



FACULTAD DE  
CIENCIAS ECONÓMICAS  
Y DE ADMINISTRACIÓN



UNIVERSIDAD  
DE LA REPÚBLICA  
URUGUAY

UNIVERSIDAD DE LA REPÚBLICA

FACULTAD DE CIENCIAS ECONÓMICAS Y DE  
ADMINISTRACIÓN

TRABAJO FINAL DE GRADO PARA OBTENER EL  
TÍTULO DE LICENCIADO EN ESTADÍSTICA:

Análisis estadístico de precipitaciones extremas en Uruguay mediante  
procesos máx-estables.

Juan Ignacio Baccino Costa

Francisco Bonora Melognio

Tutor: PhD. Leonardo Moreno

Montevideo

URUGUAY

2022

---

UNIVERSIDAD DE LA REPÚBLICA

El tribunal docente integrado por los abajo firmantes aprueba el trabajo final de grado:

**Análisis estadístico de precipitaciones extremas en Uruguay  
mediante procesos máx-estables.**

**Juan Ignacio Baccino Costa y Francisco Bonora Melognio**

Tutor:

Leonardo Moreno Romero

Licenciatura en Estadística

**Puntaje** .....

**Tribunal**

Profesor Dr. Marco Scavino .....(firma).

Profesor Fernando Massa .....(firma).

Profesor Leonardo Moreno .....(firma).

**Fecha** .....

Juan Ignacio Baccino Costa y Francisco Bonora Melognio

---

## Agradecimientos

En primera instancia queremos agradecer a nuestro tutor Leonardo Moreno, por el impresionante apoyo que nos brindó y la ejemplar presencia a lo largo del trabajo, tanto para ayudarnos como para darnos consejos.

A nuestras familias y amigos que en el transcurso de la carrera siempre estuvieron brindando su apoyo.

Al Instituto de Estadística, IESTA, y profesores que tuvimos a lo largo de la carrera por el apoyo en el camino de nuestra formación.

A nuestros compañeros de estudios que la mayoría se convirtieron en amigos.

Al Instituto Uruguayo de Meteorología, INUMET, por brindarnos los datos con los cuales pudimos llevar a cabo nuestro trabajo.



# Índice general

Índice general	v
Índice de figuras	vii
Índice de cuadros	xi
<b>1. Precipitaciones extremas en Uruguay y la base de datos</b>	<b>3</b>
1.1. Precipitaciones extremas en Uruguay . . . . .	3
1.2. La base de datos . . . . .	8
1.2.1. Filtrado inicial de la base . . . . .	9
1.3. Análisis Descriptivo de la base de datos . . . . .	15
<b>2. Metodología</b>	<b>21</b>
2.1. Teoría de Valores Extremos univariados . . . . .	21
2.1.1. Introducción . . . . .	21
2.1.2. Introducción al problema . . . . .	22
2.1.3. Distribuciones para la TVE . . . . .	24
2.1.4. Estimación de parámetros para la DGVE . . . . .	26
2.2. Procesos máx-estables . . . . .	30
2.2.1. Definiciones y teoremas principales . . . . .	30
2.2.2. Dependencia en procesos máx-estables . . . . .	33
2.2.3. Modelos aplicados . . . . .	44
2.2.4. Selección del Modelo . . . . .	50
2.2.5. Niveles de Retorno . . . . .	51

## ÍNDICE GENERAL

---

<b>3. Resultados Globales</b>	<b>53</b>
3.1. Estimación de parámetros para las DGVE . . . . .	53
3.1.1. Máxima Verosimilitud y Estimación Conjunta . . . . .	54
3.2. Estimación espacial . . . . .	63
3.2.1. Transformación a Fréchet Unitaria . . . . .	63
3.2.2. Estimación espacial mediante procesos máx-estables . . . . .	63
3.2.3. Selección de Modelos . . . . .	66
3.3. Simulaciones . . . . .	67
3.4. Niveles de Retornos estimados . . . . .	80
<b>4. Conclusiones</b>	<b>83</b>
<b>5. Anexo</b>	<b>87</b>
5.1. Anexo A - Código en el software R . . . . .	87
5.2. Anexo B - Metodología y Resultados . . . . .	93
5.2.1. Métodos de los Momentos Pesados . . . . .	93
5.2.2. Selección de Modelos . . . . .	94
5.2.3. Resultados . . . . .	97
5.2.4. Frechet . . . . .	111
5.2.5. Gráfica, Riesgos de la actualidad . . . . .	112
Referencias . . . . .	113

# Índice de figuras

1.1. Inundaciones en Santa Lucía y Durazno en el año 2019. Fuente: Santiago Porcal - Focouy y Sinae. . . . .	5
1.2. Mapa de precipitaciones acumuladas en Uruguay para el año 2016 y 2017. Fuente: Instituto Nacional de Investigación Agropecuario (INIA). . . . .	6
1.3. Mapa político e hidrográfico de Uruguay. Fuente: Álvaro Merino (2021) y <a href="https://es.liveworksheets.com/qf2316268ov">https://es.liveworksheets.com/qf2316268ov</a> . . . . .	7
1.4. Ubicación de las 19 estaciones meteorológicas y 53 estaciones pluviométricas en Uruguay. Fuente: elaboración propia en R. . . . .	9
1.5. Período activo de las estaciones climatológicas y pluviométricas. Fuente: elaboración propia en R. . . . .	11
1.6. Ubicación de las 44 estaciones meteorológicas y pluviométricas en Uruguay en el período 1982-2020 y las estaciones en colores que presentan una cantidad de datos faltantes igual o mayor a 20%. Fuente: elaboración propia en R. . . . .	13
1.7. Precipitaciones máximas anuales registradas en Uruguay en las 41 estaciones en el período 1982-2020. Fuente: elaboración propia en R. . . . .	16
1.8. Correlación de Kendall para los máximos de las 41 estaciones meteorológicas y pluviométricas según la distancia euclidiana dos a dos. Fuente: elaboración propia en R. . . . .	19
2.1. Ejemplo de un variograma experimental, que muestra los parámetros nugget, umbral y rango. Fuente: Porras (2014) . . . . .	48

## ÍNDICE DE FIGURAS

---

3.1. Estimación máxima verosímil respecto a estimación conjunta de los parámetros de posición, escala y forma del total de estaciones. Fuente: elaboración propia en R. . . . .	57
3.2. Ubicación de la estación meteorológica Piedras Coloradas y las restantes 40 estaciones pluviométricas o meteorológicas. Fuente: elaboración propia en R. . . . .	58
3.3. Estimación máxima verosímil respecto a estimación conjunta de los parámetros de posición, escala y forma en el total de estaciones sin la estación Piedras Coloradas. Fuente: elaboración propia en R. . . .	59
3.4. Intervalo de confianza para la estimación puntal máximo verosímil del parámetro de forma. Fuente: elaboración propia en R. . . . .	60
3.5. Ubicación de las estaciones meteorológicas de 25 de Agosto y San Antonio y las restantes 38 estaciones pluviométricas y meteorológicas. Fuente: elaboración propia en R. . . . .	61
3.6. Estimación del coeficiente extremal del proceso t-extremal no paramétrica (violeta), paramétrica con Bessel (azul) y paramétrica con Power Exponencial (violeta), luego de transformar los datos a una Fréchet Unitaria del período 1982-2020 de las 38 estaciones, considerando una <i>DVE</i> Fréchet. Fuente: elaboración propia en R. . . . .	65
3.7. Estimación del Madograma y F-Madograma para las 38 estaciones meteorológicas ó pluviométricas. Fuente: elaboración propia en R. . .	66
3.8. Región del territorio donde se realizaron las simulaciones. Fuente: elaboración propia en R. . . . .	68
3.9. Comparación del promedio de precipitaciones máximas simulados mediante el modelo <i>t-extremal</i> con correlaciones <i>Bessel</i> , <i>Whittle-Matérn</i> y <i>Power Exponencial</i> contra el promedio de los máximos de las precipitaciones observadas. Fuente: elaboración propia en R. . . . .	70

3.10. Comparación entre valores simulados del promedio de las precipitaciones máximas por trimestres considerando un modelo <i>t-extremal</i> con correlación <i>Bessel</i> y el promedio de las precipitaciones máximas observadas por trimestres. Fuente: elaboración propia en R. . . . .	71
3.11. Comparación entre valores simulados del promedio de las precipitaciones máximas por trimestres considerando un modelo <i>t-extremal</i> con correlación <i>Bessel</i> y el promedio de las precipitaciones máximas observadas por trimestres estación por estación. Fuente: elaboración propia en R. . . . .	73
3.12. Comparación entre valores simulados del promedio de las precipitaciones máximas por trimestres considerando un modelo <i>t-extremal</i> con correlación <i>Bessel</i> y el promedio de las precipitaciones máximas observadas por trimestres estación por estación. Fuente: elaboración propia en R. . . . .	74
3.13. Mapas de calor de las precipitaciones máximas simuladas con un modelo <i>t-extremal</i> con correlación <i>Bessel</i> en Uruguay. Fuente: elaboración propia en R. . . . .	75
3.14. Comparación del promedio de las precipitaciones máximas simuladas entre el norte y sur de Uruguay. Fuente: elaboración propia en R. . .	77
3.15. Grilla de coordenadas con la ubicación de las estaciones pluviométricas de <i>Aeropuerto Carrasco</i> y <i>Aeropuerto Melilla</i> donde se estimaron los diferentes <i>niveles de retornos</i> . Fuente: elaboración propia en R. . .	80
3.16. Niveles de retornos estimados para las precipitaciones extremas para los períodos de tiempo de 5, 10, 15, 20 y 50 años. Fuente: elaboración propia en R. . . . .	81
5.1. Coeficiente extremal estimada considerando una DVE del tipo Gumbel. Fuente: elaboración propia en R. . . . .	106
5.2. F-madograma estimada considerando una DVE del tipo Gumbel. Fuente: elaboración propia en R. . . . .	106

