



ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO
FACULTAD DE CIENCIAS
CARRERA ESTADÍSTICA

**ANÁLISIS ESPACIAL DE LAS PRECIPITACIONES EXTREMAS EN
LA PROVINCIA DE CHIMBORAZO (1990-2018)**

Trabajo de Integración Curricular

Tipo: Proyecto de Investigación

Presentado para optar al grado académico de:

INGENIERO ESTADÍSTICO

AUTOR: JUAN CRISTIAN HIDALGO LLANGARÍ

DIRECTOR: ING. AMALIA ISABEL ESCUDERO VILLA

Riobamba – Ecuador

2023

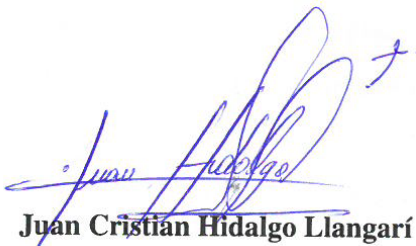
©2023, Juan Cristian Hidalgo Llangari

Autorizo la reproducción total o parcial, con fines académicos, por cualquier medio o procedimiento, incluyendo la cita bibliográfica del documento, siempre y cuando se reconozca el Derecho de Autor.

Yo, JUAN CRISTIAN HIDALGO LLANGARÍ, declaro que el presente Trabajo de Integración Curricular es de mi autoría y los resultados del mismo son auténticos. Los textos en el documento que provienen de otras fuentes están debidamente citados y referenciados.

Como autor asumo la responsabilidad legal y académica de los contenidos de este Trabajo de Integración Curricular; El patrimonio intelectual pertenece a la Escuela Superior Politécnica de Chimborazo.

Riobamba, 25 de abril del 2023


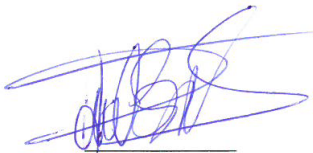
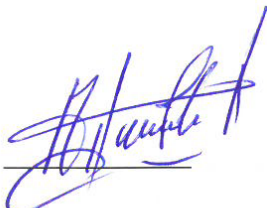


Juan Cristóbal Hidalgo Llangarí

060412726-6

ESCUELA SUPERIOR POLITÉCNICA DE CHIMBORAZO
FACULTAD DE CIENCIAS
CARRERA ESTADÍSTICA

El Tribunal del Trabajo de Integración Curricular certifica que: el Trabajo de Integración Curricular; Tipo: Proyecto de Investigación. **ANÁLISIS ESPACIAL DE LAS PRECIPITACIONES EXTREMAS EN LA PROVINCIA DE CHIMBORAZO (1990-2018)**, realizado por: **JUAN CRISTIAN HIDALGO LLANGARÍ**, ha sido minuciosamente revisado por los Miembros del Tribunal del Trabajo de Integración Curricular, el mismo que cumple con los requisitos científicos, técnicos, legales, en tal virtud el Tribunal Autoriza su presentación.

	FIRMA	FECHA
Ing. Johanna Enith Aguilar Reyes, Mgs. PRESIDENTE DEL TRIBUNAL		023-04-25
Ing. Amalia Isabel Escudero Villa DIRECTOR DEL TRABAJO DE INTEGRACIÓN CURRICULAR		2023-04-25
Dr. Arquimides Xavier Haro Velastegui ASESOR DEL TRABAJO DE INTEGRACIÓN CURRICULAR		2023-04-25

DEDICATORIA

Dedico mi trabajo a Dios por las bendiciones recibidas en cada uno de los años pasados en la Epoch y a mis padres por ser la parte fundamental y apoyarme en todas las decisiones que en el transcurso de la vida que he tomado.

Juan

AGRADECIMIENTO

En primera instancia agradezco a Dios por guiarme en el transcurso del desarrollo de este proyecto de investigación, a su vez a toda la planta docente de la carrera de Estadística que con el pasar de los semestres me fueron formando como profesional. Mi familia fue el pilar para poder terminar con un sueño que es realidad, mis padres Elizabeth y Juan que con su esfuerzo ya ven a su hijo profesional y ser un buen ejemplo para mis hermanos Jhonny, Fernando y Daniel, mis sobrinos Antony y Domenica. No olvidando a mis abuelitos que con el ejemplo dieron muchos valores a mis padres que ellos nos inculcaron cada día. Fueron muchos amigos que se hizo en la politécnica pero algunos de ellos que siempre van a estar en mi corazón.

De igual manera agradezco a mis profesores en específico a la Ing. Isabel Escudero; Directora del trabajo de titulación, al Dr. Arquímedes Haro; Miembro del Tribunal, y a la Ing. Johanna Aguilar; Presidenta del Tribunal, quienes con sus valiosos conocimientos hicieron que posible la culminación de este trabajo de titulación, gracias a cada una de ustedes por su paciencia, dedicación, apoyo incondicional y amistad.

Agradezco de sobre manera al Grupo de Investigaciones de Energías Alternativas y Ambiente (GEAA) de la ESPOCH, al INAMHI y al Ing. Cristian Paliz, por el apoyo a la realización de esta investigación.

Juan

ÍNDICE DE CONTENIDO

ÍNDICE DE TABLAS	ix
ÍNDICE DE ILUSTRACIONES	x
ÍNDICE DE ANEXOS	xi
RESUMEN	xii
SUMMARY/ABSTRACT	xiii
INTRODUCCIÓN	1

CAPÍTULO I

1. MARCO TEÓRICO REFERENCIAL	3
1.1. Planteamiento del problema	3
1.1.1. <i>Enunciado del problema</i>	3
1.2. Justificación	3
1.3. Objetivos	4
1.3.1. <i>Objetivo general</i>	4
1.3.2. <i>Objetivos específicos</i>	4

CAPÍTULO II

2. MARCO TEÓRICO	5
2.1. Características atmosféricas	5
2.2. Precipitación	5
2.2.1. <i>Tipos de precipitación</i>	6
2.2.2. <i>Medición de precipitación</i>	7
2.3. Datos geográficos	8
2.3.1. <i>Características de datos geográficos</i>	8
2.3.2. <i>Modelos geográficos</i>	8
2.4. Efectos espaciales	8
2.4.1. <i>Heterogeneidad espacial</i>	8
2.4.2. <i>Autocorrelación o dependencia espacial</i>	9
2.4.3. <i>Tipos de autocorrelación espacial</i>	9

2.4.4.	<i>Medición de autocorrelación espacial</i>	9
2.5.	Interpolación espacial	9
2.6.	Interpolación de IDW	10
2.7.	Teoría de valores extremos (TVE)	10
2.7.1.	<i>Modelos máximos por bloques</i>	10
2.7.2.	<i>Umbral</i>	11
2.7.3.	<i>Picos sobre el umbral</i>	11
2.7.4.	<i>Modelos de excedencias de umbrales</i>	13
2.8.	Muestreo de Bootstrap	13
2.9.	Distribución de remuestreo	15
2.10.	Software R para analizar datos espaciales en R	15
2.10.1.	<i>Liberías para datos espaciales</i>	15

CAPÍTULO III

3.	MARCO DE METODOLOGICO	17
3.1.	Enfoque de investigación	17
3.2.	Nivel de investigación	17
3.3.	Diseño de la investigación	17
3.3.1.	<i>Según la manipulación o no de la variable independiente</i>	17
3.4.	Tipo de estudio	18
3.5.	Población, planificación, selección y calculo del tamaño de la muestra	18
3.5.1.	<i>Población de estudio</i>	18
3.5.2.	<i>Tamaño de la muestra</i>	18
3.5.3.	<i>Localización del proyecto</i>	18
3.6.	Métodos, técnicas e instrumentos de investigación	19
3.6.1.	<i>Métodos de investigación</i>	19
3.6.2.	<i>Técnicas de investigación</i>	19
3.6.3.	<i>Instrumentos de investigación</i>	19

CAPÍTULO IV

4.	MARCO DE ANÁLISIS E INTERPRETACIÓN DE RESULTADOS	20
4.1.	Análisis exploratorio	20
4.1.1.	<i>Identificación geográfica de la zona en estudio</i>	20
4.1.2.	<i>Identificación de datos</i>	22
4.1.3.	<i>Identificación de variable en estudio</i>	24
4.1.4.	<i>Análisis de las precipitaciones</i>	26
4.1.5.	<i>Datos faltantes</i>	26
4.2.	Análisis de los umbrales	27
4.2.1.	<i>Identificación de los umbrales</i>	27
4.3.	Análisis mediante el umbral fijo	27
4.3.1.	<i>Precipitaciones extremas mediante umbral fijo</i>	31
4.4.	Mapa de riesgos extremos de precipitación en la provincia de Chimborazo con umbral fijo	32
4.5.	Análisis de valores extremos mediante el umbral variable	33
4.5.1.	<i>Identificación de umbrales</i>	33
4.5.2.	<i>Precipitaciones extremas mediante umbrales variables</i>	38
4.6.	Mapa de riesgos extremos de la provincia de Chimborazo con umbral variable	39
4.7.	Autocorrelación espacial	40
4.8.	Estrategias para el análisis recurrente de las precipitaciones extremas en la provincia de Chimborazo	40
	CONCLUSIONES	42
	RECOMENDACIONES	43
	BIBLIOGRAFÍA	
	ANEXOS	

ÍNDICE DE TABLAS

Tabla 4-1: Ubicación geográfica de las diferentes estaciones meteorológicas en la provincia de Chimborazo	21
Tabla 4-2: Precipitaciones	23
Tabla 4-3: Número de registros de precipitación en cada estación	24
Tabla 4-4: Estadísticos descriptivos de las precipitaciones registradas en las diferentes estaciones	25
Tabla 4-5: Parámetros de escala σ y forma ξ de la estación Huigra con umbral fijo (12 mm) con remuestreo de Bootstrap	29
Tabla 4-6: Parámetros de escala σ y forma ξ de la estación Chimbo DJ Pangor con umbral fijo (12 mm) con remuestreo de Bootstrap	30
Tabla 4-7: Parámetros de escala σ y forma ξ de la estación Huigra con umbral variable (6,68 mm) con remuestreo de Bootstrap	34
Tabla 4-8: Parámetros de escala σ y forma ξ de la estación Chimbo DJ Pangor con umbral variable (14,50 mm) con remuestreo de Bootstrap	36
Tabla 4-9: Umbrales y parámetros de las estaciones con mejores valores	37
Tabla 4-10: Índice de Moran y P-valor de los valores máximos, mínimos y medianas de excedentes con umbrales variables	40

ÍNDICE DE ILUSTRACIONES

Ilustración 2-1: Pluviometro	6
Ilustración 2-2: Proceso de estimación Bootstrap	13
Ilustración 3-1: Mapa de la provincia de Chimborazo	18
Ilustración 4-1: Estaciones meteorológicas en la provincia de Chimborazo	20
Ilustración 4-2: Estación ubicada en la ESPOCH	22
Ilustración 4-3: Boxplot	26
Ilustración 4-4: Datos faltantes	26
Ilustración 4-5: Identificación del umbral mediante el cuantil 95	27
Ilustración 4-6: Vida residual media	28
Ilustración 4-7: Distribución de Pareto generalizado de la estación Huigra con umbral fijo	28
Ilustración 4-8: Distribución de Pareto generalizada de la estación Chimbo DJ Pangor con umbral fijo	29
Ilustración 4-9: Precipitaciones extremas con umbral fijo mediante la interpolación IDW .	31
Ilustración 4-10: Riesgos extremos con umbral fijo	32
Ilustración 4-11: Identificación del umbral mediante el cuantil 95 de la estación Huigra .	33
Ilustración 4-12: Vida residual media de la estación Huigra	33
Ilustración 4-13: Distribución de Pareto generalizado de la estación Huigra con umbral variable	34
Ilustración 4-14: Identificación del umbral mediante el cuantil 95 de la estación Chimbo DJ Pangor	35
Ilustración 4-15: Vida residual media de la estación Chimbo DJ Pangor	35
Ilustración 4-16: Distribución de Pareto generalizado de la estación Chimbo DJ Pangor con umbral variable	36
Ilustración 4-17: Precipitaciones extremas con umbrales variables	38
Ilustración 4-18: Riesgos extremos con umbral variable	39

ÍNDICE DE ANEXOS

ANEXO A: SCRIPT DE R UTILIZADO EN LA INVESTIGACIÓN PARA OBTENER EL
ANÁLISIS EXPLORATORIO

ANEXO B: DATOS FALTANTES

ANEXO C: IDENTIFICACIÓN DE UMBRALES POR EL CUANTIL Y EL GRÁFICO DE VIDA
RESIDUAL MEDIA

ANEXO D: DISTRIBUCIÓN DE PARETO GENERALIZADA

ANEXO E: REMUESTREO BOOTSTRAP

ANEXO F: DISTRIBUCIÓN MUESTRAL

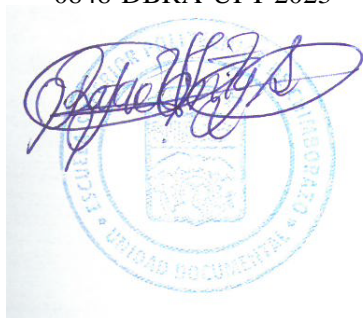
ANEXO G: INTERPOLACIÓN IDW EN VALORES EXTREMOS Y VALORES DE RIESGO

RESUMEN

En los últimos años el cambio climático ha sido un detonante en la estocasticidad del comportamiento de las variables meteorológicas. La provincia de Chimborazo al contar con microclimas presenta precipitaciones con características específicas localizadas. Investigaciones recientes se interesaron en el comportamiento de las precipitaciones en la ciudad de Riobamba, iniciando la utilización de la teoría de valores extremos desde un enfoque univariante en dicha zona, es decir utilizando los registros de una estación meteorológica. Con el propósito de dar continuidad a dicha investigación este trabajo de tesis propone estrategias para los análisis recurrentes de precipitaciones extremas en la provincia de Chimborazo. se utilizaron los registros de precipitaciones desde el 1 de enero de 1990 hasta el 31 de diciembre del 2018 proporcionadas por el GEAA-INHAMI de 23 estaciones meteorológicas. se realizó un análisis exploratorio de cada una de las estaciones, donde se observó pocas mediciones en la estación de Palmira y Chambo las cuales fueron retiradas. El índice de Moran indica que no existe autocorrelación espacial al 95% de confiabilidad. Se identificaron dos tipos de umbrales uno fijo para toda el área de estudio y otro específico para cada una de las estaciones meteorológicas, para esto se utilizaron las técnicas del cuantil y del gráfico de vida residual media. Con el remuestreo de Bootstrap se halló las distribuciones muestrales, para posteriormente estimar los parámetros de escala y forma de la distribución de Pareto generalizada en intervalos de confianza. finalmente se identificó las precipitaciones extremas y sus valores en riesgo, mismos que fueron representado a través de mapas con interpolación IDW. El uso de umbrales específico en cada estación permitió caracterizar el comportamiento de precipitaciones extremas y su valor en riesgo en forma localizada. Para todo el proceso de análisis de utilizó herramientas de R.

Palabras clave: <TEORÍA DE VALORES EXTREMOS>, <BOOTSTRAP>, <DISTRIBUCIÓN DE PARETO GENERALIZADA>, <PRECIPITACIONES>, <INTERPOLACIÓN IDW>.

0848-DBRA-UPT-2023



ABSTRACT

In recent years, climate change has been a trigger for the stochasticity of the behavior of meteorological variables. The province of Chimborazo, having microclimates presents precipitations with specific localized characteristics. Recent research studies were interested in the behavior of rainfall in the city of Riobamba, initiating the use of extreme value theory from a univariate approach in said area, that is, using the records of a weather station. For the purpose of giving continuity to this research this thesis work proposes strategies by recurring analysis of periods of extreme rainfall in the province of Chimborazo. Records of rainfall from January 1, 1990 to December 31, 2018 provided by the GEAA-INHAMI of 23 meteorological stations were used. An exploratory analysis of each station was done, where, few measurements were found at the Palmira and Chambo station thus they were withdrawn. The Moran index indicates that there is no spatial autocorrelation at 95% reliability. Two types of thresholds were identified, one fixed for the entire area of study and a specific one for each meteorological station, for this purpose the techniques of the quantile and the mean residual life graph. With Bootstrap resampling the sampling distributions were found, to later estimate the parameters of scale and shape of the generalized Pareto distribution in confidence intervals. finally identified the extreme rainfall and its values at risk, which were represented through maps with IDW interpolation. The use of specific thresholds in each station allowed to characterize the behavior of extreme rainfall and its value at risk in a localized manner. for all the analysis process R tools were used.

