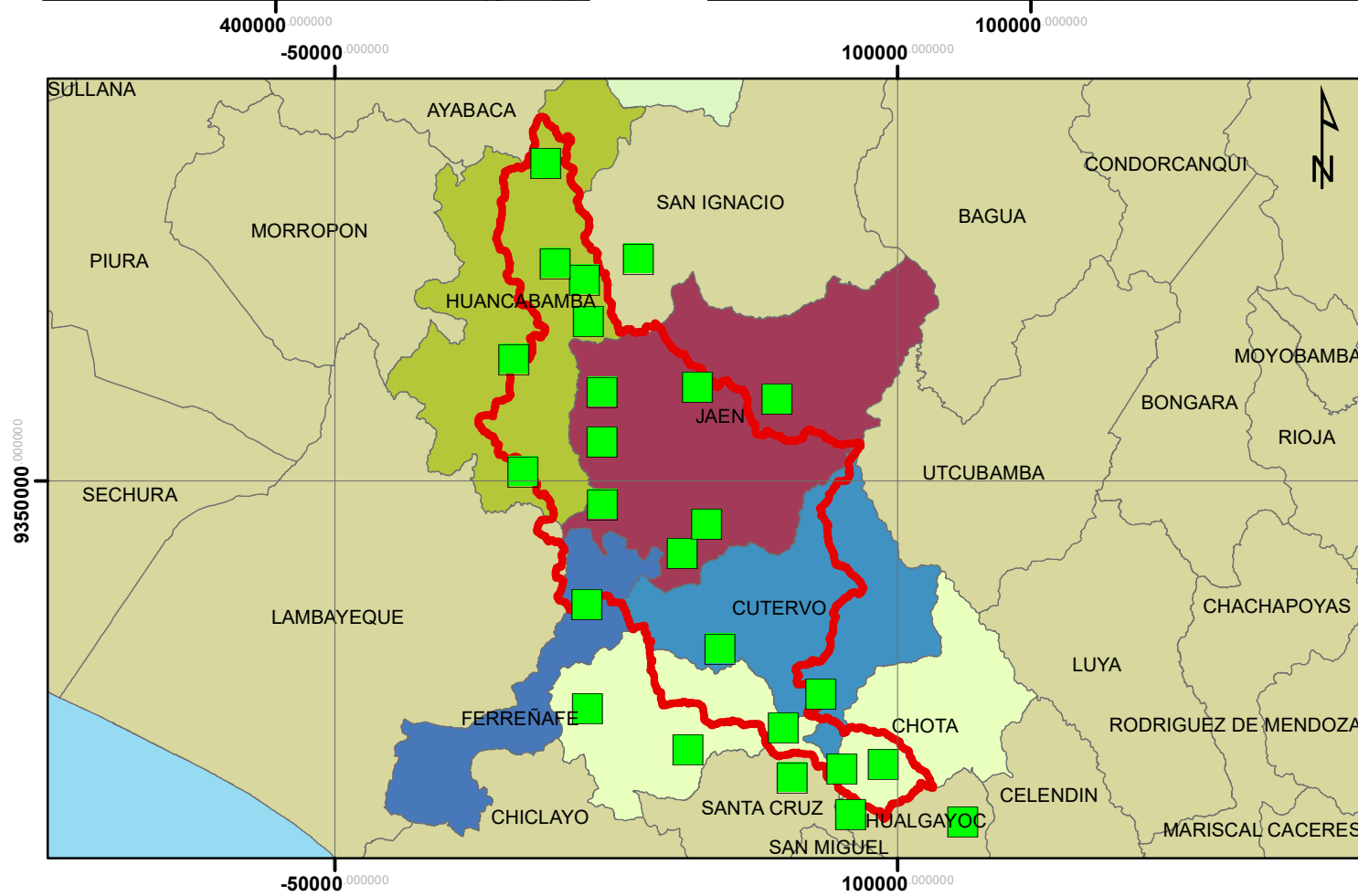
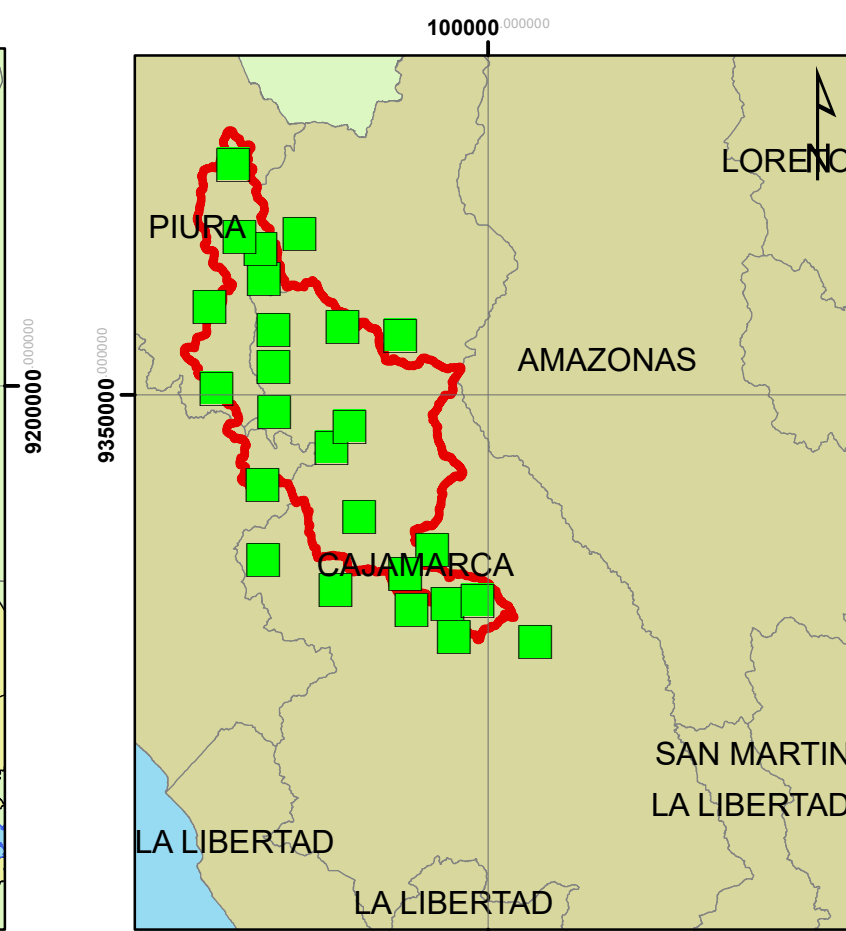
30

40

km

Mapa:

U-1



UNIVERSIDAD NACIONAL
AUTÓNOMA DE CHOTA

FACULTAD DE CIENCIAS DE LA INGENIERÍA
ESCUELA PROFESIONAL DE INGENIERÍA CIVIL

Regionalización de precipitaciones máximas diarias anuales en la cuenca del
río Chamaya, aplicando el método L-Momentos, 2024

Presentado por: Bach. Robbie Williams Ocsa Cubas
Bach. Jhensin Rivera Guamuro

Asesor:
Dr. Ing. Luis Alberto Orbegoso Navarro

Plano: MAPA DE UBICACIÓN DE LAS ESTACIONES
METEOROLÓGICAS DE LA CUENCA CHAMAYA

Fecha: Enero, 2026
Escala: 1:869,927

0 5 10 20 30 40 km

Mapa:
U-4