## ÍNDICE DE FIGURAS

---

- 5.3. Comparación entre los valores promedios de los máximos observados en los 147 trimestres en el período 1982-2020 de las 38 estaciones contra los valores simulados, considerando un modelo t-extremal con correlaciones: Bessel, Whittle-Mattem y Power Exponencial . . . . . 109
- 5.4. Capacidad destructiva y ocurrencia de eventos mundiales. Fuente: Informe de riesgos mundiales 2019. Fuente: World Economic Forum. . 112

# Índice de cuadros

1.1. Estaciones junto con los trimestres y el año de las estaciones que presentan un porcentaje igual o mayor a un 20 % de datos faltantes. . . . .	14
1.2. Máximos de precipitaciones en mm de los últimos 8 trimestres en 5 estaciones . . . . .	17
3.1. Valor del TIC para los seis mejores modelos utilizando dicho criterio de selección. . . . .	56
3.2. Valor del TIC para los cinco modelos con menor TIC en estimación espacial mediante procesos máx-estables para el período 1982-2020 en 38 estaciones meteorológicas y pluviométricas considerando una <i>DVE Fréchet</i> . . . . .	67
3.3. Estimación de los parámetros para los mejores tres modelos de los procesos máx-estables. . . . .	67
3.4. Probabilidades de superar un umbral de precipitaciones diarias máximas para un período de 1 año, 5 años y 10 años con el modelo <i>t-extremal</i> con correlaciones <i>Bessel</i> y <i>Power Exponencial</i> . . . . .	79
5.1. Estimación de los parámetros de la DVEG por el método de máxima verosimilitud y por el método de los momentos pesados en el período 1982-2020 . . . . .	97
5.2. Estimación de los parámetros de la DVEG por el método de máxima verosimilitud y por el método de los momentos pesados en el período 1982-2020 . . . . .	98

5.3. Estimación de los parámetros de la DVEG por el método de estimación conjunta para las primeras 21 estaciones meteorológicas ó pluviométricas para el periodo 1982-2020. . . . .	99
5.4. Estimación de los parámetros de la DVEG por el método de estimación conjunta para las últimas 20 estaciones meteorológicas ó pluviométricas para el periodo 1982-2020. . . . .	100
5.5. Estimación de los parámetros de la DVEG por el método de estimación conjunta para las primeras 21 estaciones meteorológicas ó pluviométricas para el periodo 1982-2020. . . . .	101
5.6. Estimación de los parámetros de la DVEG por el método de estimación conjunta para las últimas 20 estaciones meteorológicas ó pluviométricas para el periodo 1982-2020. . . . .	102
5.7. Estimación de los parámetros de la DVEG por el método de estimación conjunta para las primeras 19 estaciones meteorológicas ó pluviométricas finales para el periodo 1982-2020. . . . .	103
5.8. Estimación de los parámetros de la DVEG por el método de estimación conjunta para las últimas 19 estaciones meteorológicas ó pluviométricas finales para el período 1982-2020. . . . .	104
5.9. Valor del TIC para los seis modelos con menor TIC en estimación espacial mediante procesos máx-estables para el período 1982-2020 en 38 estaciones climatológicas y pluviométricas considerando una <i>DVE Gumbel</i> . . . . .	108
5.10. Transformación a Frechet Unitaria de los datos observados para los primeros y últimos trimestres de las 38 estaciones meteorológicas ó pluviométricas. . . . .	111

### Resumen

A lo largo del presente trabajo se realizará un análisis sobre las precipitaciones extremas en Uruguay con el objetivo de construir un modelo espacio-temporal que permita hacer inferencia sobre las mismas en todo el territorio, ya que estos eventos pueden ser considerados de alto riesgo, afectando a la población en diferentes esferas, primordialmente sociales y económicas.

En primera instancia se examinan los datos con los que se cuenta, depurándolos de posibles errores y datos faltantes. A su vez, se realiza un estudio sobre las precipitaciones en los últimos años, así como en la totalidad del período 1982-2020, para de esta forma poder empezar a comprender el comportamiento de estos eventos en nuestro país. Luego se muestran las técnicas utilizadas para lograr los resultados y se emplean las mismas. Por último se comentan las conclusiones a las que se llegaron y posibles caminos para continuar trabajando.

Las principales herramientas teóricas que se utilizaron para lograr el objetivo principal parten de la *Teoría de Valores Extremos* y la misma se toma como insumo para aplicar los *procesos máx-estables*. Estas teorías abarcan conocimientos en diferentes áreas, como conceptos de probabilidad, matemáticas, modelos estocásticos, modelos e inferencia estadística. Para cumplir con el objetivo principal se trabajó tomando diferentes técnicas las cuales fueron llevadas a la práctica mediante el software R.

Este trabajo se basó principalmente en la tesis de maestría de Leonardo Moreno, ver ([Moreno Romero, 2013](#)). Su trabajo buscó modelar las precipitaciones máximas del estado de Guanajuato, México, tanto a nivel local como global.

*Palabras claves:* Teoría de Valores Extremos, procesos máx-estables, precipitaciones, dependencia espacial.

# Capítulo 1

## Precipitaciones extremas en Uruguay y la base de datos

En este capítulo se realiza una revisión sobre los estudios referentes al fenómeno de las precipitaciones máximas en el territorio uruguayo y un breve análisis descriptivo de la base de datos.

### 1.1. Precipitaciones extremas en Uruguay

Actualmente están en auge los estudios relacionados al cambio climático y el consiguiente cambio de la distribución espacial y temporal de eventos extremos que esto ocasiona, tales como heladas, sequías, tormentas, olas de calor, temperaturas extremas mínimas o máximas, huracanes, tornados, lluvias fuertes e intensas, entre otros, ver ([Estenssoro Saavedra, 2010](#)). Estos eventos extremos violentos, repentinos y no deseados, son capaces de alterar la estructura social y económica de un país. Son muy importantes ya que tienen fuertes consecuencias sobre la vida de las personas y su entorno, por lo tanto, es vital conocerlos y poder predecirlos para tomar acciones pertinentes. Los eventos climáticos son algunos de los principales riesgos de la actualidad, ya que poseen una peligrosa combinación de capacidad destructiva y probabilidad de ocurrencia alta, [5.4](#).

## CAPÍTULO 1. PRECIPITACIONES EXTREMAS EN URUGUAY Y LA BASE DE DATOS

---

En este trabajo nos centraremos en el estudio de las precipitaciones extremas máximas en Uruguay, ya que las mismas tienen efectos negativos sobre las personas, y si se dan de forma repentina e intensa pueden llegar a causar inundaciones de magnitudes muy elevadas que dificultan actividades que benefician a las ciudades, como el turismo y perjudican la actividad económica en general. También podría afectar cierto tipo de cultivos importantes para los productores rurales y algunos campos donde se cría ganado. Esto es muy perjudicial para un país en gran parte agropecuario como Uruguay. Se calcula que aproximadamente el 93 % del territorio se destina a este sector y que la producción agrícola, ganadera y agroindustria representan el 70 % u 80 % del valor de la producción total del país, ver ([Inda y Mazzeo, 2012](#)).<sup>1</sup>

Así mismo las lluvias ocasionan inundaciones que aíslan ciudades al dejar bajo agua puentes y cortar rutas, inundan ciudades donde se debe evacuar a un gran número de personas de sus hogares dejándolas en una situación difícil, afectando su cotidianidad y provocando en algunos casos grandes pérdidas materiales e incluso vidas. Además, al afectar principalmente a las clases medias y bajas, estas catástrofes acrecientan desigualdades. Estos eventos poco habituales afectan las fuentes de energía como las represas. Uno de los antecedentes más actuales de inundaciones sucedió en el año 2019, considerada por muchas personas como la más grande de todos los tiempos junto con la ocurrida en el año 1986, ya que en respectivos años se registraron en un día particular lluvias de 204,9mm y 239mm.

Existe una alta demanda de investigación estadística para lograr estudiar estos eventos extremos. Ya que, en varios sectores de la industria son de suma importancia la aplicación de diferentes técnicas que permitan los cálculos de probabilidades de estos sucesos. De esta forma contemplar la posibilidad de que, en caso de ocurrir, las estructuras de ciertas construcciones estén preparadas, por ejemplo, estimar el alto de una represa o el tamaño de un desagüe en una calle<sup>2</sup>. Por estas razones se plantea

---

<sup>1</sup><http://http://www.ainfo.inia.uy/>

<sup>2</sup><https://ladiaria.com.uy/opinion/articulo/2019/10/con-los-pies-en-el-agua>

## 1.1. Precipitaciones extremas en Uruguay

---

como objetivo principal modelar a nivel global, con la mayor precisión posible las precipitaciones extremas máximas diarias en Uruguay.