Keywords: <EXTREME VALUE THEORY>, <BOOTSTRAP>, <GENERALIZED DISTRIBUTION PARETO>, <PRECIPITATION>, <IDW INTERPOLATION>.



Edgar Mesias Jaramillo Moyano
0603497397

INTRODUCCIÓN

En las últimas décadas, los cambios climáticos en el planeta se han acelerado debido al calentamiento global, como consecuencia, ha surgido el aumento de la estocasticidad en las mediciones meteorológicas, entre estas las precipitaciones, Según (Aguayo, et al, 2020, p.158) este desorden climático ha causado varios problemas a nivel mundial, como deshielos en los polos árticos, poniendo en peligro la existencia de muchas especies animales y vegetales, afectando su hábitat natural.

Los métodos estadísticos adecuados para analizar el comportamiento espacial de las variables meteorológicas como por ejemplo la temperatura, la precipitación y la velocidad del viento son de suma importancia, especialmente para la implementación de acciones que reduzcan los efectos negativos debido a sus comportamientos extremos (Blanco, et al, 2014, p.27-34).

El desarrollo de la teoría de valores extremos aplicada en los diversos campos de la ciencia tiene alrededor de 80 años de investigación, aunque es más antigua. Se han realizado aplicaciones como, por ejemplo, en la astronomía donde es utilizada para tratar observaciones atípicas en sus proyectos. Según (Ibañez, 2011, p.1)), no fue hasta 1950 cuando se propuso una metodología seria para modelizar sucesos extremos, siendo las primeras aplicaciones en el campo de la ingeniería civil debido a la necesidad de diseñar estructuras que soporten distintas fuerzas, especialmente las de la naturaleza.

La teoría de valores extremos es una rama de la estadística enfocada en el comportamiento de las observaciones máximas o mínimas de una variable aleatoria, esta técnica atrae a la comunidad científica por sus diversas aplicaciones en áreas como: finanzas, ingeniería ambiental, comunicación, salud, longevidad, resistencia de materiales, meteorología, entre otros (Chávez, 2009, p.17). Dicha teoría busca obtener modelos estadísticos que describan el comportamiento de las observaciones extremas utilizando varias técnicas. El Ecuador es un país que cuenta con cuatro regiones naturales que son: Costa, Sierra, Oriente y la Región Insular o Galápagos. La región sierra posee pisos climáticos que oscilan entre los 1.000 y 4.000 metros sobre el nivel del mar, esto conlleva a poner atención en zonas específicas, está conformada por una serie de cadenas montañosas que conectan los principales ramales de la cordillera de los Andes, sobresaliendo el nudo de Tiocajas, límite de las hoyas del Chambo y el Chanchán. Dando lugar a diferentes ambientes que van desde el glacial en lo más elevado de las cumbres, y cálido hasta el subtropical en las tierras más bajas. En la zona centro de esta región se ubica la provincia de Chimborazo, en donde se encuentra las elevaciones más importantes del país, como el Chimborazo con una altura de 6.310 m.s.n.m., el Altar con 5.320

m.s.n.m. y el Carihuairazo con 5.020 m.s.n.m.(INPC, 2010, p.21)), lo que contribuye a la existencia de microclimas. Fue reconocida el 25 de junio de 1824 por la Ley de División Territorial de la Gran Colombia, limitando al norte con la provincia de Tungurahua, al sur con Cañar, al este con Morona Santiago y al oeste con Guayas y Bolívar. Su territorio se divide en diez cantones: Alausí, Colta, Chunchi, Guamote, Guano, Penipe, Pallatanga, Chambo, Cumandá y Riobamba (INPC, 2010, p.21). Este trabajo de tesis pretende caracterizar las precipitaciones extremas en la provincia de Chimborazo y establecer estrategias para su análisis recurrente. Se utiliza la teoría de valores extremos mediante la distribución de Pareto generalizada, remuestreo de bootstrap, comportamiento espacial, interpolación espacial e identificación de los valores en riesgo. Para este estudio se dispuso de los registros de 23 estaciones meteorológicas proporcionadas por el GEEA-INHAMI del 1 de enero de 1990 hasta el 31 de diciembre del 2018.

Este manuscrito contiene cuatro que se detallan a continuación:

El capítulo I contiene el marco teórico referencial en el que se expone los antecedentes, el problema, la justificación y los objetivos planteados.

El capítulo II contiene el marco teórico donde se describe las conceptualizaciones relacionadas a la precipitación, análisis espacial, remuestreo de Bootstrap, las técnicas estadísticas que engloban la teoría de valores extremos y las principales librerías de R utilizadas.

El capítulo III contiene el marco metodológico de la investigación este describe la metodología investigativa, la localización de la zona de estudio y las principales técnicas estadísticas previas al análisis de datos.

El capítulo VI contiene el marco de análisis e interpretación de resultados, en el que se explica el análisis exploratorio la identificación de umbrales, estimación de parámetros, distribuciones muestrales, interpolación de precipitaciones extremas e identificación de los valores en riesgos.

Finalmente se exponen las conclusiones, recomendaciones, la bibliografía actualizada y el código R utilizado.

CAPÍTULO I

1. MARCO TEÓRICO REFERENCIAL

1.1. Planteamiento del problema

1.1.1. *Enunciado del problema*

La literatura limitada referente a la caracterización y descripción metodológica del comportamiento de las precipitaciones extremas en la provincia de Chimborazo ha generado un déficit de información, especialmente para los proyectos de monitoreo de lluvias que realiza el Grupo de Energías Alternativas y Ambiente (GEAA) conjuntamente con el INAMHI. Así como también para la toma de decisiones realizadas por los organismos gubernamentales competentes y la población en general en cuanto a planes de contingencia para hacer frente a las lluvias extremas. Incrementando de este modo la probabilidad de no estar debidamente preparados para reducir por ejemplo los daños ocasionados por deslaves, deslizamientos de tierra, desborde de ríos y quebradas, inundaciones en las ciudades, erosiones e infertilidad del suelo, sequía de las plantaciones, desaparición de ríos.

1.2. Justificación

El trabajo de tesis " Teoría de valores extremos para el análisis de la precipitación de la estación meteorológica ESPOCH (1976-2019) " enfoca la teoría de valores extremos en datos de precipitación de una sola estación meteorológica, es decir, en forma univariante, sin considerar precipitaciones de otros lugares y mucho menos identificar algún tipo de dependencia espacial. Motivados por estas deficiencias, esta investigación pretende proponer estrategias para análisis recurrentes de las precipitaciones extremas y su interacción espacial, en base a los registros en las estaciones meteorológicas distribuidas en zonas estratégicas de la provincia de Chimborazo. Para ello se abordan los principios fundamentales de la teoría de valores extremos, identificación de umbrales tanto fijos como variables, método Bootstrap para estimar los parámetros de escala y forma de la distribución de Pareto generalizado en intervalos de confianza.

1.3. Objetivos

1.3.1. *Objetivo general*

Generar estrategias para el análisis estadístico recurrente de las precipitaciones extremas en la provincia de Chimborazo con registros desde 1990 al 2018, mediante la teoría de valores extremos.

1.3.2. *Objetivos específicos*

- Realizar un análisis exploratorio general y espacial de los registros de precipitaciones.
- Identificar los umbrales fijos y variables de precipitaciones en la provincia de Chimborazo.
- Utilizar el método Bootstrap para encontrar las distribuciones muestrales y estimación de los parámetros de escala y de forma de la distribución de Pareto generalizado en intervalos de confianza.
- Generar estrategias para el análisis recurrente de las precipitaciones extremas en la provincia de Chimborazo, en función de la identificación del tipo de umbral que proporciona mejores resultados.
- Elaborar al menos un mapa de interpolación espacial, que represente el comportamiento de las precipitaciones extremas en la Provincia de Chimborazo.

CAPÍTULO II

2. MARCO TEÓRICO

2.1. Características atmosféricas

La troposfera

La troposfera se extiende hasta un límite superior llamado tropopausa, a 9 kilómetros de los polos y a 18 kilómetros del Ecuador. Tiene importantes movimientos verticales y horizontales de masas de aire (vientos) y es relativamente abundante en agua debido a su proximidad a la hidrosfera. Sin embargo, es la zona de las nubes y de los fenómenos climáticos: lluvia, viento, cambios de temperatura, etc. Esta es la capa más interesante de la meteorología. En la troposfera, la temperatura sube hasta alcanzar el límite superior de -70°C (Sarochar, 2009, p.3).

La Estratosfera

La estratosfera sigue a la troposfera y alcanza el límite superior conocido como estratosfera, que tiene 50 km de altura. En esta capa, la temperatura sube a alrededor de 0°C en la estratosfera. Casi no hay movimiento de aire vertical, pero los vientos horizontales suelen alcanzar velocidades de 200 km/h. En esta parte de la atmósfera, a una distancia de 30 a 50 km, se encuentra el ozono, que juega un papel tan importante en la absorción de la dañina radiación de onda corta (Sarochar, 2009, p.3).

La mesosfera y la termosfera

La mesosfera y la termosfera se encuentran en la estratosfera. En él, el aire es tan delgado que la densidad es muy baja. Estos son los lugares donde aparecen las auroras y se reflejan las ondas de radio, pero su trabajo tiene poco efecto sobre los seres vivos (Sarochar, 2009, p.3).

2.2. Precipitación

La precipitación se produce por la caída directa de gotas de agua o de cristales de hielo, estas gotas son mayores cuanto más alta está la nube y más elevada es la humedad del aire, ya que se condensa sobre ellas el vapor de las capas que van atravesando. Durante el largo recorrido, muchas gotas llegan a juntarse, este fenómeno también se presenta cuando está en forma de cristales de hielo. Estas gotas caen en virtud de su peso, y lo hacen a una velocidad que varía entre 4 y 8 m/s, según sea su tamaño y la dirección del viento. La lluvia resulta del ascenso y enfriamiento del aire húmedo, ya que a menos temperatura no puede retener todo su vapor de agua, parte del cual se condensa rápidamente, lo que ocurre de varias formas (Sarochar, 2009, p.13).

La precipitación en países tropicales se ha caracterizado por tener variabilidad natural ya sea por estaciones o temporadas, los cambios en las precipitaciones han causado un efecto irreversible en el cambio climático donde en la inestabilidad de temperatura en muchos lugares (Ceballos, et al, 2013, p.13).

Los Pluviómetros son sensores que se encarga en recoger el líquido o sólido de la condensación del vapor de agua que cae de las nubes, estos instrumentos están ubicados en las estaciones meteorológicas repartidas en la provincia de Chimborazo (CEEA-ESPOCH, 2022, p.1).



Ilustración 2-1: Pluviometro

Fuente: CEEA-ESPOCH, 2022

2.2.1. Tipos de precipitación

Agua líquida: este es el tipo más común y es el resultado directo de caer de una nube. Ciertas ocasiones la precipitación comienza como hielo, pero se derrite a medida que desciende. Consiste en gotas de más de 0,5 mm de diámetro que caen a velocidades superiores a 3 m/s. Las nubes que la producen son nimboestratos y se denominan lluvias fuertes, moderadas o ligeras según su intensidad (Sarochar, 2009, p.14-17).

Lluvia general: son precipitaciones prolongadas (3 a 48 horas) y fuertes (10 a 200 mm) sobre un área grande, generalmente una provincia o todo el país. Es la lluvia que se produce cuando se elevan masas de aire cálido y húmedo (Sarochar, 2009, p.14-17).

Llovizna: por lo general, es de larga duración, pero de muy baja intensidad, no más de 1 mm por hora. La llovizna o lluvia fina se produce en las mismas condiciones que la precipitación normal, pero procede de masas de aire muy húmedas o no muy altas. Es una precipitación uniforme formada por gotas de agua de menos de 0,5 mm de diámetro que parecen flotar en el aire debido a su baja velocidad de caída (Sarochar, 2009, p.14-17).

Garúa: es una llovizna de niebla. Poco efecto sobre la visibilidad. Debido a su pequeño peso y tamaño, las gotas se mueven al azar en todas las direcciones, por lo que parecen húmedas por todas partes. La cantidad de agua que sedimenta es muy baja. Las garúas es a menudo llovizna, especialmente en agosto y septiembre, cuando la niebla suele tener un espesor de 300 a 400 metros (Sarochar, 2009, p.14-17).

Chaparrón o chubasco: son lluvias de corta duración, pero la intensidad es muy alta y puede alcanzar decenas de milímetros por hora (Sarochar, 2009, p.14-17).

Aguacero: es también una lluvia breve, pero a diferencia del chaparrón, mucho más abundante. Sus variedades más importantes son:

a) **Aguacero fuerte:** es una lluvia breve, muy fuerte, reducida a una pequeña zona.

b) **Aguacero torrencial:** es breve, pero fortísimo. Suele producir daños materiales.

Nieve: la nieve cae en copos más o menos grandes, que contienen estructuras cristalinas de varias formas, la mayoría de las veces tienen forma de estrella de seis puntas. La nieve se forma cuando la temperatura es lo suficientemente baja como para que el agua se congele (Sarochar, 2009, p.14-17).

Granizo: Partículas sólidas de hielo que caen de una nube. Estos corpúsculos suelen variar en tamaño desde unos pocos milímetros hasta dos o más centímetros. A diferencia de la nieve, que casi siempre se produce en invierno o en zonas con formación favorable de hielo, el granizo suele producirse tanto en verano como en invierno (Sarochar, 2009, p.14-17).

2.2.2. Medición de precipitación

Un pluviómetro es un instrumento utilizado por los centros de investigación meteorológica para recoger y medir la cantidad de lluvia que ha caído. La unidad de medida de la precipitación es la precipitación mm, que equivale a una hoja de papel de 1 metro cuadrado de superficie y 1 mm de espesor. Volumétricamente representa 1 litro de agua por metro cuadrado (Sarochar, 2009, p.4).

El mecanismo de precipitación de estas intensas partículas de hielo está relacionado con las tormentas, donde la convección interviene como elemento esencial en su formación, y con los fenómenos eléctricos (Sarochar, 2009, p.14-17).

2.3. Datos geográficos

2.3.1. Características de datos geográficos

- **Posición:** conjunto de coordenadas referenciadas en un contexto que pueda interpretarlas.
- **Atributos no espaciales:** describen al dato geográfico.
- **Relaciones espaciales:** muestran relaciones entre los datos geográficos.
- **Tiempo:** momento o período del dato geográfico (Cano & Trilla, 2007, p.7-8).

2.3.2. Modelos geográficos

Ráster

Es la posición del objeto, dividido en celdas, generalmente dentro de un cuadro definido por la posición exacta entre la fila y la columna medidas (Cano & Trilla, 2007, p.8).

Técnicas de compresión de datos en un ráster

- **Run length encoding:** agrupa las celdas adyacentes con un mismo valor.
- **Quadtree:** se divide en cuadrantes de igual tamaño, si el cuadrante no contiene el mismo valor se vuelve a subdividir en otros cuatro (Cano & Trilla, 2007, p.8).

Vector

Es la ubicación de los objetos geográficos que están ubicados en un sistema de referencia o también llamado plano cartesiano con el eje X y el eje Y (Cano & Trilla, 2007, p.9).

2.4. Efectos espaciales

2.4.1. Heterogeneidad espacial

Es la variación de las relaciones de las medidas en el espacio, la cual se difiere en 2 partes:

- **La inestabilidad estructural.**

La inestabilidad estructural es la falta de precisión en los datos medidos, ya que puede ser periódico o inestable según el lugar donde se mida, por lo que no será igual entre mediciones una al lado de la otra. Esto suele ocurrir cuando se utilizan para medidas extremas (Moreno & Vayá, 2000, p.211-288).