En la Figura 1.1 se observan dos imágenes de los daños causados por estas lluvias en dos sectores del territorio uruguayo (Santa Lucia y Durazno). Más precisamente ocurridas en el año 2019, donde en Santa Lucia se consideró la inundación más grande de la historia.



Figura 1.1: Inundaciones en Santa Lucia y Durazno en el año 2019. Fuente: Santiago Porcal - Focouy y Sinae.

Uruguay está comprendido dentro de las latitudes subtropicales, en un área intermedia entre los trópicos y los extratropicales. Es un país que, a pesar de ser pequeño en territorio, con tan solo 176.215 kilómetros cuadrados, tiene variación climática entre norte y sur. Por un lado, los vientos del norte muy húmedos provenientes del Amazonas. Por otro lado, desde el sureste, los vientos provenientes del Océano Atlántico también están relacionados con las lluvias. Estas características se deben tener en cuenta a la hora de evaluar los diferentes resultados obtenidos y a su vez, podría ser un punto de partida para realizar un análisis comparativo de precipitaciones máximas entre el sur y norte de Uruguay. <sup>3</sup>.

---

-inundaciones-en-santa-lucia/

<sup>3</sup><https://ladiaria.com.uy/ciencia/articulo/2021/3/las-lluvias-extremas-de-nuestra-primavera-estan-conectadas-con-lo-que-sucede-dias-antes-en-indonesia-y-filipinas/>

## CAPÍTULO 1. PRECIPITACIONES EXTREMAS EN URUGUAY Y LA BASE DE DATOS

En los últimos años, en el territorio uruguayo se ha registrado un patrón, el cual indica a nivel de precipitaciones acumuladas que a medida que nos desplazamos hacia el norte del país los registros de precipitaciones acumuladas van en aumento.<sup>4</sup> En la imagen 1.2, se puede observar las precipitaciones acumuladas en Uruguay para dos años particulares, 2016 y 2017, observando el patrón anteriormente mencionado.

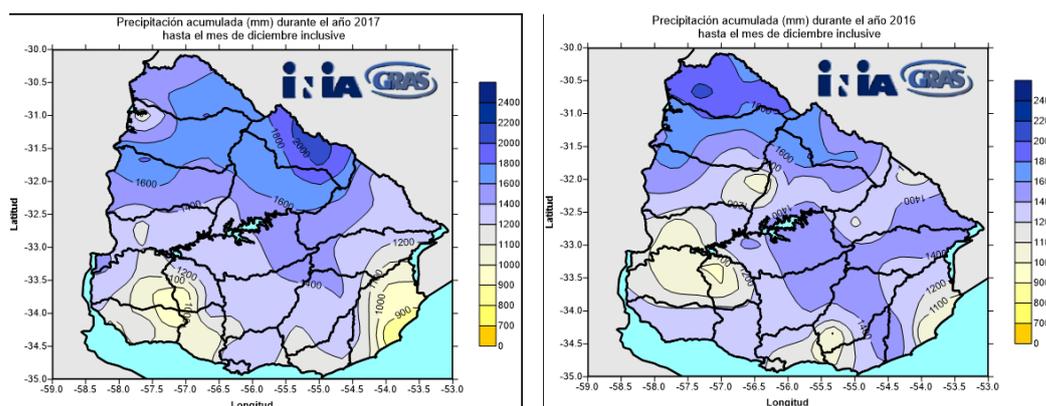


Figura 1.2: Mapa de precipitaciones acumuladas en Uruguay para el año 2016 y 2017. Fuente: Instituto Nacional de Investigación Agropecuario (INIA).

En la Figura 1.3 se puede apreciar el mapa de Uruguay dividido por sus Departamentos y el mapa hidrográfico. En este último, se observa una gran cantidad de arterias fluviales que atraviesan el país. Esto conlleva, en general, que ciertas precipitaciones de un caudal medio o alto determinen inundaciones en territorios y poblados en las cercanías de los ríos.

La estadística no ha permanecido ajena a esta problemática. Sin embargo, la búsqueda de la modelización de la cola de la distribución, donde hay pocos o incluso ningún dato, es un problema desafiante. Para poder hacer inferencia en estos casos se imponen ciertos supuestos (a veces implícitamente) que permiten la extrapolación de información. Diversos son los trabajos que abordan estos problemas. Como referencias generales, ver (Embrechts, Klüppelberg, y Mikosch, 2013; Beir-

[http://www.inia.org.uy/disciplinas/agroclima/fpta\\_pron\\_clim/vi\\_reunion/lluvias.html](http://www.inia.org.uy/disciplinas/agroclima/fpta_pron_clim/vi_reunion/lluvias.html)

<sup>4</sup><http://www.inia.org.uy/online/site/684740I1.php>

## 1.1. Precipitaciones extremas en Uruguay

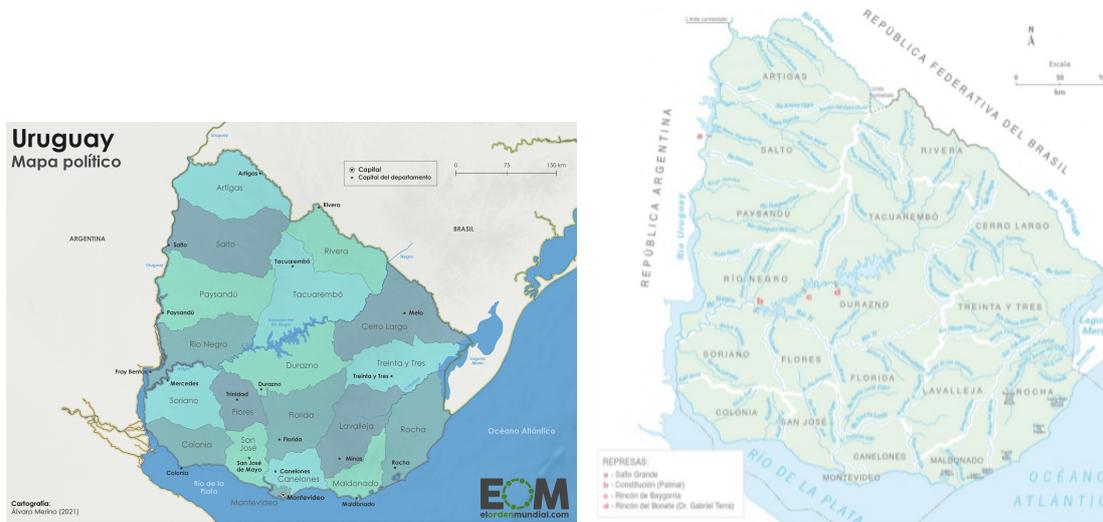


Figura 1.3: Mapa político e hidrográfico de Uruguay. Fuente: Álvaro Merino (2021) y <https://es.liveworksheets.com/qf2316268ov>

lant, Goegebeur, Segers, y Teugels, 2004; Smith, 1990a). Vemos que esta tipología de datos es frecuente en diversas áreas, ver por ejemplo (Salado, Huerta, Guzmán, Lazalde, y Salado, 2013; Dutfoy, Parey, y Roche, 2014). En particular el problema de la modelización de precipitaciones extremas es un problema de actual interés, y en donde se encuentran diversos estudios a nivel univariado, multivariado, o donde es modelizado el campo aleatorio de lluvias extremas en una determinada región, ver (Moreno Romero, 2013; E. M. González y Macías, 2011).

Existen diversos estudios en la región sobre las precipitaciones extremas, (D. S. González, 2011; Villacis Rivadeneira y Marrero de León, 2017; Khan, Kuhn, Ganguly, Erickson III, y Ostrouchov, 2007; Aparicio-Effen y cols., 2016), sin embargo son escasos los realizados específicamente en Uruguay sobre extremos. Por ejemplo en el 2019, Mathías Cardarello y Lorena Luraghi realizaron un estudio con el objetivo de modelizar las temperaturas mínimas extremas a partir de los mínimos diarios en la estación “La Estanzuela”, ubicada en el departamento de Colonia, Uruguay. Para lograr esto aplicaron Método de Valores Extremos por Bloques y Método de Excesos sobre un Umbral, ver (Cardarello y Luraghi, 2019). En referencia a precipitaciones extremas Florencia Santiñaque en su tesis de Maestría (2020) se planteó como prin-

## CAPÍTULO 1. PRECIPITACIONES EXTREMAS EN URUGUAY Y LA BASE DE DATOS

---

El principal objetivo en estudiar la existencia de patrones espaciales de las precipitaciones máximas anuales diarias en nuestro país utilizando diferentes técnicas. En primera instancia estudiando las distribuciones límite marginales de los valores extremos en cada estación. Luego aplicando técnicas espaciales de cluster. En dicho trabajo se utilizó una base de datos de precipitaciones diarias en 20 localizaciones del Uruguay para el período de enero 1981 a diciembre 2013. Y al igual que en el presente trabajo se utilizó una técnica llamada bloques de máximos, con la diferencia del tamaño de los mismos, en su trabajo se consideraron bloques anuales. Uno de los hallazgos fue que, en 18 localizaciones de las estudiadas, la distribución de Valores Extremos Generalizada (GEV) que mejor ajusta a los datos es del tipo Gumbel, ver (Santiñaque Mesones, 2020).