- **Heterocedasticidad.**

Aunque la heterocedasticidad ocurre cuando faltan o faltan datos, ya sea por causas naturales u otros errores, los errores existen (Moreno & Vayá, 2000, p-211-288).

2.4.2. Autocorrelación o dependencia espacial

La correlación espacial o autocorrelación es el resultado de la relación entre una característica que ocurre en un cierto punto en el espacio y una característica que ocurre en otro lugar de medición. Por tanto, podemos decir que habrá diferencias entre medidas y diferentes funciones entre medidas adyacentes, por lo que es fácil decir que no cumplen el supuesto de independencia, por lo que podemos decir que tenemos dependencia espacial, lo cual es imposible. cambiar la ubicación de los valores de las variables sin afectar la información de la muestra (Moreno & Vayá, 2000, p.211-288).

2.4.3. Tipos de autocorrelación espacial

Autocorrelación espacial positiva

Se dice que se tiene un fenómeno cuando existe en un área registrada o medida, lo que extiende ese fenómeno a otros lugares a su alrededor y así facilita la investigación (Moreno & Vayá, 2000, p.211-288).

Autocorrelación espacial negativa

Está presente cuando un fenómeno registrado cuenta con diferencia en la región y no se extiende a otras regiones vecinas por lo que impide su aparición (Moreno & Vayá, 2000, p.211-288).

2.4.4. Medición de autocorrelación espacial

La estadística I-Moran es una medida global de autocorrelación espacial que indica si existe autocorrelación en todas las unidades geográficas de la muestra. Las estadísticas de los indicadores de autocorrelación espacial local (LISA) se pueden utilizar en la estadística de I-Moran. Esta métrica se puede usar para identificar grupos de valores altos (o bajos) para una variable por ubicación (Aguayo & Mendellín, 2014, p.72-75).

2.5. Interpolación espacial

La interpolación espacial es una técnica matemática utilizada para predecir el valor de un atributo en una ubicación precisa a partir de valores obtenidos de puntos vecinos que se encuentran dentro de la misma área de estudio. La interpolación se utiliza para transformar un número finito de

observaciones, como elevaciones, en un continuo espacial de modo que su patrón espacial sea comparable al representado por las observaciones del punto base (Miranda & Condal, 2016, p.4).

2.6. Interpolación de IDW

El método IDW es similar al Kriging ordinario, ya que da más peso a los valores cercanos a un punto, pero posee una menor complejidad del cálculo, El IDW utiliza un algoritmo simple basado en distancias. Diversos autores han comparado el interpolador Kriging con el IDW en condiciones no tropicales (Villatoro, et al, 2008, p.25-30).

2.7. Teoría de valores extremos (TVE)

La teoría de valores extremos tomó importancia en los años 20 con problemas relacionados principalmente con la hidrología y dieron lugar al primer teorema fundamental de Fisher-Tippet (1928) y Gnedenko (1948) que caracteriza la distribución asintótica del máximo observado (Serra, et al, 2013, p.20).

En la literatura técnica se han propuesto varios procedimientos para llevar a cabo el proceso de identificación de distribuciones de valores extremos, que no requieren estimar los parámetros de estas distribuciones (Villaseñor, 1997, p.57).

Esta teoría se desarrolla a partir de los resultados de Fréchet (1927), Fisher y Tippett (1928) y von Mises (1936), que sientan las bases de la teoría probabilística de valores extremos unidimensional; sin embargo, su desarrollo teórico no tendrá lugar hasta principios de los años 70 con la tesis doctoral de Haan (1970) y los resultados de inferencia estadística desarrollados por Pickands (1975); estos resultados supusieron las primeras contribuciones a la Teoría de Valores Extremos Multivariante y motivaron el desarrollo de modelos alternativos basados en excedencias de umbrales (Medialdea, 2016, p.30-36).

2.7.1. Modelos máximos por bloques

La familia de distribuciones Valores Extremos será útil para modelizar la distribución de los máximos por bloques. El procedimiento consiste en agrupar los datos en bloques de igual tamaño y, a continuación, ajustar la distribución de valores extremos al conjunto de los máximos correspondientes a cada uno de los bloques. El principal problema que presenta este método reside en la elección del tamaño de los bloques, para la cual habrá que encontrar un equilibrio entre el sesgo y la varianza (Medialdea, 2016, p.13).

2.7.2. Umbral

Es el valor mínimo o máximo de alguna medición en investigación ya sea variables meteorológicas o económicas (Aritizábal, et al, 2011, p.98).

2.7.3. Picos sobre el umbral

En el método de picos sobre un umbral se consideran como valores extremos aquellos que sobrepasen un cierto nivel o umbral dado, es el de picos sobre un umbral o POT.

Se dice que X es una excedencia sobre u sí $X > u$. En ese caso al valor $Y = X - u$ se le conoce como el exceso de X sobre u . Para este modelo se hace uso de la función de distribución de los excesos (F_u). Se define a la función de distribución de los excesos sobre por la probabilidad condicional (Avilés, et al, 2014, p.2).

$$F_u = P(X - u \leq y | X > u) \quad , \quad 0 \leq y \leq w_F - u$$

Esta función también es llamada la función de distribución de excesos condicional, pues reduce el espacio muestral a aquellas variables que sobrepasan el nivel u . La distribución relacionada con el modelo de picos sobre un umbral es la Distribución Generalizada de Pareto (DGP) la cual se define como:

$$G_{\xi, \beta}(y) = 1 - \left(1 + \frac{\xi(y)}{\beta}\right)^{-\frac{1}{\xi}}$$

$$\xi = \text{Parametro de Escala}$$

$$\beta = \text{Parametro de Forma}$$

Para $y > 0$ cuando $\xi > 0$; y para $0 \leq y \leq -\frac{\beta}{\xi}$.

$$H_{\xi, \beta}(y) = 1 - \exp\left\{-\frac{y}{\beta}\right\} \quad , \quad \xi = 0$$

Uno de los principales teoremas dentro de la teoría de valores extremos es el propuesto por Pickands, Balkema y de Haan (McNeil 2005) el cual relaciona al modelo de picos sobre un umbral a la Distribución de Pareto generalizada (DGP).

El teorema dice que si la función de distribución, entonces la DGP es la distribución $FED(G_{\xi, u, \sigma})$ límite de los excesos cuando el umbral u tiende a w_F . Es decir, si F_u es la función de distribución de los excesos sobre u entonces:

$$\lim_{x \rightarrow w_f} \sup_{0 \leq y < w_{F-u} | F_u(y) \cdot H_{\xi, \beta}(y) | = 0$$

Si y sólo si $F \in D(G)$. Este resultado es la base para la aproximación de la distribución de los excesos sobre umbrales suficientemente altos mediante la DGP. El teorema nos da bases teóricas para esperar que, se elige un umbral adecuado, las observaciones que se encuentren por encima mostrarán un comportamiento tipo Pareto (Avilés, et al, 2014, p.3-8).

Se conocen varias propiedades de la DGP, una de ellas es sobre su función media de excesos que se define como:

$$e(u) = E(X - u | X > u)$$

Es decir, la esperanza o media de los excesos sobre μ condicionada a las excedencias. Si una variable aleatoria X sigue una DGP, entonces su función media de excesos tiene la siguiente expresión (Avilés, et al, 2014, p.3-8).

$$e(u) = \frac{\beta + \xi \mu}{1 - \xi}$$

El problema inicial es la elección de un umbral adecuado para que la aproximación de F_u puede hacerse mediante la DGP, el cual no debe ser elegido de manera arbitraria. Sin embargo, no existe un método para estimarlo puntualmente, por lo que esta tarea requiere de un análisis cuidadoso, ya que, de elegirse un umbral demasiado bajo, se violaría la suposición del Teorema, y un umbral demasiado alto, dejará muy pocas observaciones con las cuales pueda hacerse la inferencia sobre la DGP. Existen diversas herramientas disponibles que ayudan a elegir un umbral adecuado, la cual es la gráfica de medias de excesos que consiste en los puntos de la forma:

$$\left\{ \left(u, \frac{1}{n_u} \sum_{i=1}^{n_u} x_{(i)} \right); u < x_{max} \right\}$$

Donde $x_{(1)}, x_{(2)}, x_{(3)}, \dots, x_{(n)}$ son las n_u excedencias observadas sobre u , y max es el valor máximo de las $x_{(n)}$. Es decir, esta gráfica relaciona a cada umbral con la esperanza empírica de los excesos sobre dicho umbral. Por la propiedad de la DGP antes mencionada, se sabe que sobre un umbral para el cual la DGP proporcione una aproximación válida a la distribución de los excesos, la función media de excesos debería comportarse de manera lineal en u (Avilés, et al, 2014, p.3-8).

2.7.4. Modelos de excedencias de umbrales

Mediante el estudio (Medialdea, 2016, p.13) indica que modelizar únicamente máximos de bloques es una aproximación poco eficiente en el análisis de valores extremos si algunos de los bloques contienen eventos más extremos que el resto. En este caso, será adecuado el uso de modelos de excedencias de umbrales.

2.8. Muestreo de Bootstrap

Los procedimientos basados en los métodos Bootstrap implican obviar los supuestos sobre la distribución teórica que siguen los estadísticos. En su lugar, la distribución del estadístico se determina simulando un número elevado de muestras aleatorias construidas directamente a partir de los datos observados. Es decir, utilizamos la muestra original para generar a partir de ella nuevas muestras que sirvan de base para estimar inductivamente la forma de la distribución muestral de los estadísticos, en lugar de partir de una distribución teórica asumida a priori (Araya, 2011, p.32).

Este enfoque tiene su antecedente inmediato en las técnicas de simulación Monte Carlo, las cuales consisten en extraer un número elevado de muestras aleatorias de una población conocida, para calcular a partir de ellas el valor del estadístico cuya distribución muestral pretende ser estimada (Araya, 2011, p.30-32).

La Ilustración 2-2 muestra el proceso de estimación de estadísticos Bootstrap a partir de una muestra. Consideremos que disponemos de una muestra (X) a partir de la cual calculamos un estadístico de interés que estime algún parámetro poblacional ($f(X)$) (Alpízar, 2011, p.45).

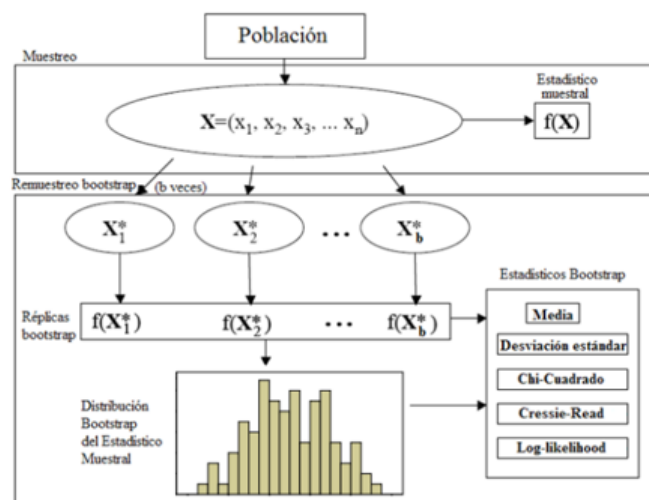


Ilustración 2-2: Proceso de estimación Bootstrap

Fuente: Araya, 2011.

La publicación en 1979 del primer trabajo de Bradley Efron sobre los métodos Bootstrap constituyó uno de los sucesos de mayor relieve en el campo de la estadística en la década de los años 80. La idea de reemplazar complicadas aproximaciones y a veces poco precisas, a sesgos, varianzas y otras medidas de incertidumbre por simulaciones con el ordenador, fue ampliamente utilizada por investigadores teóricos y usuarios de los métodos estadísticos. Pasado el inicial escepticismo sobre la viabilidad de esta metodología, los investigadores empezaron a poner de manifiesto que estos métodos en muchos casos aproximan mejor que los convencionales (Montesdeoca, 2003, p.45).

El Bootstrap fue concebido por Efron en el contexto de diseños de tamaño fijo y su planteamiento del Bootstrap es muy simple. Dada una muestra aleatoria $X = (X_1, \dots, X_n)$ de una distribución de probabilidad F , el objetivo es estimar la distribución de probabilidad de una variable aleatoria especificada $R = R(X, F)$, posiblemente dependiente de X y de la distribución de probabilidad desconocida, utilizando los datos observados. Estas cantidades R son a menudo sesgos, varianzas, pivotaes o test estadísticos, cuyas distribuciones de probabilidad exacta son difíciles de determinar, o el error de la aproximación es muy considerable.

La aproximación de sumas de variables aleatorias independientes y con varianza finita por la distribución normal puede tener errores en la aproximación del orden $n^{-1/2}$.

La aproximación Bootstrap a la distribución probabilística de R , dada por Efron en el referido trabajo de 1979 es como sigue:

- Construir a partir de los datos $X = (X_1, \dots, X_n)$ la distribución empírica $\widehat{F}_n = \frac{1}{n} \sum_{i=1}^n I(X_i \leq t)$.
- Extraer una muestra aleatoria de tamaño n , $X^* = (X_1^*, \dots, X_n^*)$ de la distribución empírica $\widehat{F}_n(t)$, la cual recibe el nombre de distribución Bootstrap. La muestra Bootstrap, a diferencia del jackknife, es aquí una muestra aleatoria con reemplazamiento del conjunto de datos observados $X = (X_1, \dots, X_n)$. En el jackknife, se seleccionan muestras aleatorias sin reemplazamiento de tamaño $n - 1$.
- La distribución de $R(X, F)$ se aproxima finalmente por la distribución Bootstrap de $R^* = R(X^*, F)$.

Parece claro que la distribución de probabilidad de R^* puede calcularse de la distribución empírica en el mismo modo que la distribución de R se calcula de la distribución original F . Ahora bien, el recurrir al Bootstrap en líneas generales se deberá a la dificultad de obtener la distribución exacta de R . La clave del método Bootstrap consiste en que, tomando repetidas muestras de la distribución empírica mediante el método de Monte Carlo, se obtiene una muestra de valores de R que en si

misma constituye la aproximación Bootstrap buscada. Naturalmente, esto es posible gracias a la capacidad de los ordenadores modernos.

Precisamente los métodos Bootstrap reciben también el nombre de métodos de computación intensiva. La idea de generar nuevos datos a partir de los datos observados recordaba al Barón de Munchausen, cuando estando en el fondo de un lago logró salir tirando el mismo de los cordones de sus botas (Montesdeoca, 2003, p.46).

2.9. Distribución de remuestreo

La distribución empírica se utiliza en el algoritmo dado por Efron como distribución de remuestreo por ser un estimador consistente de la verdadera distribución generadora de los datos. Para diseños de tamaño fijo, $\widehat{F}_n(t)$ tiene las siguientes propiedades como estimador de $F(t)$.

- $1E \left[\widehat{F}_n(t) \right] = F(t)$, para cualquier valor t .
- $var \left(\widehat{F}_n(t) \right) = \frac{F(t)(1-F(t))}{n}$
- $3 \left\| \widehat{F}_n - F \right\| \rightarrow 0$, para $n \rightarrow \infty$ (Teorema de Glivenko – Cantelli)

Estas propiedades en general no se mantienen cuando los datos se generan a través de un diseño secuencial (Montesdeoca, 2003, p.44-47).

Es interesante destacar la forma que adopta la distribución empírica para datos binarios.

Puede alternativamente considerarse también como distribución de remuestreo estimadores de núcleo de la función de distribución, los cuales tienen la forma:

$$\widehat{F}_n(x, h) = \frac{1}{n} \sum_{i=1}^n W \left(\frac{x - X_i}{h} \right).$$

2.10. Software R para analizar datos espaciales en R

R es un lenguaje de programación moderadamente complejo, de libre acceso, centrado en el análisis estadístico. R es el resultado de los esfuerzos concertados de muchas personas que, desde mediados de 1977, han mejorado y desarrollado el lenguaje a través de un entorno de programación integrado (IDE) diseñado específicamente para R (Pacheco, 2016, p.2).

2.10.1. *Librerías para datos espaciales*

Es importante el uso de software para el estudio y representación gráfica de datos geográficos, y el software estadístico utilizado es R, el cual contiene librerías óptimas para el análisis de variables meteorológicas (Vargas, 2021, p.12-20).

Para visualizar datos espaciales tenemos las librerías mas comunes son :

- *Maptools*
- *Google Vis*
- *Ggplot*
- *Ggmap*
- *Extremes*

Para el análisis con valores extremos la librería mas común es *extremes* que contiene un conjunto de funciones diferentes para el análisis de valores extremos. Una característica importante es evitar que se produzcan picos en bloques largos o umbrales altos (Vargas, 2021, p.12-20).

El código R para obtener el archivo shapefile, de forma importante para convertirlo en un objeto tipo data frame. Usando las funciones *readOGR* y *spTransform* de la librería shapefile de forma *rgdal* para leer los datos geoespaciales, También requería la función *gpclibPermit* de la biblioteca *Maptools* (Pacheco, 2016, p.4).

CAPÍTULO III

3. MARCO DE METODOLOGICO

3.1. Enfoque de investigación

Esta investigación es de tipo cuantitativo, debido a que se estudió los registros de precipitaciones que están registrados en valores numéricos y su unidad de medida son milímetros. Según el objetivo la investigación fue de tipo aplicativo puesto que se aplica la teoría a los datos de precipitaciones extremas. según el objetivo la investigación fue de tipo explicativa, puesto que pretende explicar el comportamiento de precipitaciones extremas en la provincia de Chimborazo. Según la manipulación de la variable fue no experimental puesto que no se ha hecho ninguna manipulación en sus valores. Según el periodo temporal es de tipo transversal, puesto que se estudia las mediciones de precipitaciones registradas durante el periodo temporal (1990-2018).

3.2. Nivel de investigación

Se realizó un análisis estadístico exploratorio, se aplicó la teoría de valores extremos en la variable precipitación, donde se identificó un umbral fijo para todas las estaciones y un umbral específico en cada una de ellas, mediante las técnicas del cuantil y el gráfico de vida residual media. Se halló las distribuciones muestrales mediante el remuestro de Bootstrap lo que permitió el cálculo de los parámetros de escala y de forma de la distribución de Pareto generalizado con intervalos de confianza. Se identifica la dependencia espacial mediante el índice de Moran. se halló los valores en riesgos de precipitaciones y finalmente se construyen mapas de interpolación para representarlos. El estudio se realizó en R, principalmente con las librerías POT y raster.

3.3. Diseño de la investigación

3.3.1. *Según la manipulación o no de la variable independiente*

Según el diseño de investigación, este trabajo de tesis es cuantitativo no experimental debido a que la variable principal de estudio que es la precipitación no se manipulo de ningún modo.

3.4. Tipo de estudio

Según el tipo de estudio, esta investigación es de campo, debido a que los datos de precipitaciones en las 23 estaciones, son datos reales registradas desde el 1 de enero de 1990 hasta el 31 de diciembre del 2018 en la provincia de Chimborazo.

3.5. Población, planificación, selección y calculo del tamaño de la muestra

3.5.1. Población de estudio

Las precipitaciones registradas en las 23 estaciones meteorológicas instaladas en zonas estratégicas de la Provincia de Chimborazo, monitoreadas por el GEAA-INAMHI, desde el 1 de enero de 1990 hasta el 31 de diciembre del 2018, siendo un total de 164041 datos.

3.5.2. Tamaño de la muestra

La investigación se realizó con todos los datos del GEAA-INHAMI siendo un total de 164041.

3.5.3. Localización del proyecto

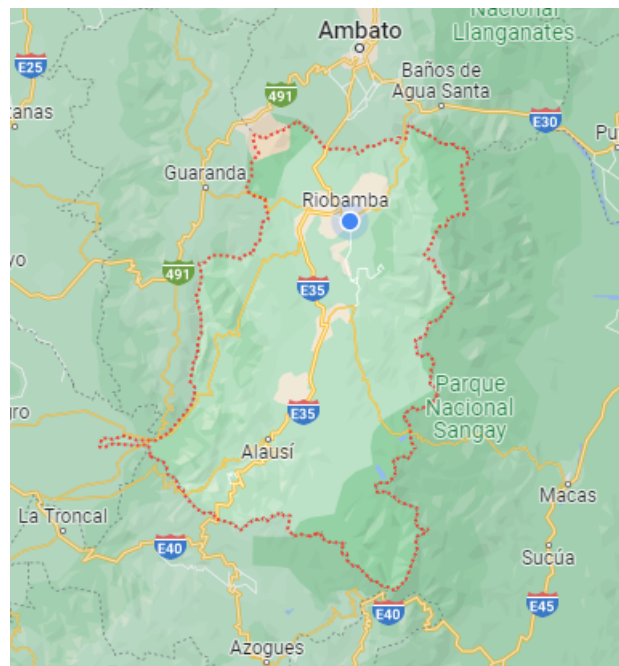


Ilustración 3-1: Mapa de la provincia de Chimborazo

Fuente: Google maps

La investigación se lo llevo a cabo con datos de las 23 estaciones meteorológicas ubicadas en la provincia de Chimborazo, monitoreadas por el INAMHI y el Grupo de Energías Alternativas y Ambiente (GEAA) perteneciente a la ESPOCH en la Facultad de Ciencias.

3.6. Métodos, técnicas e instrumentos de investigación

3.6.1. *Métodos de investigación*

El tipo de inferencia inductiva-deductiva esta dado por el estudio y la observación, el comportamiento espacial de los valores extremos en la provincia de Chimborazo. Se analizó las precipitaciones desde el 1 de enero de 1990 hasta el 31 de diciembre del 2018 en las 23 estaciones meteorológicas.

3.6.2. *Técnicas de investigación*

- **Técnicas de recolección de datos**

Debido que los datos son proporcionados por el GEEA-INHAMI, se realizaron entrevistas a los técnicos encargados de monitorear y registrar las precipitaciones.

- **Método de muestreo**

No se aplica ninguna tecnica de muestreo debido a que se analizan todos los registros proporcionados por e GEAA-INHAMI..

3.6.3. *Instrumentos de investigación*

Los datos proporcionados por el GEAA y el INAMHI son precipitaciones en forma digital (archivos excel) registrados en los dataloger instalados en cada estación meteorológica,

CAPÍTULO IV

4. MARCO DE ANÁLISIS E INTERPRETACIÓN DE RESULTADOS

4.1. Análisis exploratorio

4.1.1. Identificación geográfica de la zona en estudio

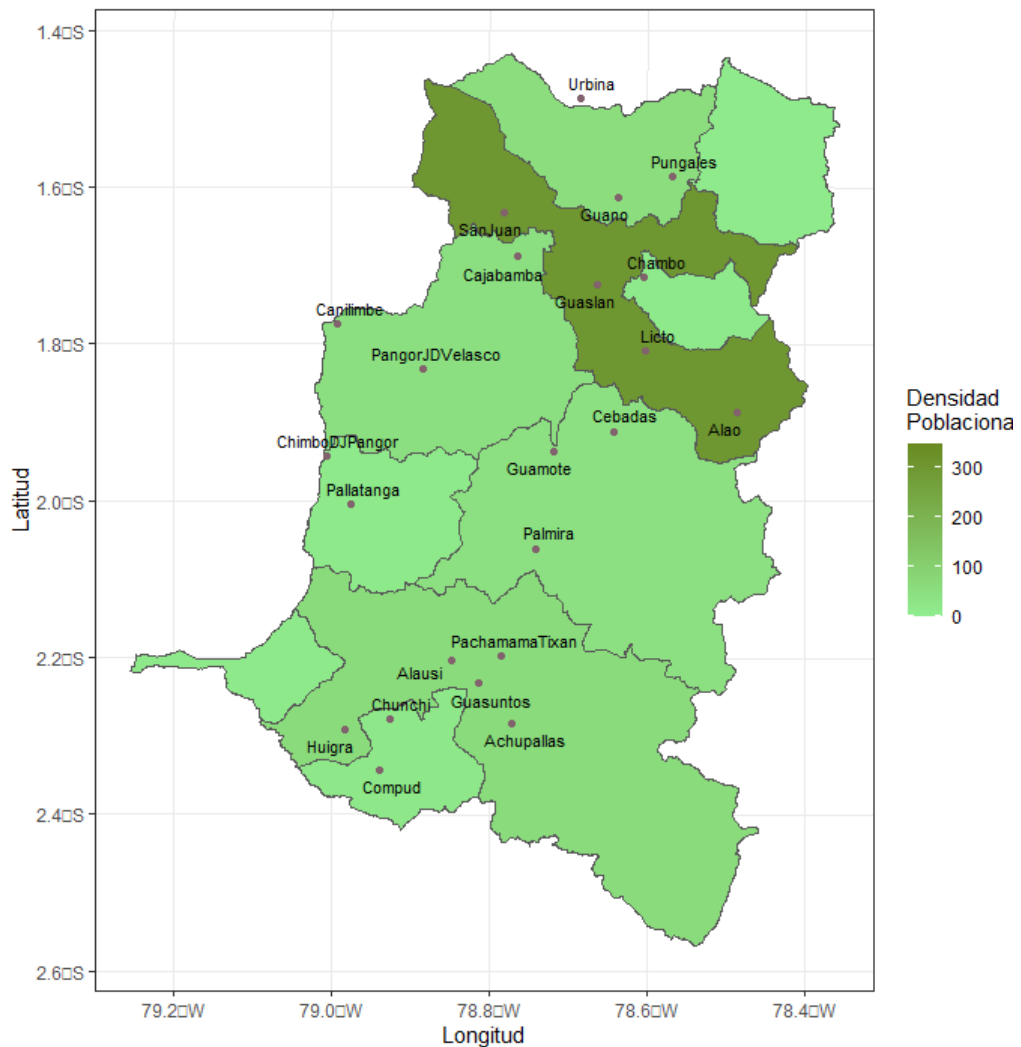


Ilustración 4-1: Estaciones meteorológicas en la provincia de Chimborazo

Fuente: Grupo de energías alternativas y ambiente (GEAA-INHAMI)

En la (Figura 4-1) se muestra la densidad poblacional de los diferentes cantones mediante el degradado de colores, donde se observa mayor número de población en el cantón Riobamba según el INEC 2021.

Tabla 4-1: Ubicación geográfica de las diferentes estaciones meteorológicas en la provincia de Chimborazo

CÓDIGO	ESTACIÓN	LATITUD	LONGITUD	ALTITUD
E1	Guaslán	-1,72	-78,66	2850,00
E2	Guamote	-1,93	-78,72	3020,00
E3	Pachamama-Tixán	-2,19	-78,78	3690,00
E4	Chunchi	-2,27	-78,92	2177,00
E5	Pungales	-1,58	-78,56	2550,00
E6	Urbina	-1,48	-78,68	3610,00
E7	Pallatanga	-2,00	-78,97	1523,00
E8	Huigra	-2,28	-78,98	1255,00
E9	San Juan - Chimborazo	-1,63	-78,78	3220,00
E10	Cajabamba	-1,68	-78,76	3160,00
E11	Cebadas	-1,90	-78,64	2930,00
E12	Alao	-1,88	-78,48	3200,00
E13	Compud	-2,34	-78,93	2402,00
E14	Palmira	-2,06	-78,73	3180,00
E15	Achupallas	-2,28	-78,76	3178,00
E16	Chimbo DJ Pangor	-1,94	-79,00	1452,00
E17	Alausi	-2,19	-78,85	2267,00
E18	Cañi-limbe	-1,77	-78,99	2800,00
E19	Guasuntos	-2,23	-78,81	2438,00
E20	Chambo	-1,71	-78,60	2640,00
E21	Licto	-1,80	-78,60	2840,00
E22	Guano	-1,61	-78,63	2620,00
E23	Pangor- JD Velasco	-1,82	-78,88	3109,00

Fuente: Grupo de energías alternativas y ambiente (GEAA-INMAHI)

Realizado por: Juan C, Hidalgo LL, 2023

4.1.2. Identificación de datos

La ESPOCH y el Instituto Nacional de Meteorología e Hidrología (INAMHI), conjuntamente han trabajado para desarrollar proyectos con el grupo de investigación GEAA, donde se ha logrado la medición automática de precipitaciones en diferentes lugares de la provincia de Chimborazo.

Se toman los datos de precipitaciones en las 23 estaciones meteorológicas, información proporcionada por el GEAA-INHAMI, desde el 1 de enero de 1990 hasta el 31 de diciembre del 2018.

La precipitación esta medida en milímetros cuando su estado es líquido, donde el rango operacional de las dimensiones observadas tiene una cantidad total de agua líquida que va de 0 a 400 milímetros, la toma de la muestra es automática, registrando un total de precipitación y acumulando toda la lluvia que cae durante el día.

La posición de este instrumento es a una altura del suelo de 1,50 metros sobre el suelo para que no le llegue la humedad, el mantenimiento del pluviómetro en las estaciones meteorológicas no requiere de mano de obra especializada, es decir es suficiente con un mantenimiento respectivo para evitar incidentes en las mediciones.



Ilustración 4-2: Estación ubicada en la ESPOCH

Fuente: CEEA-ESPOCH, 2022

Tabla 4-2: Precipitaciones

Estación	Guaslán (E1)	Guamote (E2)	Tixán (E3)	Chunchi (E4)	...	Chambo (E20)	Licto (E21)	Guano (E22)	Pangor DJV (E23)
1/1/1990	0	0	0	0	...	0	3	0	0
2/1/1990	0	0	0	0	...	0	3	0	0
3/1/1990	0	0	0	0,4	...	0	5,1	0	0
.	.	.	⋮
.
.
31/12/2018					...				

Fuente: Grupo de Energías Alternativas y Ambiente (GEAA-INHAMI)

Realizado por: Juan C, Hidalgo LL, 2023

Para facilitar el manejo de las estaciones se codifica a cada estación con la letra E y su número correspondiente, cada estación meteorológica tiene sus propias características debido a las condiciones y mantenimientos técnicos de los sensores. No se enfatiza el análisis temporal.

4.1.3. Identificación de variable en estudio

Tabla 4-3: Número de registros de precipitación en cada estación

ESTACIÓN	NÚMERO DE REGISTRO	ESTACIÓN	NÚMERO DE REGISTRO
Guaslán	7604	Compud	8994
Guamote	2260	Palmira	850
Pachamama-Tixán	3168	Achupallas	9714
Chunchi	8884	Chimbo D.J. Pangor	8995
Pungales	3009	Alausí	9572
Urbina	5715	Cani-limbe	9689
Pallatanga	8741	Guasuntos	9689
Huigra	5505	Chambo	134
San Juan	10452	Licto	9081
Cajabamba	3429	Guano	9917
Cebadas	9463	Pangor D.J. Velasco	9825
Alao	9382		

Fuente: Grupo de Energías Alternativas y Ambiente (GEAA-INMAHI)

Realizado por: Juan C, Hidalgo LL, 2023

El número de registro de precipitación tomadas por el GEEA-INHAMI son distintos en cada estación meteorológica. Según la (Tabla 4-3) San Juan (10452) y Guano (9917) respectivamente tienen el mayor número de registros, mientras que Palmira y Chambo con 850 y 134 registros respectivamente presentan un numero menor datos recolectados.