### 1.2. La base de datos

El objetivo principal es estimar la cola de la distribución, por lo tanto, es necesario contar con una base de datos de gran tamaño y una alta precisión. Las técnicas que se utilizarán funcionan con alta eficiencia en aquellas zonas donde hay una gran cantidad de observaciones en un período de tiempo considerablemente prolongado.

Los datos de las precipitaciones diarias de las diferentes estaciones fueron brindados por el Instituto Uruguayo de Meteorología (INUMET). La base de datos es abierta para el ámbito académico bajo pedido.<sup>5</sup>

Se cuenta con un total de 72 estaciones, de las mismas 19 pertenecían a estaciones meteorológicas y las otras 53 a estaciones pluviométricas. En la Figura 1.4 se observa la disposición del total de estaciones en el territorio uruguayo.

---

<sup>5</sup><https://www.inumet.gub.uy/institucional/transparencia/solicitud-informacion-meteorologica>.

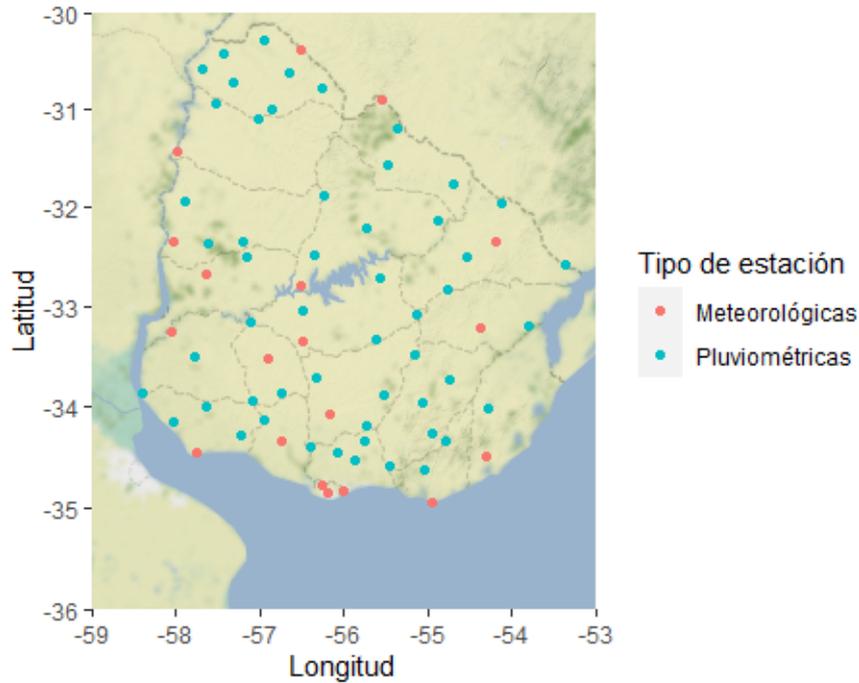


Figura 1.4: Ubicación de las 19 estaciones meteorológicas y 53 estaciones pluviométricas en Uruguay. Fuente: elaboración propia en R.

### 1.2.1. Filtrado inicial de la base

Para una correcta modelización de datos extremos, en una estación en particular, es necesario contar con una serie lo suficientemente larga y con una proporción baja de datos faltantes. Por otro lado, como nuestro objetivo es la modelización del campo extremo, que engloba al total de estaciones, también es necesario contar con un conjunto de las mismas que reporte las lluvias extremas en los mismos períodos de tiempo, es decir, trabajar en un lapso de tiempo donde todas las estaciones estén en funcionamiento simultáneamente. Para ello se realiza una primera depuración de la base en referencia a las estaciones y al período de tiempo para el análisis.

Mediante un primer estudio y analizando las Figura 1.5 la cual muestra el período activo diario de las estaciones. Se observó que en líneas generales eligiendo el período desde el año 1982 al año 2020 se conservan la mayor cantidad de estaciones que

## CAPÍTULO 1. PRECIPITACIONES EXTREMAS EN URUGUAY Y LA BASE DE DATOS

---

comprendan el mismo período, quedando 44 estaciones de las 72 iniciales. En general las estaciones en el período que estuvieron activas presentan registros diariamente interrumpidos, en algunos sucede que algún día no se logró un registro, pero en ninguna estación sucedió que por un período consecutivo de días (consideramos 15 días continuos) no se realizaron registros. En la Figura 1.5 se aprecian muy pocos intervalos con saltos. Lo que sucede es que un trimestre específico faltaron registrar una cantidad considerable de días. Pero estos días fueron espaciados en el tiempo, y por lo tanto, estos en las Figura 1.5 no se aprecian. A continuación, se realizó un análisis más profundo para los datos faltantes.

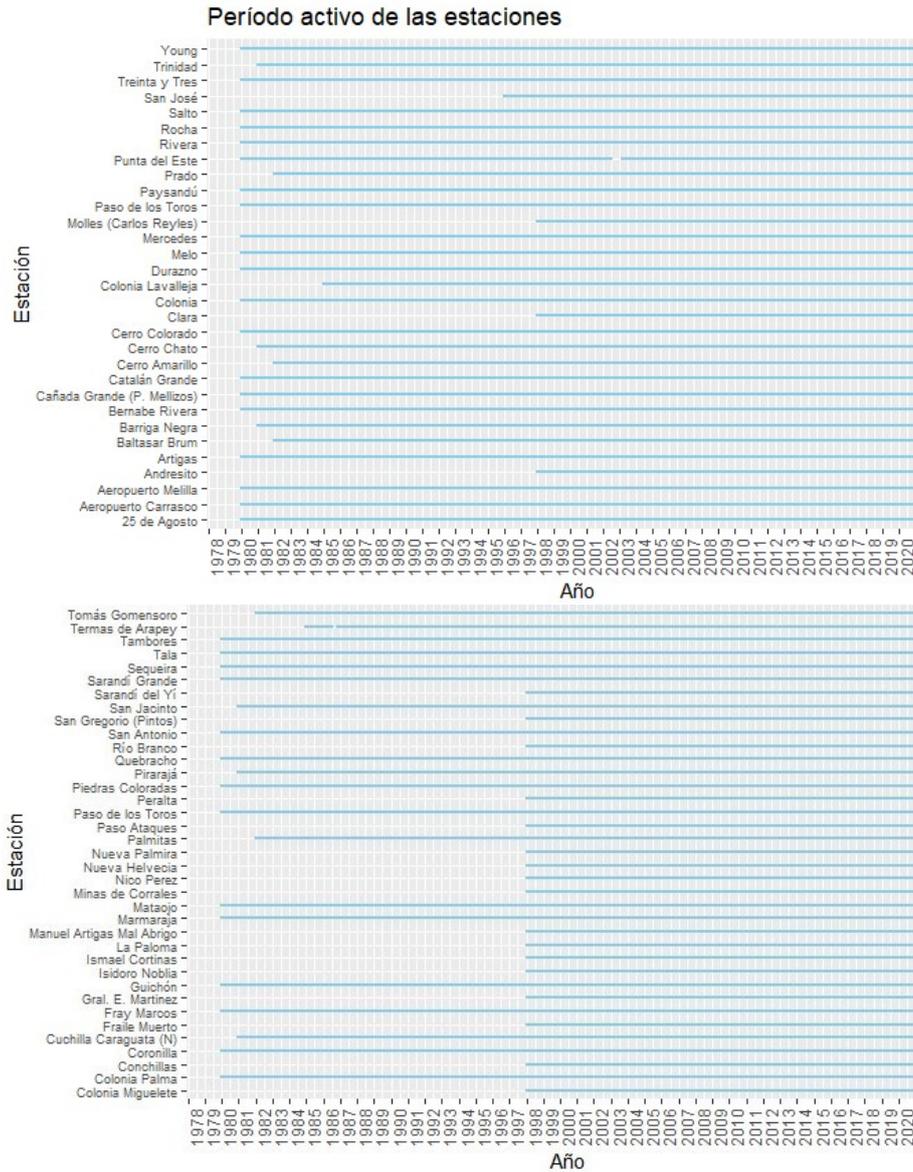


Figura 1.5: Período activo de las estaciones climatológicas y pluviométricas. Fuente: elaboración propia en R.

Una vez definido el período de estudio (1982-2020), se realiza un análisis descriptivo de los datos faltantes y datos atípicos. Este último punto es relevante para la correcta aplicación de la *Teoría de Valores Extremos (TVE)* y en los *procesos máx-estables*, puesto que un único outlier puede ocasionar un sesgo sustancial en las estimaciones. La metodología de estas técnicas es explicada en profundidad en el Capítulo 2, y en particular por el enfoque del *Método de Bloques*, el cual es de-

## CAPÍTULO 1. PRECIPITACIONES EXTREMAS EN URUGUAY Y LA BASE DE DATOS

---

sarrollado también en el Capítulo 2, ver 2, donde los máximos en cada estación son calculados en períodos sucesivos de tiempo. La selección de estos períodos (bloques) tienen que tener un balance, de forma que la cantidad de observaciones (precipitaciones diarias) dentro de cada bloque y a su vez el número de bloques sea lo suficientemente grande de forma de equilibrar el sesgo y la varianza del error de estimación, ver Capítulo 2.