Tabla 4-4: Estadísticos descriptivos de las precipitaciones registradas en las diferentes estaciones

Código	Estación	Media	Mediana	Desvia.	Var.	Asime.	Curtosis	Máx.
E1	Guaslán	1,65	0,00	4,57	20,93	7,15	94,66	100,5
E2	Guamote	1,17	0,00	3,92	15,37	5,89	52,05	64,20
E3	Tixán	2,17	0,00	5,21	27,11	3,91	20,71	51,30
E4	Chunchi	1,22	0,00	3,34	11,21	5,68	57,10	72,10
E5	Pungales	1,43	0,00	3,85	14,87	4,38	25,07	41,70
E6	Urbina	2,72	1,40	3,99	15,98	3,63	20,43	43,10
E7	Pallatanga	2,85	0,00	6,64	44,14	4,78	37,56	120,40
E8	Huigra	1,18	0,00	3,78	14,33	6,05	47,38	50,20
E9	San Juan	1,83	0,00	4,42	19,57	4,09	22,11	54,8
E10	Cajabamba	1,35	0,00	3,13	9,78	3,97	24,27	41,00
E11	Cebadas	1,81	0,00	4,45	19,79	4,16	24,36	60,00
E12	Alao	3,07	0,40	4,76	22,72	2,39	8,78	52,30
E13	Compud	3,12	0,50	5,78	33,42	3,57	19,71	80,40
E14	Palmira	1,25	0,00	5,86	34,36	5,61	33,63	50,00
E15	Achupallas	2,21	0,00	4,91	24,10	3,16	12,74	48,90
E16	Ch. Pangor	2,95	0,00	6,29	39,66	4,27	29,33	97,70
E17	Alausi	1,50	0,00	3,49	12,18	5,14	67,41	94,00
E18	Cañi-limbe	3,57	0,00	6,53	42,67	2,84	10,88	67,2
E19	Guasuntos	1,24	0,00	3,12	9,73	4,536	29,505	42,50
E20	Chambo	1,86	0,20	3,56	12,67	2,67	7,58	20,00
E21	Licto	3,05	0,00	5,95	35,43	3,15	17,18	99,10
E22	Guano	1,21	0,00	3,83	14,74	4,48	24,86	46,70
E23	P. Velasco	3,95	2,00	5,05	25,52	2,18	17,77	104,00

Fuente: Grupo de Energías Alternativas y Ambiente (GEAA-INHAMI)

Realizado por: Juan C, Hidalgo LL, 2023

Según la (Tabla 4-4), la estación que presentó mayor precipitación promedio fue Pallatanga con una media de 2,85 mm y una desviación estándar de 3,12 mm, en cambio, Guasuntos con una media de 1,24 mm y una desviación estándar de 6,64 mm siendo la mas baja.

4.1.4. Análisis de las precipitaciones

En la (Figura 4-3) se visualiza los datos extremos de cada una de las estaciones, siendo Pallatanga (E7) la estación con el extremo mas alto.

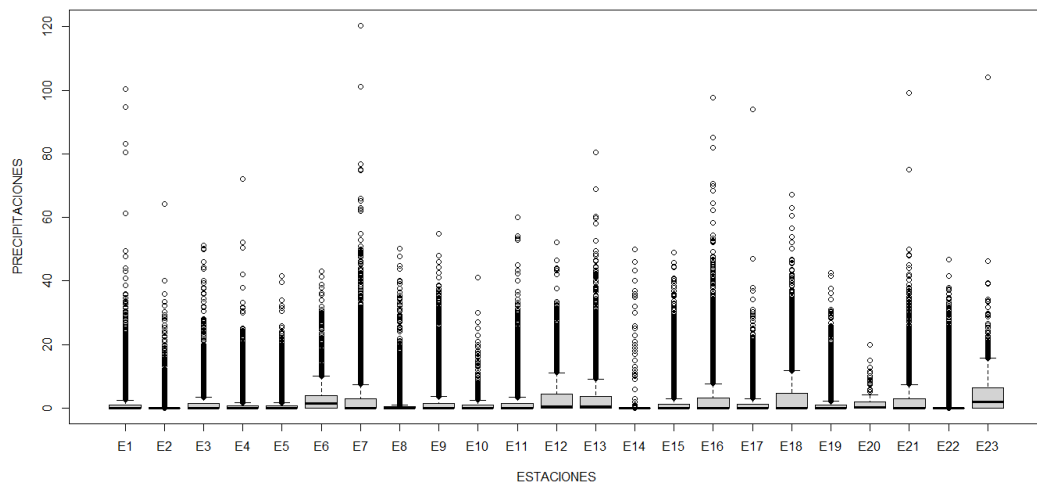


Ilustración 4-3: Boxplot

Fuente: Grupo de Energías Alternativas y Ambiente (GEAA-INHAMI)

4.1.5. Datos faltantes



Ilustración 4-4: Datos faltantes

Fuente: Grupo de Energías Alternativas y Ambiente (GEAA-INHAMI)

En la (Figura 4-4) se visualiza en color rojo los datos faltantes en cada estación, desde el 1 de enero de 1990 y el 31 de diciembre del 2018, identificando a las estaciones Chambo y Palmira con

un estimado del 85 % de ausencia, por esta razón estas 2 estaciones no fueron consideradas en el estudio.

4.2. Análisis de los umbrales

La literatura indica que el método más común para identificar umbrales es el de pico sobre umbral, conocidas por sus siglas en inglés POT (Peaks Over Threshold). El análisis del umbral se hace desde dos enfoques, con umbral fijo y variables. El propósito es identificar cual es la estrategia que caracterizan de mejor forma las precipitaciones extremas en la zona de estudio.

4.2.1. Identificación de los umbrales

Para la identificación de los umbrales existen varios métodos, sin embargo en esta investigación se utiliza la identificación mediante el cuantil 95 y el método del gráfico de la vida residual media.

4.3. Análisis mediante el umbral fijo

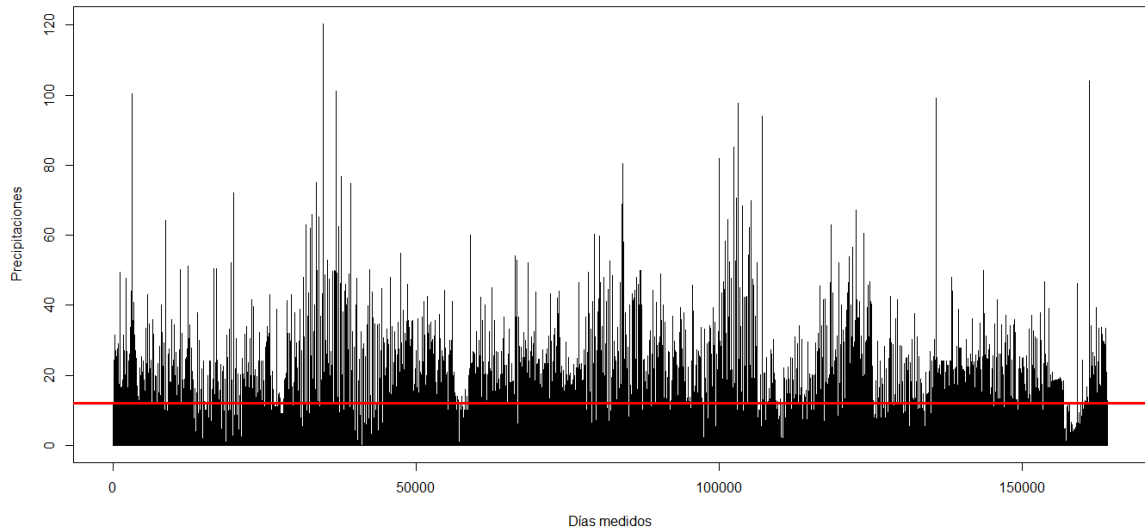


Ilustración 4-5: Identificación del umbral mediante el cuantil 95

Realizado por: Juan C, Hidalgo LL, 2023

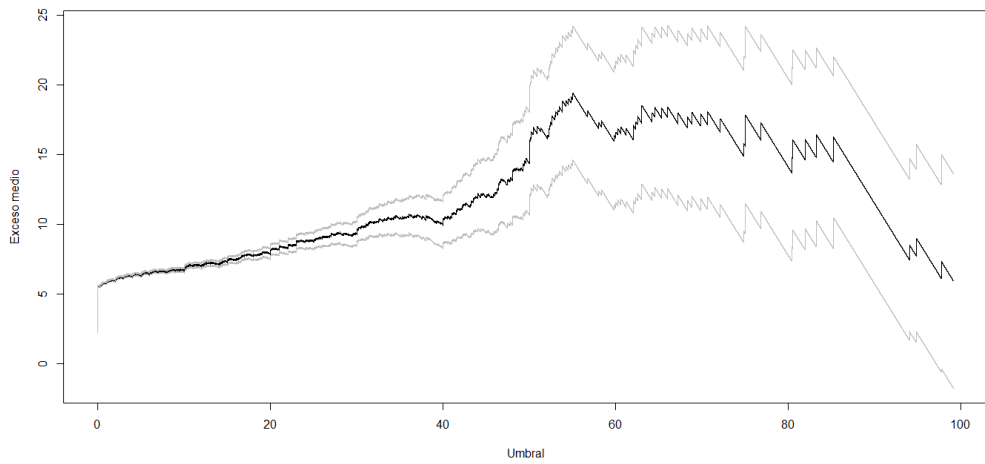


Ilustración 4-6: Vida residual media

Realizado por: Juan C, Hidalgo LL, 2023

Según la (Figura 4-5) del cuantil 95 el umbral óptimo es 12 mm, mientras que la (Figura 4-6) de Vida residual media la primera estabilidad y ruptura se da en [8,15].

Como el valor máximo esta dentro del intervalo [8,15] se corrobora que 12 mm es un umbral adecuado para el estudio.

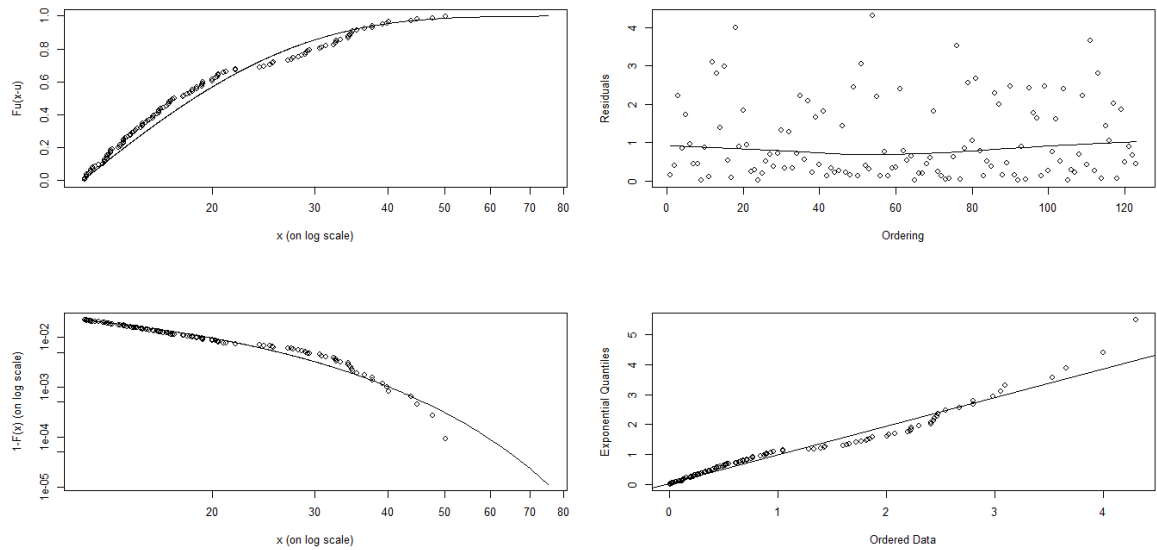


Ilustración 4-7: Distribución de Pareto generalizado de la estación Huigra con umbral fijo

Realizado por: Juan C, Hidalgo LL, 2023

Tabla 4-5: Parámetros de escala σ y forma ξ de la estación Huigra con umbral fijo (12 mm) con remuestreo de Bootstrap

Código	Umbral	Escala	Forma
E8	12,00	29,12	0,15
E8	12,00	17,46	-0,08
.	.	.	.
.	.	.	.
.	.	.	.
E8	12,00	7,73	0,13
PROMEDIO	12,00	17,47	-0,04

Realizado por: Juan C, Hidalgo LL, 2023

En la (Tabla 4-5) se muestra las estimaciones de los parámetros de escala σ y de forma ξ de los remuestros de Bootstrap (100 repeticiones) con umbral fijo en la estación Huigra.

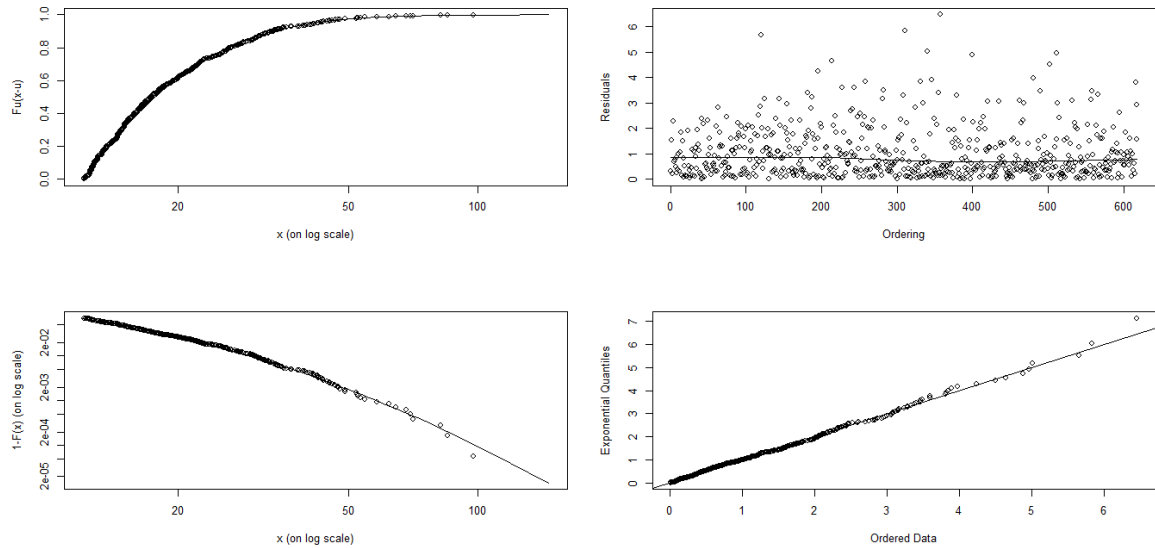


Ilustración 4-8: Distribución de Pareto generalizada de la estación Chimbo DJ Pangor con umbral fijo

Realizado por: Juan C, Hidalgo LL, 2023

Tabla 4-6: Parámetros de escala σ y forma ξ de la estación Chimbo DJ Pangor con umbral fijo (12 mm) con remuestreo de Bootstrap

Código	Umbral	Escala	Forma
E16	12,00	6,09	0,12
E16	12,00	4,57	0,16
.	.	.	.
.	.	.	.
.	.	.	.
E16	12,00	4,01	0,20
PROMEDIO	12,00	5,13	0,14

Realizado por: Juan C, Hidalgo LL, 2023

En la (Tabla 4-6) se muestra las estimaciones de los parámetros de escala σ y de forma ξ de los remuestreos de Bootstrap (100 repeticiones) con umbral fijo en la estación Chimbo DJ Pangor.

4.3.1. Precipitaciones extremas mediante umbral fijo

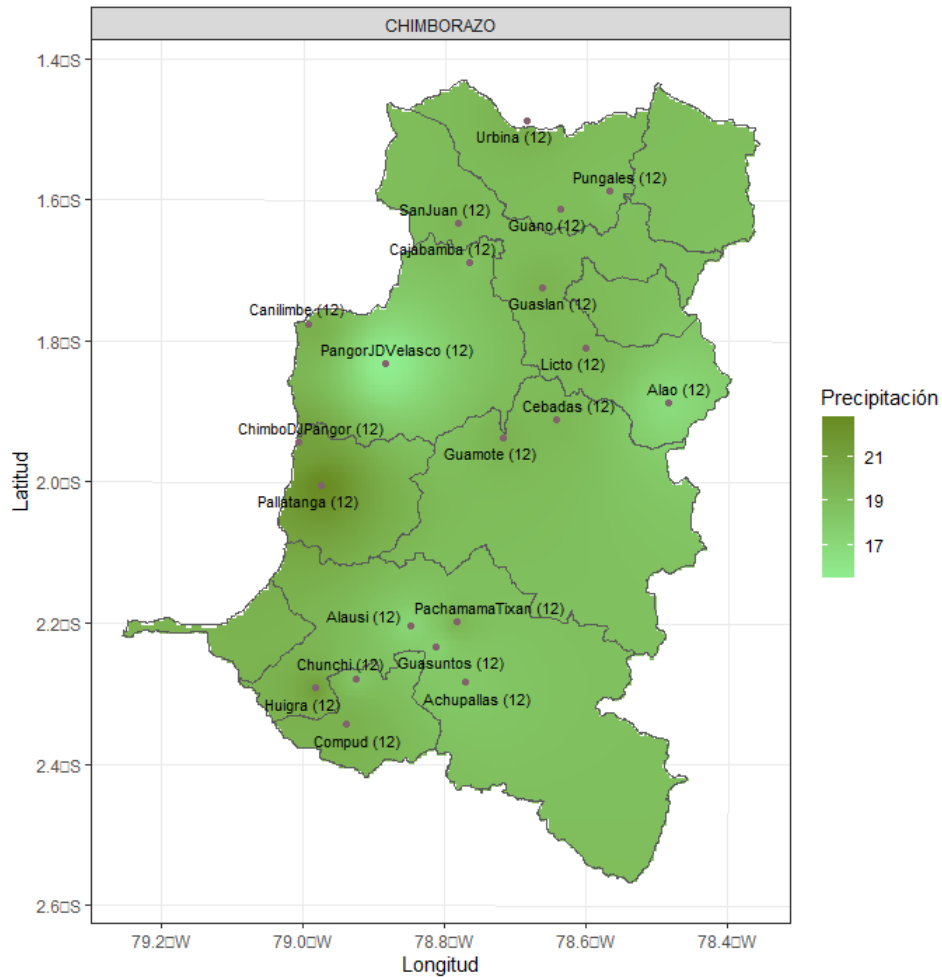


Ilustración 4-9: Precipitaciones extremas con umbral fijo mediante la interpolación IDW

Realizado por: Juan C. Hidalgo LL, 2023

En la (Figura 4-9) se identifica precipitaciones más extremas en la estación Pallatanga que se ubica en el cantón del mismo nombre, las 2 estaciones que también presentan extremos son Guaslán (Riobamba) y Huigra (Alausí). Mientras, la estación Pangor Juan de Velasco ubicado en el cantón Colta no se observa precipitaciones extremas, al igual que Guasuntos y Alao (Alausí, Riobamba) respectivamente presentan pocos valores que sobrepasen el umbral.

4.4. Mapa de riesgos extremos de precipitación en la provincia de Chimborazo con umbral fijo

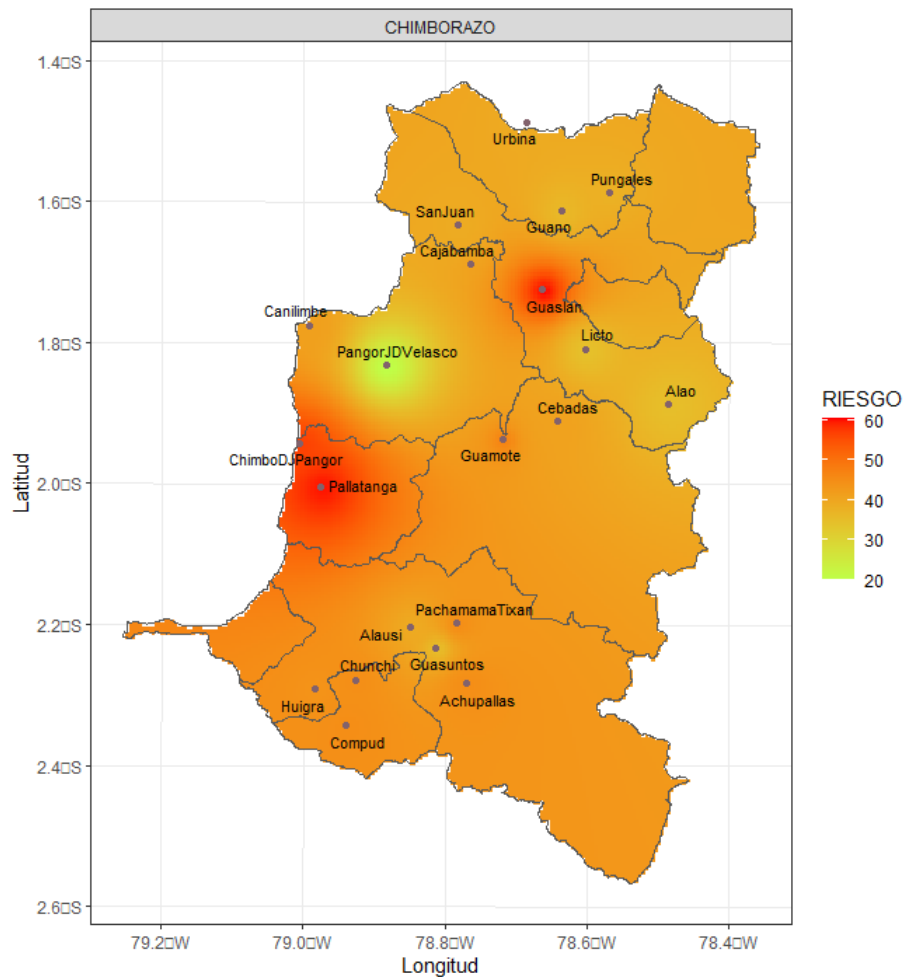


Ilustración 4-10: Riesgos extremos con umbral fijo

Realizado por: Juan C. Hidalgo LL, 2023

En la (Figura 4-10) se identifica los riesgos extremos en las estaciones de Pallatanga y Guaslán. Mientras la estación Pangor Juan de Velasco ubicado en el cantón Colta y Guasuntos en Alausí no se observan riesgos extremos.

4.5. Análisis de valores extremos mediante el umbral variable

4.5.1. Identificación de umbrales

Se analiza las precipitaciones extremas por cada estación meteorológicas con umbrales variables. A continuación, se muestra los resultados obtenidos para las estaciones de Huigra y Chimbo DJ Pangor, los resultados de las demás estaciones se puede ver en el (ANEXO C).

Estación Huigra

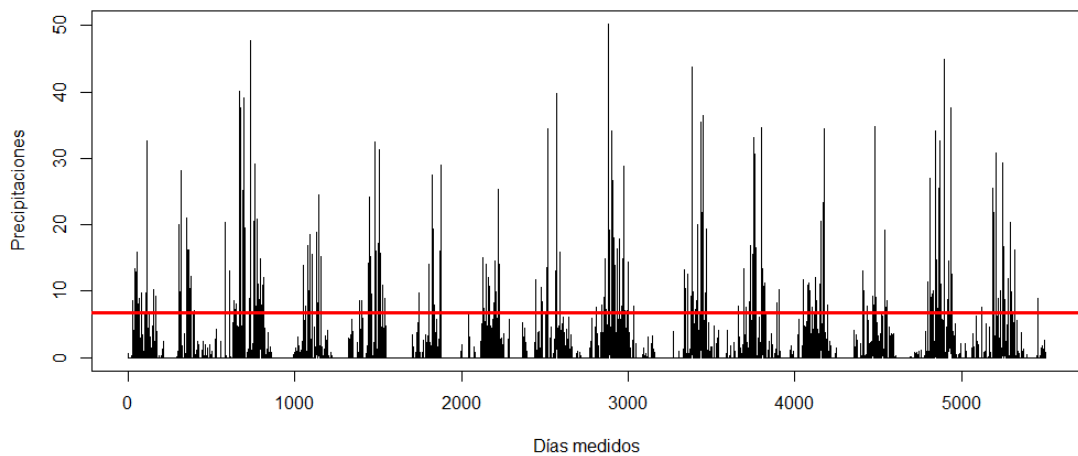


Ilustración 4-11: Identificación del umbral mediante el cuantil 95 de la estación Huigra

Realizado por: Juan C, Hidalgo LL, 2023

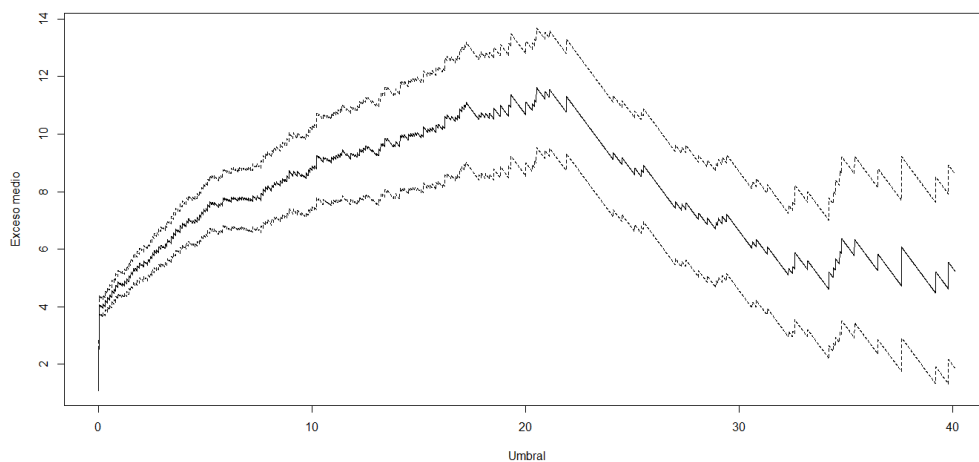


Ilustración 4-12: Vida residual media de la estación Huigra

Realizado por: Juan C, Hidalgo LL, 2023

Según la (Figura 4-11) del cuantil 95 el umbral óptimo es 6,68 mm, mientras que la (Figura 4-12) de Vida residual media la primera estabilidad y ruptura se da en [4,9].

Como el valor máximo esta dentro del intervalo [4,9] se corrobora que 6,68 mm es un umbral adecuado para el estudio.

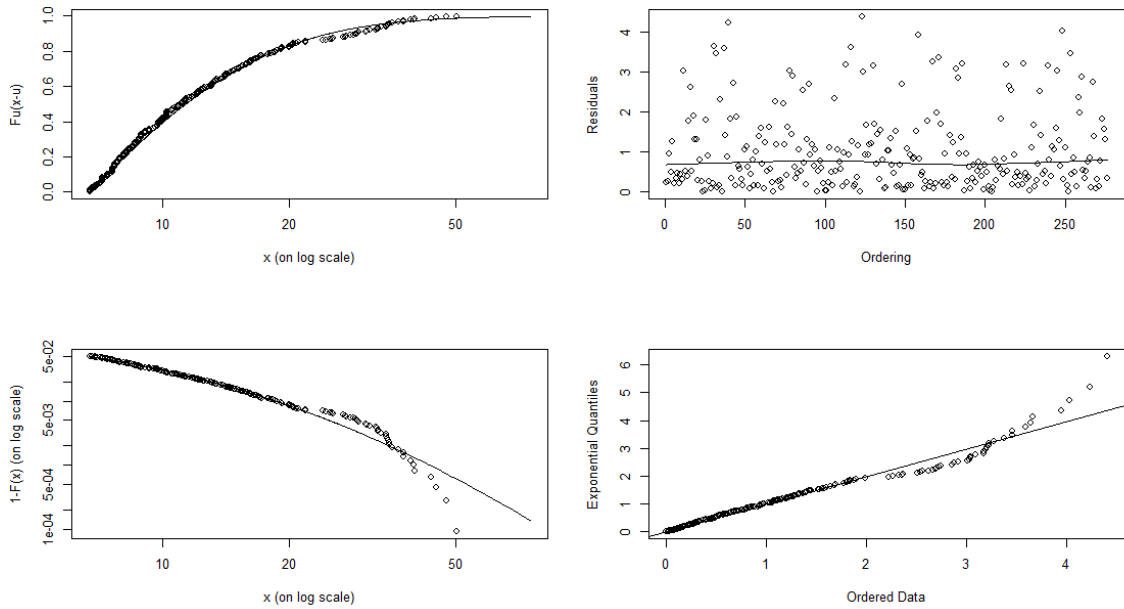


Ilustración 4-13: Distribución de Pareto generalizado de la estación Huigra con umbral variable
Realizado por: Juan C, Hidalgo LL, 2023

Tabla 4-7: Parámetros de escala σ y forma ξ
de la estación Huigra con umbral
variable (6,68 mm) con
remuestreo de Bootstrap

Código	Umbral	Escala	Forma
E8	6,68	4,34	0,14
E8	6,68	3,72	0,16
.	.	.	.
.	.	.	.
.	.	.	.
E8	6,68	2,24	0,15
PROMEDIO	6,68	3,66	0.17

Realizado por: Juan C, Hidalgo LL, 2023

En la (Tabla 4-7) se muestra las estimaciones de los parámetros de escala σ y de forma ξ de los remuestros de Bootstrap (100 repeticiones) con umbral variable en la estación Huigra.

Estación Chimbo DJ Pangor

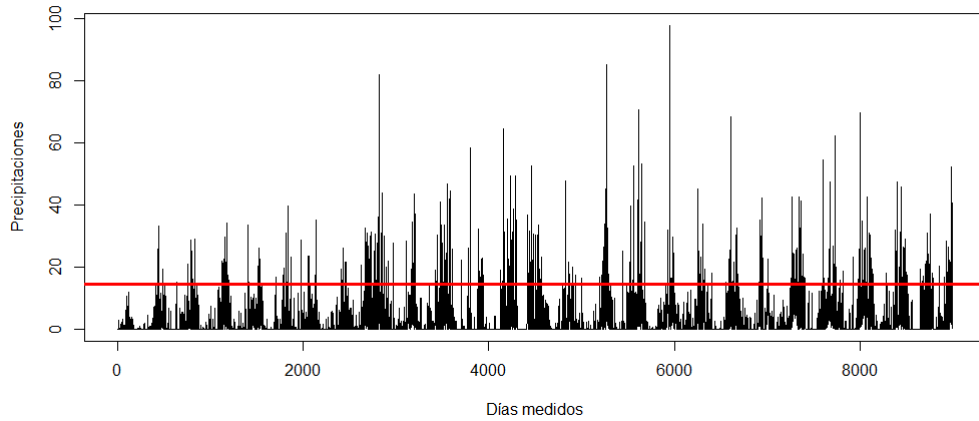


Ilustración 4-14: Identificación del umbral mediante el cuantil 95 de la estación Chimbo DJ Pangor

Realizado por: Juan C, Hidalgo LL, 2023

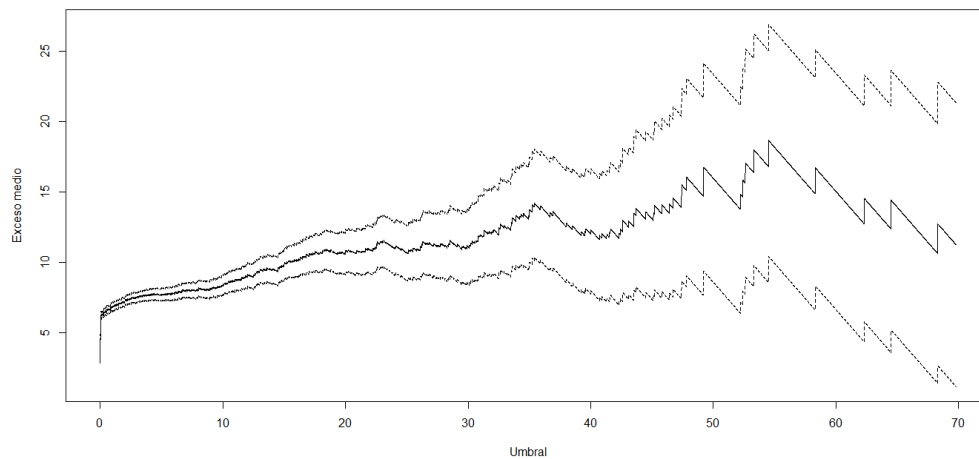


Ilustración 4-15: Vida residual media de la estación Chimbo DJ Pangor

Realizado por: Juan C, Hidalgo LL, 2023

Según la (Figura 4-14) del cuantil 95 el umbral óptimo es 14,50 mm, mientras que la (Figura 4-15) de Vida residual media la primera estabilidad y ruptura se da en [12,16].