Considerando este último punto se opta por trabajar con bloques trimestrales. En cada uno de ellos se estudia el porcentaje de datos faltantes, para estudiar su validez. Luego de filtrar los bloques se calcula en cada uno de ellos el máximo y se guarda en una nueva matriz de datos máximos.

Al estudiar la cantidad de datos faltantes por trimestres se considera una regla empírica la cual indica que aquellos trimestres que presentan un porcentaje de datos faltantes igual o mayor a 20 % no se deben considerar en el análisis y se estudia la posibilidad de eliminarlos. Para ello se plantean dos alternativas posibles, la primera es quitar este trimestre, por lo tanto, si se decide esto, hay que retirarlo de todas las estaciones para mantener la misma cantidad de bloques en cada estación. El segundo camino es sacar la estación meteorológica, de esta manera se conservaría el bloque, pero se perdería un punto de estimación del mapa de Uruguay. Para elegir la alternativa mas conveniente hay que tener en cuenta la ubicación de las estaciones, si la misma se encuentra relativamente cerca de otras estaciones es más plausible optar por quitarla, de lo contrario, si la estación cubre un lugar importante del territorio, es decir, donde se tienen pocas estaciones, será conveniente dejarla, teniendo en cuenta que hay que eliminar los trimestres o bloques correspondientes en todas las estaciones meteorológicas y pluviométricas. Para realizar este análisis de forma correcta nos basamos en la Figura 1.6 y en la Tabla 1.1.

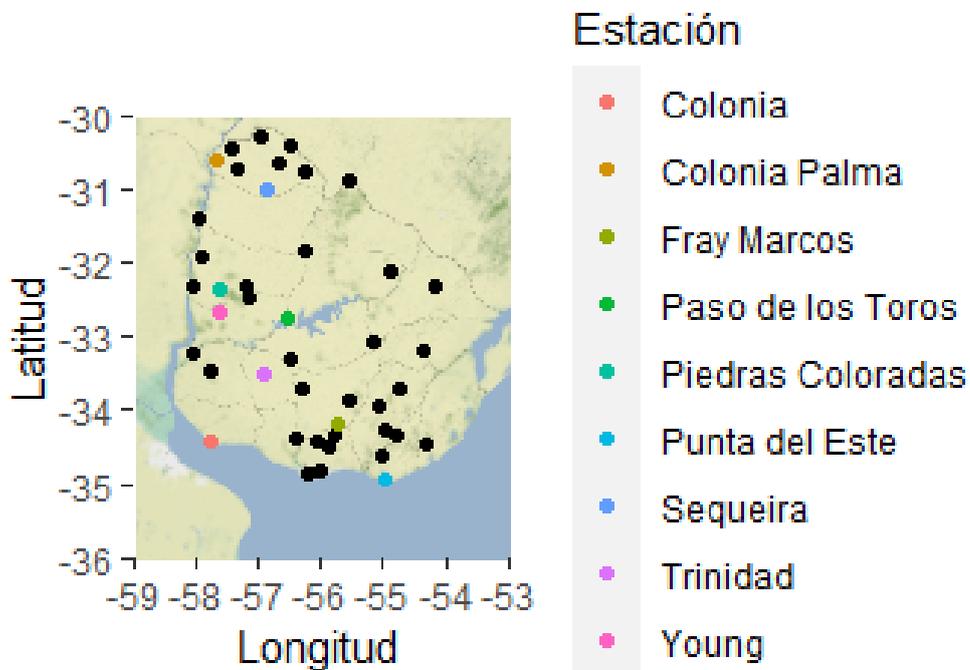


Figura 1.6: Ubicación de las 44 estaciones meteorológicas y pluviométricas en Uruguay en el período 1982-2020 y las estaciones en colores que presentan una cantidad de datos faltantes igual o mayor a 20%. Fuente: elaboración propia en R.

Mediante el software *R* se construyó una función que nos permite ver cuáles trimestres y en que estaciones se presentan un porcentaje igual o mayor a 20% de datos faltantes, ver Anexo 5.

CAPÍTULO 1. PRECIPITACIONES EXTREMAS EN URUGUAY Y LA BASE DE DATOS

	Estación	Año	Trimestre	% Datos faltantes
1	Colonia	2006	4	59 %
2	Colonia Palma	1993	3	36 %
3	Colonia Palma	1996	2	33 %
4	Colonia Palma	1996	3	34 %
5	Fray Marcos	1996	4	24 %
6	Paso de los Toros	2020	2	33 %
7	Paso de los Toros	2020	3	23 %
8	Punta del Este	2002	3	84 %
9	Punta del Este	2002	4	100 %
10	Punta del Este	2003	1	66 %
11	Sequeira	1990	2	33 %
12	Sequeira	1993	3	33 %
13	Sequeira	1996	3	34 %
14	Trinidad	2017	1	23 %
15	Young	2019	2	30 %
16	Young	2019	3	28 %
17	Young	2019	4	38 %

Cuadro 1.1: Estaciones junto con los trimestres y el año de las estaciones que presentan un porcentaje igual o mayor a un 20 % de datos faltantes.

En la Tabla 1.1 se observan 17 trimestres y 8 estaciones con un alto porcentaje de datos faltantes que superan el 20 %. Entonces, se estudia la ubicación de dichas estaciones mediante la Figura 1.6 y se decide retirar del análisis *Colonia Palma*, *Fray Marcos* y *Sequeira*, ya que las mismas están ubicadas cerca de otras estaciones meteorológicas o pluviométricas que pueden remplazarlas.

Por otra parte, se conservan las estaciones *Colonia*, *Paso de los Toros*, *Punta del Este*, *Trinidad* y *Young*. *Colonia* porque alrededor no hay una cantidad considerable

de estaciones y solo para un trimestre presenta una cantidad mayor al 20% de datos faltantes. La estación de *Trinidad*, al igual que la de *Colonia*, solo presenta problemas en un bloque por lo que no se quitó.

Por otro lado, la estación *Colonia* y *Punta del Este*, son estaciones costeras y se podría pensar que el clima en estas zonas se comporta distinto al interior del territorio, por lo tanto, se mantienen, ya que es importante capturar este comportamiento.

Por último, las estaciones de *Paso de los Toros* y *Young* se conservan por encontrarse alejadas de las demás estaciones.

La estación *Piedras Coloradas*, la cual también aparece en la Figura 1.6 es un caso atípico, no presenta problemas con los datos faltantes, sin embargo más adelante en el Capítulo 3 se explica porque se retira del análisis.

Conservando la estaciones antes mencionadas debemos retirar de nuestro análisis, el trimestre 4 del año 2006, los trimestres 2 y 3 del año 2020, los trimestres 3 y 4 del año 2002, el trimestre 1 del año 2003, el trimestre 1 del año 2017 y los trimestres 2,3 y 4 del año 2019 de todas las estaciones. Finalmente hay un total de 41 estaciones meteorológicas o pluviométricas y 147 trimestres (bloques).

## 1.3. Análisis Descriptivo de la base de datos

Una vez elegidas las estaciones y los trimestres en el período 1982-2020, se construyó una función en el software *R* denominada *máximo*, ver Anexo 5. La misma permite calcular los máximos de precipitaciones (en mm) por trimestres y estación, para luego poder aplicar la *Teoría de Valores Extremos* y llegar al objetivo final de nuestro trabajo aplicando los *procesos máx-estables* para modelar el campo de lluvias extremas máximas diarias en Uruguay.

Para observar el comportamiento de las lluvias máximas anuales y tener un pri-

## CAPÍTULO 1. PRECIPITACIONES EXTREMAS EN URUGUAY Y LA BASE DE DATOS

mer acercamiento a como se desarrollaron en el tiempo se presenta la Figura 1.7. Se logra ver, que el día en que se presentó la lluvia máxima en el período 1982-2020 se da en el año 1997. En varios años de dicho período al menos un día se registró una lluvia que supera los 200mm. También se aprecia que el año 2020 las precipitaciones fueron escasas y con registros relativamente bajos en comparación a los demás años. En líneas generales, la variabilidad anual de las precipitaciones máximas es homogénea, aunque en algunos años estas precipitaciones máximas presentaron una variabilidad superior a los otros años. Existen fenómenos climáticos vinculados con las precipitaciones, por ejemplo, el fenómeno del Niño y el fenómeno de la Niña, los cuales se podrían relacionar con las lluvias anuales, sin embargo no fueron estudiados con profundidad en este trabajo.

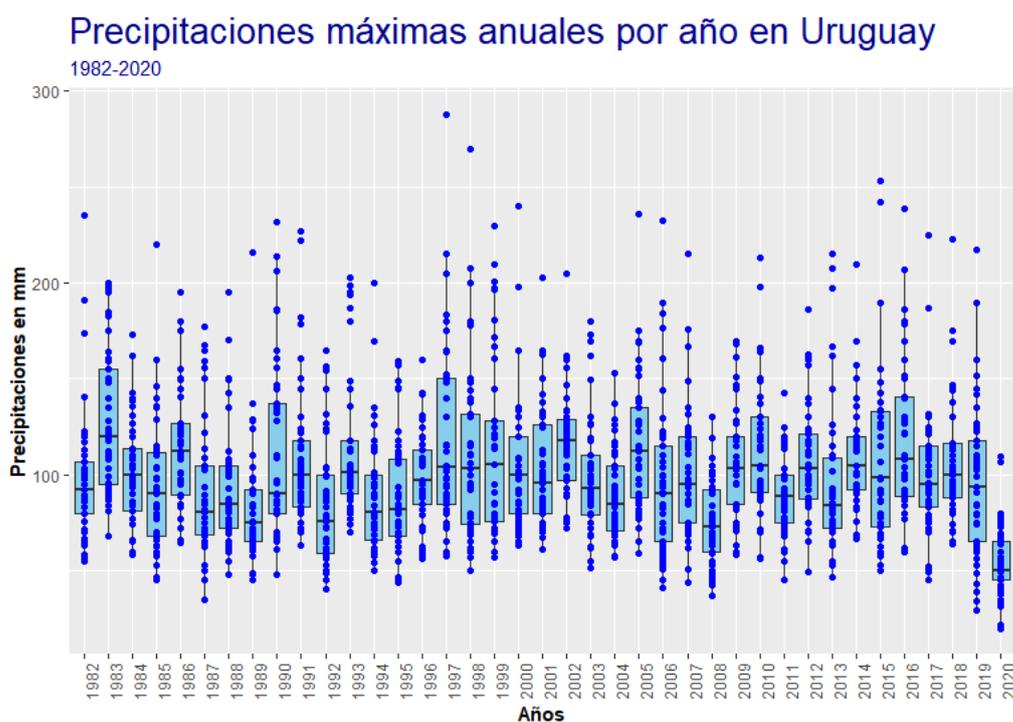


Figura 1.7: Precipitaciones máximas anuales registradas en Uruguay en las 41 estaciones en el período 1982-2020. Fuente: elaboración propia en R.

A continuación, en el Cuadro 1.2 a modo de resumen se muestran las máximas precipitaciones en mm en algunas estaciones climatológicas y pluviométricas. Se

### 1.3. Análisis Descriptivo de la base de datos

presentan los trimestres 1, 2, 3 y 4 en los años 1982 y 1983 para las estaciones de *25 de Agosto*, *Aeropuerto Carrasco*, *Aeropuerto Melilla*, *Artigas* y *Baltasar Brum*. Se observa que la máxima precipitación en estos primeros 8 trimestres se da en la estación *Artigas*, en el primer trimestre del año 1983, con un valor de 185 mm. Por otro lado, en el Cuadro 1.2 también se exponen los últimos 8 trimestres con los que se contaban. El trimestre 4 del año 2017, los trimestres 1, 2, 3 y 4 del año 2018, el trimestre 1 de los años 2019 y 2020 y el trimestre 4 del año 2020. La máxima lluvia que se registró en estos últimos 8 trimestres se dio en el trimestre 1 del año 2019 en la estación *Tomás Gomensoro*, donde se registró una precipitación de 123mm.

	1982-1	1982-2	1982-3	1982-4	1983-1	1983-2	1983-3	1983-4
25 de Agosto	42.00	73.00	86.00	39.00	55.00	34.00	112.00	49.00
Aeropuerto Carrasco	57.10	46.20	46.20	33.20	33.40	43.10	120.00	114.40
Aeropuerto Melilla	48.20	54.50	58.30	31.00	55.00	64.40	95.00	96.00
Artigas	121.40	62.70	83.20	102.40	185.00	138.80	26.40	42.50
Baltasar Brum	50.00	45.00	50.00	100.00	69.00	195.00	80.00	50.00
	2017-4	2018-1	2018-2	2018-3	2018-4	2019-1	2020-1	2020-4
Tambores	114.00	36.00	85.00	103.00	81.00	65.00	55.00	51.00
Tomás Gomensoro	28.00	46.00	92.00	58.00	60.00	123.00	54.00	57.00
Treinta y Tres	67.50	35.80	64.00	54.00	48.20	74.40	42.30	21.90
Trinidad	45.50	66.00	42.60	36.50	105.00	59.50	73.50	20.00
Young	48.80	32.80	46.00	77.00	107.20	108.70	109.40	31.00

Cuadro 1.2: Máximos de precipitaciones en mm de los últimos 8 trimestres en 5 estaciones

Con los datos obtenidos, máximos de las estaciones para cada trimestre se decidió realizar un estudio de la correlación entre estaciones. Para este objetivo se utilizó la correlación de *Kendall*, ver ([Badii, Guillen, Lugo Serrato, y Aguilar Garnica, 2014](#)). La misma se define de la siguiente forma:

$$\tau = \frac{(S_a - S_b)}{(n(n - 1))/2} \quad (1.1)$$

Donde:

- ▶  $S_a$ , es la sumatoria de los rangos más altos.
- ▶  $S_b$ , es la sumatoria de los rangos más bajos.
- ▶  $n$ , es la cantidad de máximos registrados por estación.

Las distribuciones de extremos en general no cuentan con segundo momento, para calcular el coeficiente de correlación clásico de Pearson, este segundo momento es necesario y es por esto que se usa el coeficiente de correlación de Kendall, debido a que el mismo no necesita que el segundo momento de las distribuciones exista.

Observando la Figura 1.8 se apreció que las estaciones meteorológicas o pluviométricas que presentan una distancia euclidiana más pequeña presenta una *correlación de Kendall* más grande. Por lo tanto, las precipitaciones máximas de estaciones que se encuentran más cercanas presentan un comportamiento similar y a medida que la distancia que las separa es mayor, las precipitaciones máximas no tiene un comportamiento semejante. De esto se puede inferir que tiene sentido la modelización conjunta de las estaciones y no de manera independiente.

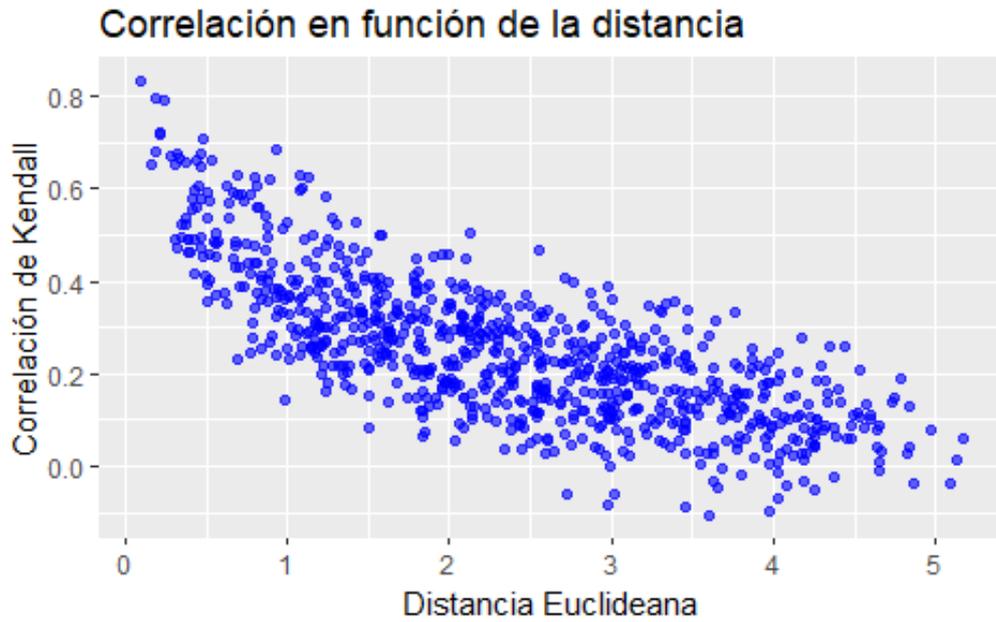


Figura 1.8: Correlación de Kendall para los máximos de las 41 estaciones meteorológicas y pluviométricas según la distancia euclidiana dos a dos. Fuente: elaboración propia en R.

# CAPÍTULO 1. PRECIPITACIONES EXTREMAS EN URUGUAY Y LA BASE DE DATOS

---

# Capítulo 2

## Metodología

En este capítulo se hace referencia y se detallan los conceptos teóricos vinculados al desarrollo de las técnicas utilizadas en esta monografía. El capítulo se divide en cuatro secciones. En la primera se realiza un breve resumen sobre la modelización de *valores extremos univariados*, en la segunda se desarrollan los *modelos máx-estables*, en la tercera se explican algunas diferentes medidas de dependencia en los *procesos máx-estables* y en la última sección se presentan algunos *criterios para seleccionar los modelos* que mejor se ajustan para las precipitaciones máximas diarias en Uruguay.