Como el valor máximo está dentro del intervalo [12,16] se corrobora que 14,50 mm es un umbral adecuado para el estudio.

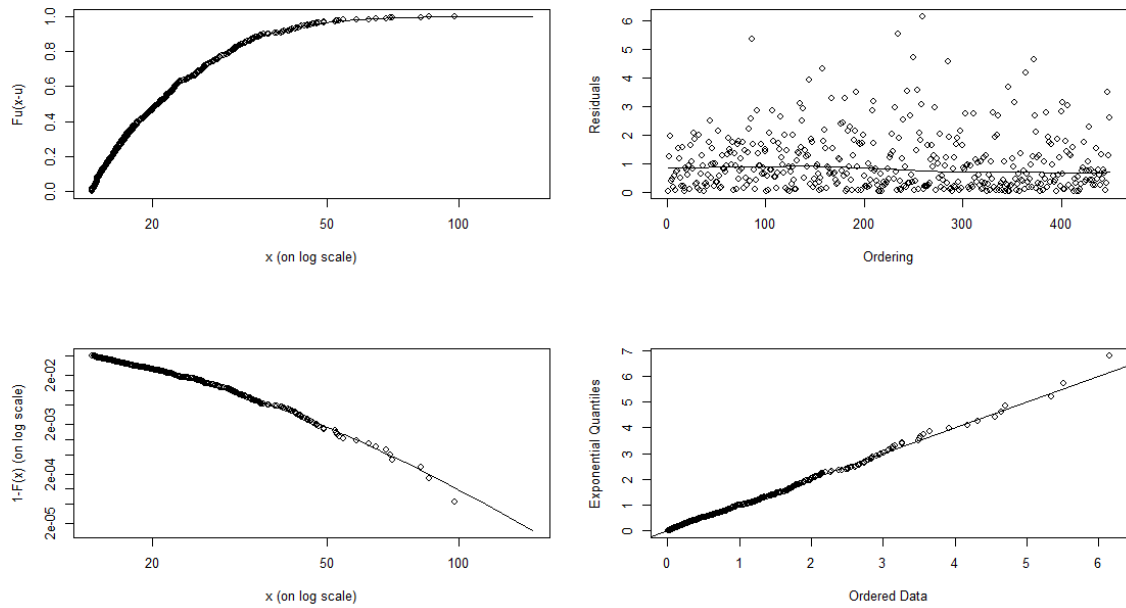


Ilustración 4-16: Distribución de Pareto generalizado de la estación Chimbo DJ Pangor con umbral variable

Realizado por: Juan C, Hidalgo, 2023

Tabla 4-8: Parámetros de escala σ y forma ξ de la estación Chimbo DJ Pangor con umbral variable (14,50 mm) con remuestreo de Bootstrap

Código	Umbral	Escala	Forma
E16	14,50	4,82	0,17
E16	14,50	5,41	0,13
.	.	.	.
.	.	.	.
.	.	.	.
E16	14,50	3,35	0,24
PROMEDIO	14,50	5,17	0,16

Realizado por: Juan C, Hidalgo LL, 2023

En la (Tabla 4-8) se muestra las estimaciones de los parámetros de escala σ y de forma ξ de los remuestros de Bootstrap (100 repeticiones) con umbral variable en la estación Chimbo DJ Pangor.

Tabla 4-9: Umbrales y parámetros de las estaciones con mejores valores

Código	Estación	Umbral	Parámetro de escala	Parámetro de forma
E1	Guaslán	9,50	2,84	0,24
E2	Guamote	7,80	5,58	0,07
E3	Pachamama-Tixán	12,30	5,79	0,07
E4	Chunchi	7,50	3,26	0,13
E5	Pungales	8,60	9,17	-0,07
E6	Urbina	9,13	3,24	0,17
E7	Pallatanga	14,40	8,39	0,07
E8	Huigra	6,68	3,55	0,19
E9	San Juan - Chimborazo	10,10	8,71	-0,04
E10	Cajabamba	8,00	3,29	0,09
E11	Cebadas	10,50	6,73	0,02
E12	Alao	12,70	3,99	0,06
E13	Compud	14,44	5,02	50,11
E15	Achupallas	12,60	6,67	-0,02
E16	Chimbo DJ Pangor	14,60	5,17	0,16
E17	Alausi	8,30	3,10	0,11
E18	Cañi-limbe	17,60	6,69	0,03
E19	Guasuntos	7,10	3,94	0,07
E21	Licto	17,00	4,46	0,07
E22	Guano	8,90	7,03	-0,02
E23	Pangor JD Velasco	14,50	1,59	0,15

Fuente: Grupo de energías alternativas y ambiente (GEAA-INMAHI)

Realizado por: Juan C, Hidalgo LL, 2023

En la (Tabla 4-9) se observa las estimaciones de los parámetros de escala σ y de forma ξ de las 23 estaciones del grupo de investigación GEAA-INHAMI con umbrales individuales.

4.5.2. Precipitaciones extremas mediante umbrales variables

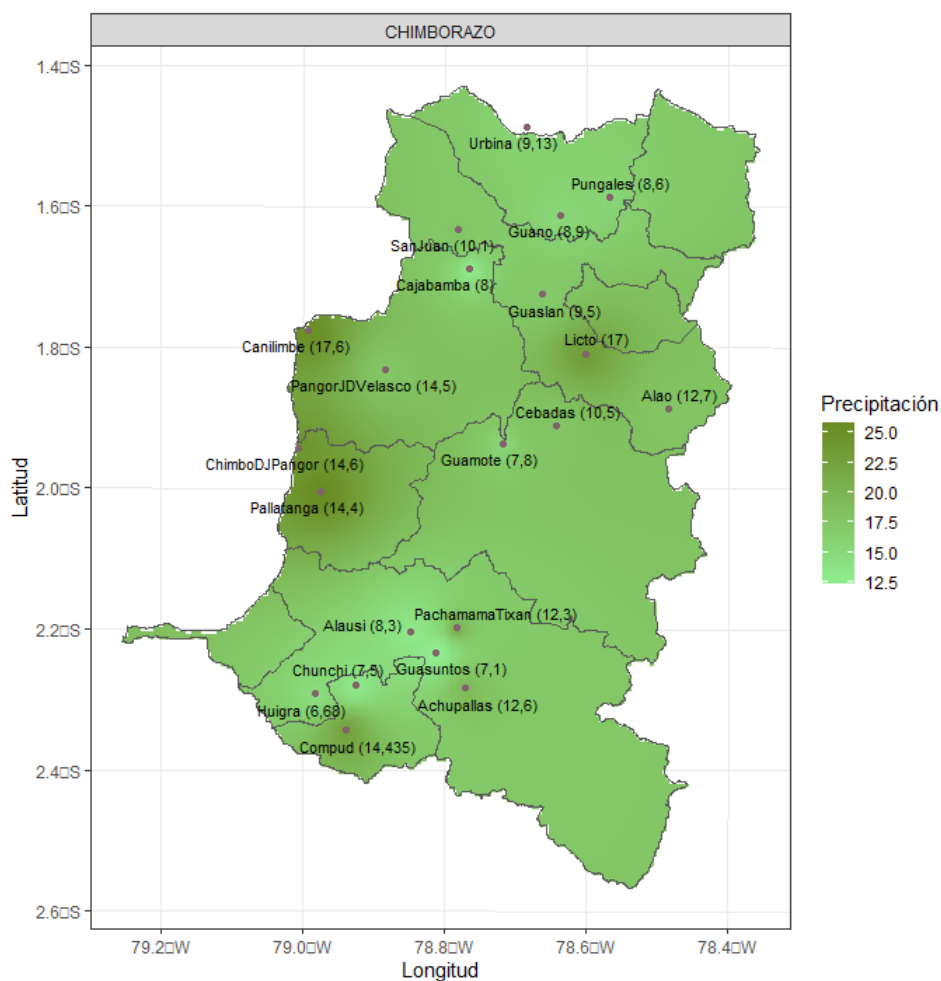


Ilustración 4-17: Precipitaciones extremas con umbrales variables

Realizado por: Juan C, Hidalgo LL, 2023

En la (Figura 4-17) se identifica precipitaciones más extremas en las estaciones Pallatanga, Licto y Cani-limbe que se ubica en los cantones Pallatanga, Riobamba y Colta respectivamente. Mientras, la estación Cajabamba ubicado en el cantón Colta y Guasuntos en el cantón Alausí presentan una disminución en los valores que sobrepasen el umbral.

4.6. Mapa de riesgos extremos de la provincia de Chimborazo con umbral variable

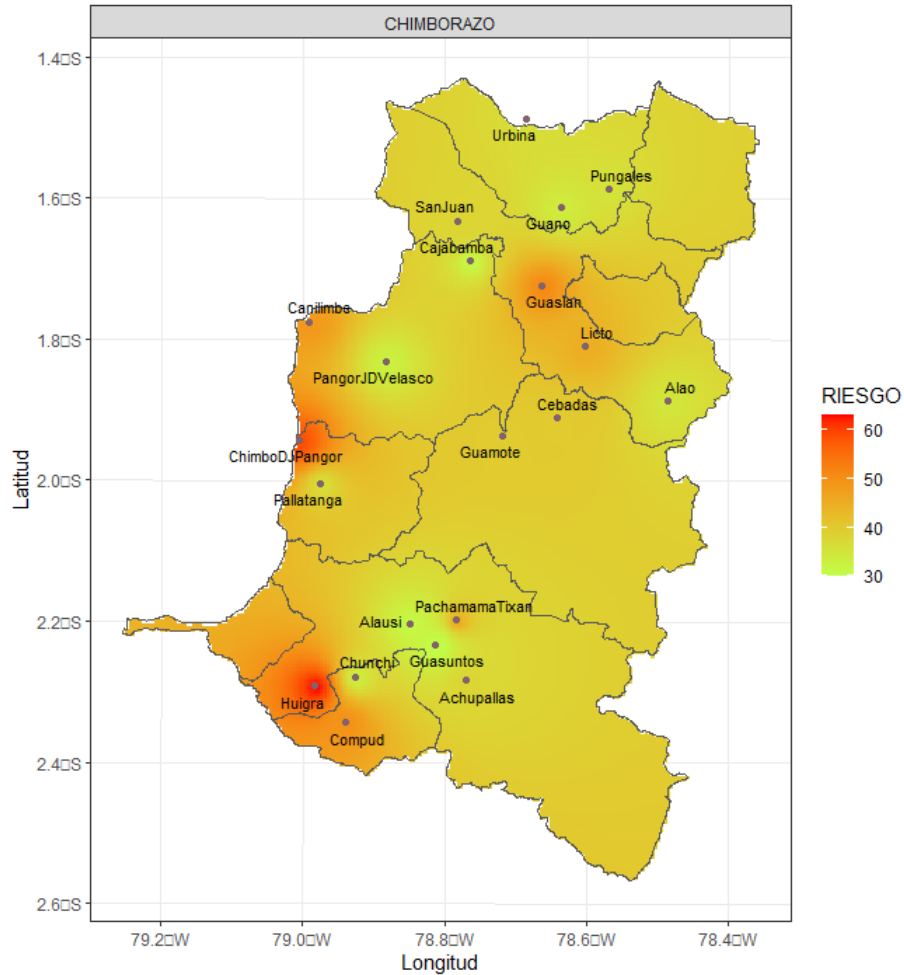


Ilustración 4-18: Riesgos extremos con umbral variable

Realizado por: Juan C. Hidalgo LL, 2023

En la (Figura 4-18) se identifica riesgos extremos en las estaciones de Huigra y Chimbo DJ Pangor, seguido por las estaciones Guaslán y Compud, mientras que, menor riesgos en las estaciones Pangor Juan de Velasco y Guasuntos.

Se identificó el comportamiento espacial de las precipitaciones extremas, considerando un análisis general mediante un umbral fijo para toda la zona de estudio y umbrales específicos por estación debido a sus características propias.

Con el umbral fijo se identifica mayor riesgo por precipitaciones en las estaciones Guaslán y Pallatanga debido a que no consideran características específicas de cada zona de estudio, mientras que, al usar umbrales específicos para cada estación (umbrales variables) se identificó mayor riesgo en las estaciones Chimbo DJ Pangor, Pallatanga, Licto y Cani-limbe, ya que esta metodología si

toma en cuenta la existencia de microclimas existentes en la zona donde se encuentran instaladas las estaciones.

4.7. Autocorrelación espacial

Para identificar la autocorrelación espacial se utiliza el índice de Moran basado en las ubicaciones de las estaciones y los valores de precipitaciones, es decir, evalúa si el patrón expresado está agrupado, disperso o es aleatorio. Se analiza la autocorrelación espacial desde el enfoque de precipitaciones máximas, mínimas y medianas.

Tabla 4-10: Índice de Moran y P-valor de los valores máximos, mínimos y medianas de excedentes con umbrales variables

Precipitaciones	I. Moran	P-Valor
Máximos	-0.02270042	0.6353563
Mediana	-0.07452825	0.6583534
Mínimos	-0.0888921	0.4908604

Realizado por: Juan C, Hidalgo LL, 2023

Como se puede observar en la (Tabla 4-10) las precipitaciones extremas máximas, medianas y mínimas están distribuidas de forma aleatoria entre estaciones meteorológicas, al 95% de confiabilidad.

4.8. Estrategias para el análisis recurrente de las precipitaciones extremas en la provincia de Chimborazo

La teoría de valores extremos es una técnica con gran interés en meteorología, su objetivo es analizar la distribución de valores extremos. Esta teoría ha sido ampliamente aplicada en diversos campos de la ciencia, como, por ejemplo, inundaciones, sequías, ráfagas de vientos, temperatura y precipitaciones.

Las precipitaciones se pueden analizar desde el punto de vista de valores extremos inferiores y superiores. En el primero caso se trataría desde la sequía hasta la llovizna, mientras que en el

segundo hace referencia a la lluvia que pasa de lo normal a la excesiva. La provincia de Chimborazo presenta microclimas (Bustamante, 2017, p.18-19), motivo por el cual el GEAA-INAMHI mantienen varias estaciones meteorológicas instaladas en puntos estratégicos de la provincia, lo que permite registrar las precipitaciones en diferentes zonas.

Para esta investigación se dispuso de los registros de precipitaciones de 23 estaciones meteorológicas instaladas en la provincia de Chimborazo, desde el 1 de enero de 1990 hasta el 31 de diciembre del 2018. En base a esta información se proporciona algunas estrategias para realizar análisis estadísticos recurrentes de precipitaciones extremas superiores.

1. Realizar un análisis exploratorio de los datos, que incluya la identificación geográfica de la zona en estudio, identificación de datos, estadísticos descriptivos de las precipitaciones, análisis gráfico (Boxplot) y datos faltantes.
2. Identificar el umbral (este puede ser único para todas las estaciones o específico para cada una, esto depende de las características geográficas de la zona de estudio) mediante el método de los cuantiles.
3. Usar el método de bootstrap para obtener las distribuciones muestrales.
4. Estimar los parámetros de escala σ y forma ξ de la distribución de Pareto generalizado y sus intervalos de confianza.
5. Realizar simulaciones de valores de precipitaciones excedentes en función de sus distribuciones muestrales.
6. Realizar mapas de precipitaciones extremas mediante interpolación IDW (método recomendado en función de las características de los datos analizados en esta investigación).
7. Hallar los valores en riesgo de las precipitaciones extremas y representarlos mediante mapas de interpolación.
8. En el caso de utilizar R emplea las librerías como por ejemplo *sf*, *raster*, *rgdal*, *rgeos*, *wesanderson*, *sp*, *mapview*, *gstat*, *tidyverse*, *readxl*.