### 2.1. Teoría de Valores Extremos univariados

#### 2.1.1. Introducción

La *Teoría de Valores Extremos* (TVE) es una rama de la estadística que estudia el comportamiento de los valores extremos de una distribución. Consiste en un conjunto de técnicas estadísticas para la identificación y modelización de los máximos o mínimos de una variable aleatoria. El objetivo principal de dicha teoría es la extrapolación de información. La teoría tiene aplicaciones en varias áreas y diversas ramas de la ciencia, ingeniería, las finanzas pero principalmente en ciencias ambientales, más específicamente en el estudio de la *climatología*. En su versión más sencilla, el problema es el siguiente: dada una muestra independiente  $X_1, X_2, \dots$ ,

$X_n$  de una distribución desconocida  $F$ , queremos estimar la cola de la misma. Esto presenta ciertas dificultades, ya que se busca modelar las colas de la distribución  $F$  y en dichas colas las observaciones por definición van a ser pocas. Esto deriva en que las técnicas de estimación de densidades ajustan bien donde los datos tienen mayor densidad, pero pueden tener sesgos importantes al estimar las colas. Esta sección fue desarrollada en base a las notas del Dr. Joaquín Ortega Sánchez, sobre el curso "Valores Extremos" (Ortega Sanchez, 2010), la tesis de "Magister en Ingeniería Matemática" desarrollada por Leonardo Moreno sobre "Precipitaciones Máximas en el Estado de Guanajuato, México" (Moreno Romero, 2013), también nos basamos en el libro "*Extreme Value Theory*" desarrollado por Richard L. Smith (Smith, 1990a) y en el libro "*Extreme Value Theory: an introduction*" que brinda una introducción a la *Teoría de Valores Extremos* publicado por Laurens Hann y Ana Ferreira (De Haan y Ferreira, 2006).

### 2.1.2. Introducción al problema

Sea  $X_n$ ,  $n \geq 1$  una sucesión de variables independientes e idénticamente distribuidas con una función de distribución común  $F(x)$  desconocida, se define,

$$M_n = \max \{X_1, X_2, \dots, X_n\} \quad (2.1)$$

Uno de los objetivos principales como se mencionó en el *Capítulo 1* es estudiar la distribución del máximo de precipitaciones diarias para varias estaciones meteorológicas o pluviométricas ubicadas en Uruguay.  $M_n$  es la variable aleatoria máximo de las  $X_i$ , más precisamente el máximo trimestral de precipitación registrada en los 147 trimestres en el período 1982-2020 como se aludió en el *Capítulo 1*. Con  $X_i$  la precipitación acumulada en el día  $i$  en una estación determinada. Mediante esta teoría, se busca, para cada estación, estimar los parámetros de los modelos que a

continuación se mencionan, para poder usar estas estimaciones en las modelizaciones de los procesos marginales del *proceso máx-estable*.

La distribución de esta variable aleatoria es sencilla de calcular en función de  $F$ ,

$$P(M_n \leq x) = P(X_1 \leq x, \dots, X_n \leq x) = \prod_{i=1}^n P(X_i \leq x) = F^n(x). \quad (2.2)$$

Siendo  $F(x)$  la función de distribución de  $X_1$ . Pero en general dicha distribución de las  $X_i$  no es conocida, por lo tanto, la distribución del máximo  $M_n$  tampoco lo es, por ello, se buscan aproximaciones asintóticas. La *Teoría de Valores Extremos*, muestra que bajo ciertas características y comportamientos, ver (De Haan y Ferreira, 2006), se puede aproximar la distribución de  $M_n$  cuando  $n \rightarrow \infty$ .

Se denotan los extremos del soporte de la variable  $X_1$  como,

$$\alpha(F) = \inf\{x : F(x) > 0\} \geq -\infty \quad (2.3)$$

$$\omega(F) = \sup\{x : F(x) < 1\} \leq \infty \quad (2.4)$$

Por lo tanto, es evidente que  $P(M_n \leq x) = F^n(x)$  converge en distribución a la constante  $\omega(F)$  y al ser una constante, también se cumple la convergencia en probabilidad. Sumado a esto, como la sucesión  $M_n$  es creciente podemos afirmar la convergencia casi segura.

En primer instancia, si se conoce  $F$  la distribución de  $M_n$  es conocida, pero las expresiones analíticas para  $M_n$  pueden ser bastantes complejas. En general, la distribución de  $F$  es desconocida y si se desea tener alguna aproximación de la distribución

de  $M_n$ , es decir, buscamos alguna distribución límite que sirva de aproximación a  $F^n(x)$ . Este enfoque es similar a la práctica utilizada para aproximar la distribución de las medias de una muestra por una distribución Normal, como resultado del *Teorema de Limite Central (TLC)*.

### 2.1.3. Distribuciones para la TVE

Una vez que se cuenta con la matriz de máximos se debe considerar una distribución para los máximos por trimestres de precipitaciones de cada estación. Para seleccionar dicha distribución nos basamos en el siguiente Teorema, ver (Fisher y Tippett, 1928), donde la versión final del mismo viene dada por Gnedenko en 1943, ver (Gnedenko, 1943).

**Teorema 2.1.1** (*Teorema de Fisher-Tippett*) Dada una sucesión de variables aleatorias  $\{X_n\}_{n \geq 1}$  i.i.d, tal que,  $X_i \sim F$  y sea  $M_n = \max_{1 \leq i \leq n} \{X_i\}$ . Si existen sucesiones reales  $\{a_n\}_{n \geq 1}$  positiva y  $\{b_n\}_{n \geq 1}$  se debe cumplir lo siguiente,

$$P\left(\frac{M_n - b_n}{a_n} \leq x\right) = F^n(a_n x + b_n) \rightarrow G(x), \quad (2.5)$$

converge débilmente cuando  $n \rightarrow \infty$  y donde  $G$  es propia y no degenerada. Además,  $G$  pertenece a alguna de las siguientes distribuciones. A partir de la estimación de los parámetros, más específicamente estudiando la estimación del parámetro de forma  $\hat{\epsilon}$  se selecciona una de las tres distribuciones, según el siguiente criterio:

► Si  $\hat{\epsilon} < 0$  entonces se debe trabajar con la distribución de *Weibull*, cuya función de distribución es de la siguiente forma,

$$G_1(x, \mu, \sigma, \epsilon) = \begin{cases} \exp\left\{-\left(\frac{\mu - x}{\sigma}\right)^{-\frac{1}{\epsilon}}\right\} & \text{si } x \leq \mu, \epsilon < 0 \\ 0 & \text{en otro caso} \end{cases} \quad (2.6)$$

► Si  $\hat{\epsilon} > 0$  se trabaja con la distribución de *Frechet*, donde la función de distribución es la siguiente,

$$G_2(x, \mu, \sigma, \epsilon) = \begin{cases} \exp \left\{ - \left( \frac{x - \mu}{\sigma} \right)^{-\frac{1}{\epsilon}} \right\} & \text{si } x \geq \mu, \epsilon > 0 \\ 0 & \text{en otro caso} \end{cases} \quad (2.7)$$

► Por último, si  $\hat{\epsilon} = 0$ , ( $\hat{\epsilon} \rightarrow 0$ ) debemos trabajar con una distribución de *Gumbel*, y su función de distribución es,

$$G_3(x, \mu, \sigma) = \exp \left\{ - \exp \left( \frac{\mu - x}{\sigma} \right) \right\}; \sigma \neq 0 \quad (2.8)$$

Dentro de la *TVE* y considerando la familia de distribuciones de probabilidad continua, se puede combinar las distribuciones de *Weibull*, *Frechet* y *Gumbel* para conseguir una distribución conocida como *Distribución Generalizada de Valores Extremos (DGVE)*. El origen de la forma funcional común para las tres distribuciones se remonta al menos a Jenkinson en 1955, ver ([Jenkinson, 1955](#)), aunque también podría haber sido dado por Von Mises en 1936, ver ([Alves y Neves, 2011](#)). La *DGVE* es de la siguiente forma,

$$G_{\epsilon, \mu, \sigma}(x) = \exp \left\{ - \left( 1 + \epsilon \left( \frac{x - \mu}{\sigma} \right) \right)_+^{-\frac{1}{\epsilon}} \right\}, \mu \in \mathbb{R}, \sigma > 0, \epsilon \neq 0 \quad (2.9)$$

Cualquiera de las tres distribuciones se puede transformar a una distribución *Frechet unitaria*, cuya función de distribución es la siguiente,

$$G_{1,0,1}(x) = \begin{cases} \exp \left( -\frac{1}{x} \right) & \text{si } x \geq 0 \\ 0 & \text{en otro caso.} \end{cases} \quad (2.10)$$

**Propiedad 2.1.1 *Propiedad Frechet Unitaria.*** *Los siguientes tres enunciados son equivalentes:*

- ▶  $X \sim \text{Frechet}(\mu, \sigma, \epsilon)$ .
- ▶  $\log X^\alpha \sim \text{Gumbel}(\mu, \sigma)$ ,  $\sigma = 1/\epsilon$ .
- ▶  $-X^{-1} \sim \text{Weibull}(\mu, \sigma, \epsilon)$ .

#### 2.1.4. Estimación de parámetros para la DGVE

La *Distribución de Valores Extremos Generalizada (DGVE)* es de la forma (2.9).

Siendo el vector de parámetros  $\theta = (\epsilon, \mu, \sigma) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}^+$ , donde se encuentran el parámetro de forma  $\epsilon$ , el parámetro de ubicación  $\mu$  y el parámetro de escala  $\sigma$ .

Se obtiene una muestra de máximos agrupando los datos en conjuntos disjuntos de igual longitud mediante el *Método de Bloques*, ver 2.1.4.1, un método clásico de *Estadística de Extremos*.