CONCLUSIONES

- El análisis exploratorio se identificó: las localizaciones de las estaciones meteorológicas instaladas en la provincia de Chimborazo; los estadísticos descriptivos de cada estación, en donde los registros de precipitaciones más extremos se observaron en la estación E7 (Pallatanga); excesivos datos faltantes en las estaciones Palmira y Chambo, motivo el cual fueron separadas de la investigación.
- Se identificó los umbrales mediante las técnicas del cuantil y el gráfico de vida residual media. La primera proporciona un estimador puntual del umbral, mientras que el segundo un intervalo que contiene al estimador puntual. Sin embargo, para esta investigación se tomó la técnica del cuantil a un 95 %, valor que puede variar según las necesidades del investigador.
- Se utilizó el remuestreo de Bootstrap para encontrar las distribuciones muestrales de los excesos de precipitaciones en las diferentes estaciones meteorológicas, lo cual permitió hallar los parámetros de escala σ y de forma ξ de la distribución de Pareto generalizada en intervalos de confianza para cada estación meteorológica, tanto para el umbral fijo para toda la zona de estudio y específico para cada estación. Se realizó 100 remuestreos para cada proceso.
- Una vez realizada el análisis de los datos y valorado las técnicas utilizadas, se propone algunas estrategias para la realización de los análisis recurrentes de los registros de precipitaciones extremas en la provincia de Chimborazo, así como también la identificación de los valores en riesgo de cada estación.
- Mediante la interpolación espacial IDW con umbral fijo se identificó extremos y valores en riesgo mayores por precipitaciones en las estaciones Huigra y Pallatanga, mientras que con umbrales específicos en cada estación (umbrales variables) se identificó extremos mayores en las estaciones Pallatanga y Chimbo DJ Pangor, esto se debe a que el uso de un umbral fijo no permite capturar las características geográficas propias de cada estación meteorológica. Finalmente se corrobora mediante el índice de Moran, que las precipitaciones extremas en la provincia de Chimborazo no están autocorrelacionadas espacialmente.

RECOMENDACIONES

- Considerar nuevas estrategias para la recopilación de precipitaciones en cada estación, incrementando la frecuencia del mantenimiento en instrumentos de recolección de datos y reduciendo la pérdida de información.
- Considerar las características de la zona de estudio, especialmente en la Provincia de Chimborazo, debido a la presencia de microclimas, lo que conlleva a necesitar análisis localizados.
- considerar el coste computacional al realizar el remuestreo de Bootstrap, debido a que un número elevado de remuestras proporciona simulaciones ajustadas a los datos reales.
- Dar continuidad a esta investigación debido a que es un tema sumamente interesante y los resultados pueden ayudar de manera significativa en la toma de decisiones frente a fenómenos meteorológicos, además de ser aplicable en diversos campos de la ciencia.
- Incluir en las líneas de investigación en la Carrera de Estadística la teoría de valores extremos y sus aplicaciones debido a su amplia aplicabilidad.

BIBLIOGRAFÍA

AGUAYO , E. ; et al *Dependencia espacial de la delincuencia en Monterrey, México*. *Ecós de Economía* [en línea], 2014, (Monterrey-México) 18(38), pp. 63-92. [Consulta: 17 diciembre 2022]. ISSN 1657-4206. Disponible en: http://www.scielo.org.co/scielo.php?script=sci_abstract&pid=S1657-42062014000100003&lng=en&nrm=iso&tlng=es

ALVARADO, M. ; et al *Diseño y Desarrollo de Cartillas Informativas Culturales y Turísticas Enfocadas en la Cultura Puruhá*. *Adj. CDÁ's Multimedia: Promo Chimborazo y Anexo no. 18 de Fotografías*. [En línea] (Trabajo de titulación). (Ingeniería) Escuela Superior Politécnica de Chimborazo, Riobamba, ECUADOR, 2010. pp. 150-160. [Consulta: 17 enero 2022]. Disponible en: <http://dspace.esPOCH.edu.ec/handle/123456789/150>

ARAYA, C. *El modelo de clases latentes en presencia de tablas poco ocupadas: aplicación del método Bootstrap*. [En línea] (Trabajo de titulación). Universidad de Salamanca, Salamanca, ESPAÑA, 2011. pp. 33-44. [Consulta: 10 febrero 2022]. Disponible en: <https://kerwa.ucr.ac.cr/handle/10669/544>

ARISTIZÁBAL, E. ; et al *Análisis de umbrales empíricos de lluvia para el pronóstico de movimientos en masa en el valle de Aburrá, Colombia*. [en línea], 2011, (Aburrá-COLOMBIA), pp. 98-100. [Consulta: 17 noviembre 2022]. Disponible en: http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S1794-12372011000100009

CEBALLOS, A. ; et al *Análisis de la variabilidad espacio-temporal de las precipitaciones en el sector español de la cuenca del Duero (1961-2005)*. [en línea], 2013, pp. 13. [Consulta: 16 septiembre 2022]. Disponible en: <https://dialnet.unirioja.es/servlet/articulo?codigo=4157738>

BLANCO, M. ; et al *Metodología para investigar tendencias espacio-temporales en eventos meteorológicos extremos: caso Durango, México*. [en línea], 2014, (Durango-MEXICO), pp. 27-34. [Consulta: 05 septiembre 2022]. Disponible en: http://www.scielo.org.mx/scielo.php?script=sci_abstract&pid=S2007-24222014000600002&lng=es&nrm=iso&tlng=es

BORJA, C. *Análisis de la variabilidad espacio-temporal del índice de vegetación mejorado (EVI) y su relación con índices atmosféricos globales en el páramo ecuatoriano.* [en línea], 2020, (ECUADOR), pp. 44. [Consulta: 05 agosto 2022]. Disponible en: <http://bibdigital.epn.edu.ec/handle/15000/20950>

BUSTAMANTE, D. *Escenario de cambio climático a nivel de subcuencas hidrográficas para el año 2050 de la provincia de Chimborazo-Ecuador.* [en línea], 2017, (Chimborazo-ECUADOR), pp. 17-18. [Consulta: 05 agosto 2022]. Disponible en: http://scielo.senescyt.gob.ec/scielo.php?script=sci_abstract&pid=S1390-85962017000100015&lng=es&nrm=iso&tlng=es

CANO, D. ; et al *Un modelo OO para manipulación de datos espaciales.* [en línea], Universidad Nacional de La Plata, ARGENTINA, 2007, pp. 7-10. [Consulta: 26 agosto 2022]. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/3965>

CEEA-ESPOCH *Centro de Energías Alternativas y Ambiente.* [en línea], 2022, (Chimborazo-ECUADOR), pp. 7-8. [Consulta: 05 agosto 2022]. Disponible en: <https://www.espoch.edu.ec/index.php/component/k2/item/1595-ceaa-de-la-epoch-desarrolla-su-plataforma-digital.html>

AVILÉS, D. ; et al *Aplicación de la teoría de valores extremos en el análisis de eventos hidrometeorológicos.* [en línea], Universidad Nacional Autónoma de México, MÉXICO, 2014, pp. 7-10. [Consulta: 29 agosto 2022]. Disponible en: <https://www.researchgate.net/profile/Rodolfo-Silva-9/publication/324546038Aplicaciondelateoriadevaloresextremos>

IBÁÑES, A. *Análisis estadístico de valores extremos y aplicaciones.* [en línea], Universidad de Granada, ESPAÑA, 2011, pp. 1-10. [Consulta: 29 agosto 2022]. Disponible en: <https://docplayer.es/22442547>

INAMHI *Instituto Nacional de Meteorología e Hidrología.* [en línea], 2021, (ECUADOR), pp. 7-8. [Consulta: 01 agosto 2022]. Disponible en: http://extension.fcaglp.unlp.edu.ar/content/obs/descargas/Introduccion_a_la_Meteorologia_I.pdf

INPC *Instituto Nacional de Patrimonio Cultural.* [en línea], 2010, (ECUADOR), pp. 7-8. [Consulta: 01 agosto 2022]. Disponible en: <https://downloads.arqueo-ecuatoriana.ec/ayhpwxgv/noticias/publicaciones/INPC-X-GuiaChimborazo.pdf>

MEDIALDEA, A. *Análisis de valores extremos.* [en línea], Universidad de Granada, ESPAÑA, 2016, pp. 1-10. [Consulta: 29 julio 2022]. Disponible en: <https://docplayer.es/22442547-Realizado-por-alejandro-ibanez-rosales-trabajo-de-investigacion>

MIRANDA, M. ; et al *Importancia del análisis estadístico exploratorio en el proceso de interpolación espacial: caso de estudio Reserva Forestal Valdivia.* [en línea], 2003, pp. 58. [Consulta: 26 junio 2022]. Disponible en: http://www.scielo.cl/scielo.php?script=sci_abstract&pid=S0717-92002003000200004&lng=es&nrm=iso&tlng=pt

MOROTE, Á. ; et al *El estudio del cambio climático en la Educación Primaria: una exploración a partir de los manuales escolares de Ciencias Sociales de la Comunidad Valenciana.* [en línea], 2020, pp. 158. [Consulta: 26 junio 2022]. Disponible en: <https://roderic.uv.es/handle/10550/78315>

QUINTANA, M. *Estimación Bootstrap con datos secuenciales.* [en línea], 2003, pp. 45. [Consulta: 26 junio 2022]. Disponible en: <https://accedacris.ulpgc.es/jspui/handle/10553/19572>

SAROCHAR, R. *Precipitación convectiva y estratiforme en la estación Buenos Aires - observatorio central: una aproximación más precisa a su discriminación.* [en línea], 2009, pp. 56. [Consulta: 26 julio 2022]. Disponible en: http://www.scielo.org.ar/scielo.php?script=sci_abstract&pid=S1850-468X2016000100004&lng=es&nrm=iso&tlng=pt

SERRA, I. ; et al *Modelos estadísticos para valores extremos y aplicaciones.* [en línea], Universidad Autónoma de Barcelona, ESPAÑA, 2014, pp. 106. [Consulta: 26 mayo 2022]. Disponible en: <https://ddd.uab.cat/record/116414>

MORENO, R. ; et al *Técnicas econométricas para el tratamiento de datos espaciales : la econometría espacial.* [en línea], Universidad Autónoma de Barcelona, ESPAÑA, 2000, pp. 60. [Consulta: 26 mayo 2022]. Disponible en: https://books.google.es/books?id=Udh_wcm75GwC

PACHECO, I. *Visualización de datos espaciales en R: elecciones gubernamentales 2016 en Zacatecas.* [en línea], 2016, pp. 160. [Consulta: 26 abril 2022]. Disponible en: https://www.researchgate.net/profile/Ivan-Pacheco-Soto/publication/323666587_Visualizacion_de_datos_espaciales_en_R

VARGAS, L. *Teoría de valores extremos para el análisis de la precipitación de la estación meteorológica ESPOCH (1976-2019).* [en línea], 2021, (Chimborazo-ECUADOR), pp. 60. [Consulta: 26 abril 2022]. Disponible en: <http://dspace.esPOCH.edu.ec/handle/123456789/14824>

RAYNAL, J. *Sobre el uso del dominio de atracción para la identificación de distribuciones de valores extremos para máximos.* [en línea], 1997, pp. 57-62. [Consulta: 26 abril 2022]. Disponible en: <http://revistatyca.org.mx/ojs/index.php/tyca/article/view/783>

VITTURINI, M. *Modelos de Datos para Datos Espaciales.* [en línea], 2003, pp. 11. [Consulta: 26 abril 2022]. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/21569>

CHÁVEZ, J. *Uso de la distribución generalizada de Pareto Multivariada Para Modelar Riesgo Operativo.* [en línea], 2009, pp. 59-88. [Consulta: 26 agosto 2022]. Disponible en: <https://repositorio.tec.mx/bitstream/handle/11285/629367/33068001081687.pdf?sequence=1>



ANEXOS

ANEXO A: SCRIPT DE R UTILIZADO EN LA INVESTIGACIÓN PARA OBTENER EL ANÁLISIS EXPLORATORIO

```
chi <- shapefile(" Cantones . shp ")
plot(chi)
showMethods(" mapView ")
mapview(chi)
pts <- lts1 [[1]]
coordinates(pts) <- ~ X + Y
crs(pts) <- crs(chi)
plot(pts)
ecuador_map <- st_read(" Cantones . shp ")
unique(ecuador_map$DPA_DESPRO)
chimborazo_map <- cbind(chimborazo_map ,
POBLACION <- c(306.231,64.059,50.431,17.089,20.587,46.659,
56.254,18.327,13.655,18.129))
mapa <- ggplot(chimborazo_map) +
geom_sf(aes(fill=POBLACION), chimborazo_map ,
col='black', size=0.5) +
guides(fill=guide_colorbar(title = "Densidad
Poblacional"))+
theme(plot.title = element_text(family=" Arial "),
legend.position = "right",
panel.background = element_rect(fill = "white"))+
geom_sf(data = st_as_sf(chi), fill= NA)+
geom_sf(data = st_as_sf(pts), color = "#82626C")+
xlab(" Longitud ") + ylab(" Latitud ") +
legend.position = "right",
panel.background = element_rect(fill = "white"))

mapa
```

ANEXO B: DATOS FALTANTES

```
aggr(est , prop=FALSE,  
      numbers=TRUE, border =NA,  
      combine=TRUE, cex.lab = 1, cex.axis = 1)
```

ANEXO C: IDENTIFICACIÓN DE UMBRALES POR EL CUANTIL Y EL GRÁFICO DE VIDA RESIDUAL MEDIA

Método del cuantil

```
q <- quantile(E1, probs = 0.95)  
par(mfrow = c(1, 1))  
quantile(E1, c(0.05, 0.95))  
plot(E1, type="l")
```

Gráfico de vida residual media

```
library(evd)  
mrlplot(E1)
```

ANEXO D: DISTRIBUCIÓN DE PARETO GENERALIZADA

```
fit <- gpd(pareto , threshold=9.5, nextremes=NA)  
fit
```

ANEXO E: REMUESTREO BOOTSTRAP

```
alpha <- 0.05  
nSim <- 100  
set.seed(1)  
sbootE1 <- replicate(nSim, {  
  indice <- sample(1:n1, n1, replace = T)  
  fit <- gpd(E1[indice], threshold=9.5, nextremes=NA)  
  tp <- tailplot(fit)  
  unlist(tp[9])[1]  
})  
sbootE1
```

ANEXO F: DISTRIBUCIÓN MUESTRAL

```
g <- function( f,e) {  
  mues <- 1-((1+(f*E1)/e)^-(1/f))  
  return(mues)  
}  
t1 <- g(f,e)
```

ANEXO G: INTERPOLACIÓN IDW EN VALORES EXTREMOS Y VALORES DE RIESGO

```
chi <- shapefile("Cantones.shp")  
plot(chi)  
showMethods("mapView")  
mapview(chi)  
coordinates(grd1) <- ~ X + Y  
gridded(grd1) <- TRUE  
crs(grd1) <- crs(chi)  
calc_idw <- function(x){  
  print(x)  
  pts <- lts1 [[x]]  
  a <- unique(pts)  
  idw.p <- gstat::idw(PRECIPITACION ~ 1, pts , grd1)  
  idw.p <- raster(idw.p)  
  idw.p <- raster::mask(idw.p , chi)  
  dir.create("UMBRAL")  
  ord <- c(1: length(lts1))  
  rs11 <- map(.x = ord, .f = calc_idw)  
  plot(lyr)  
  tbl <- rasterToPoints(lyr , xy = TRUE) %>% as_tibble() %>%  
  gather(var , value , -x, -y)  
  head(tbl ,2)  
  valores extremos <- ggplot() +  
  geom_raster(data = tbl , mapping = aes(x=x, y=y, fill = value)) +  
  facet_wrap(~var)+ scale_color_viridis_c()  
  
valores extremos
```




epoch

Dirección de Bibliotecas y
Recursos del Aprendizaje

UNIDAD DE PROCESOS TÉCNICOS Y ANÁLISIS BIBLIOGRÁFICO Y
DOCUMENTAL

REVISIÓN DE NORMAS TÉCNICAS, RESUMEN Y BIBLIOGRAFÍA

Fecha de entrega: 19 / 06 / 2023

INFORMACIÓN DEL AUTOR/A (S)
Juan Cristian Hidalgo Llangarí
INFORMACIÓN INSTITUCIONAL
Facultad: Ciencias
Carrera: Estadística
Título a optar: Ingeniero Estadística
f. Analista de Biblioteca responsable: Ing. Rafael Inty Salto Hidalgo

0848-DBRA-UPT-2023