Para estimar los parámetros mencionados anteriormente existen diferentes enfoques, de los cuales algunos se mencionan a continuación,

- ▶ El primer método es la estimación marginal de los parámetros para cada estación por separado. Existen dos técnicas para realizar esta estimación marginal, por un lado, la *Estimación por Máxima Verosimilitud* (Zhou, 2010), y, por otro lado, la estimación por el *Métodos de los Momentos Pesados* de cada estación de forma independiente. En el presente trabajo se utilizó la *Estimación por Máxima Verosimilitud*, dado que en el período de estudio se contaba con una cantidad considerablemente grande de observaciones.

- ▶ Un segundo enfoque fue propuesto por Padoan en 2010, (Padoan, Ribatet, y Sisson, 2010). Una característica atractiva de este enfoque es que la estimación de los parámetros marginales de *DGVE* se puede realizar junto con los parámetros de dependencia en un marco unificado. Se asume cierta similitud entre los parámetros

de una estación y otra si estas son cercanas. Esta estimación es razonable pensarla mediante un modelo de *regresión lineal* donde las variables predictoras son latitud y longitud de cada una de las estaciones meteorológicas o pluviométricas, mientras que la variable dependiente serían los parámetros de la DGVE de la ecuación 2.9.

#### 2.1.4.1. Ajuste de Máximos Trimestrales en el Método de Bloques

Se considera una colección de datos que se agrupan en  $n$  conjuntos (bloques) disjuntos y de igual longitud. Para cada bloque se identifica el valor máximo. Cada bloque contiene la información correspondiente a un período de tiempo fijo, un día, un trimestre o un año. Una vez que contamos con un conjunto de datos formado por los valores máximos por bloques se ajusta la *Distribución de Valores Extremos* al último conjunto mencionado. Los bloques son los siguientes,

$$\begin{aligned} X^{(1)} &= (X_1^{(1)}, \dots, X_n^{(1)}) \\ X^{(2)} &= (X_1^{(2)}, \dots, X_n^{(2)}) \\ X^{(s)} &= (X_1^{(s)}, \dots, X_n^{(s)}) \end{aligned}$$

Donde  $s$  indica la cantidad de estaciones disponibles luego del primer filtro estudiando los datos faltantes. A la hora de estimar se parte de 41 estaciones meteorológicas o pluviométricas. El valor de  $n$  indica la cantidad de bloques (intervalo de tiempo que se escoge), en el presente trabajo 147 trimestres y se supone que las componentes de cada vector  $(X^{(s)})_{i=1}^n$  son *iid* y cada vector se modela de forma independiente (este supuesto no siempre es real).

La muestra *iid* para  $G_\theta$  con la que se hará la inferencia es entonces,

$$X_l = \text{máx}(X_l^1, \dots, X_l^n), \quad l = 1, \dots, s \quad (2.11)$$

Anteriormente, se había mencionado, que las colas de las distribuciones  $F$ , pueden contar con pocas observaciones. Entonces, la selección del tamaño del bloque juega

un papel importante. Si por ejemplo se trabaja con bloques anuales, el valor de precipitación máxima registrada en cada bloque se aproxima más al valor real de una precipitación máxima. En este caso se contaba con menos observaciones que se elige bloques más chicos, y por lo tanto, existe un trade-off entre sesgo y varianza. Por este motivo, una elección plausible en la práctica consiste en trabajar con bloques trimestrales.

#### 2.1.4.2. Estimación por Máxima Verosimilitud

A la hora de seleccionar la técnica de *Estimación por Máxima Verosimilitud* se corroboró de manera empírica (comparando con otros estudios) que la cantidad de observaciones con las que se contaba era lo suficientemente grande para que las condiciones de regularidad necesarias para que las propiedades asintóticas habituales sean válidas asociadas con el *Método de Máxima Verosimilitud*.

Sea  $g_\theta$  la función de densidad de  $G_\theta$  (DGVE) y  $\theta = (\mu, \sigma, \epsilon)$ , la verosimilitud de  $X = (X_1, \dots, X_n)$  viene dada por,

$$L(\theta, X) = \prod_{i=1}^n g_\theta(X_i) \mathbb{I} \left( \left( 1 + \epsilon \left( \frac{X_i - \mu}{\sigma} \right) \right) > 0 \right) \quad (2.12)$$

Entonces, el algoritmo de la *función de log-verosimilitud* definida como  $\mathcal{L} = \log(L(\theta, X))$  viene dada por,

$$\log(\mathcal{L}(\theta, X)) = \begin{cases} -s \log(\sigma) - (1 + \frac{1}{\epsilon}) \sum_{i=1}^s \log[1 + \epsilon(\frac{X_i - \mu}{\sigma})] - \\ \sum_{i=1}^s \log[1 + \epsilon(\frac{X_i - \mu}{\sigma})]^{\frac{-1}{\epsilon}}, \text{ si } \epsilon \neq 0 \\ -s \log(\sigma) - \sum_{i=1}^s \exp[-(\frac{X_i - \mu}{\sigma})] - \sum_{i=1}^s (\frac{X_i - \mu}{\sigma}), \text{ si } \epsilon = 0. \end{cases} \quad (2.13)$$

Donde,

$$1 + \epsilon \left( \frac{X_i - \mu}{\sigma} \right) > 0 \text{ para } i = 1, \dots, s. \quad (2.14)$$

Maximizando la ecuación (2.13) respecto al vector de parámetros  $\mu, \sigma, \epsilon$  se consiguen los *estimadores máximo verosímiles* de cada uno para *DGVE*. Dicha expresión no se puede maximizar de forma analítica, por lo tanto, se utilizan algoritmos de optimización numérica, y la librería *fExtremes* brindada por el software *R*<sup>1</sup>.

Como se mencionó anteriormente es necesario corroborar las condiciones de regularidad necesarias para el cumplimiento de las propiedades asintóticas. Esto se realiza estudiando la estimación del parámetro de forma  $\epsilon$ , según los siguientes tres casos:

► Si  $\epsilon > -0,5$  entonces los estimadores máximos verosímiles de los tres parámetros  $(\hat{\mu}, \hat{\sigma}, \hat{\epsilon})$  presentan las propiedades asintóticas usuales. La cual, dice que dichos parámetros estimados presentan aproximadamente una distribución normal multivariada con media  $(\mu, \sigma, \epsilon)$  y la matriz de covarianza viene dada por el inverso de la matriz de información observada evaluada en los estimadores máximo verosímiles.

► Si  $-1 < \epsilon < -0,5$  se puede calcular los estimadores, sin embargo, estos no presentarán las propiedades asintóticas usuales mencionadas anteriormente, ver (Hosking, Wallis, y Wood, 1985).

► Si  $\epsilon < -1$  existe la posibilidad que los estimadores no se puedan obtener, ver (Hosking y cols., 1985).

#### 2.1.4.3. Estimación conjunta

La estimación de los parámetros de cada una de las estaciones se puede realizar mediante una regresión en función a determinadas variables de medidas que caracterizan a las mismas, cada parámetro a estimar  $(\mu, \sigma, \epsilon)$  se puede expresar,

---

<sup>1</sup><https://www.rdocumentation.org/packages/evd/versions/2.3-6/topics/fgev>

$$h\{p(x_k)\} = [X\beta]_k, \quad k = 1, 2, \dots, K. \quad (2.15)$$

Donde  $K$  indica el número de estaciones,  $p(\cdot)$  representa los parámetros  $\mu(\cdot)$ ,  $\sigma(\cdot)$  y  $\epsilon(\cdot)$ ,  $h$  es la función de enlace,  $x_k = (x_{k1}, x_{k2}, \dots, x_{kI})$  es el vector de predictores,  $X_{K \times I}$  es la matriz de diseño, y  $\beta_{Ix1}$  es el vector de parámetros desconocidos.

Cuando la dependencia espacial es alta y se va atenuando linealmente en términos de las distancias entre las estaciones, y además las estimaciones obtenidas son similares a los métodos marginales esta modelización es la más adecuada.

## 2.2. Procesos máx-estables

En esta sección, se presentan los principales conceptos y algunos modelos paramétricos máx-estables existentes en orden histórico. Los trabajos desarrollados por Balkema y Resnick en 1977, Laurens de Hann en 1978 y 1984 fueron los primeros avances para el desarrollo de la teoría, ver (Brown y Resnick, 1977), (Balkema y De Haan, 1978) y (De Haan, 1984). El primer modelo que se realizó en la literatura es el llamado *proceso de Smith*, en el año 1990, ver (Smith, 1990b). El segundo modelo se introdujo unos 10 años después en el artículo seminal de Schlather en 2002, quien incorporó el *proceso de Schlather*, también denominado a veces *proceso Gaussiano Extremo*, ver (Smith, 1990b). También se presentan otros modelos como Brown-Resnick presentado en el año 2009, Geométrico Gaussiano en el año 2012 y por último, el modelo t-extremal desarrollado en 2012 por Opitz, ver (Opitz, 2012).

### 2.2.1. Definiciones y teoremas principales

**Definición 2.2.1** (*Distribuciones Sum-Estables*) Dada una sucesión de variables aleatorias  $X_1, X_2, \dots, X_n$  i.i.d tal que  $X_1 \sim F$  (no degenerada).  $F$  es una distribución sum-estable, si existen constantes  $a_n > 0$  y  $b_n \in \mathbb{R}$  tales que,

$$\frac{\sum_{i=1}^n X_i - b_n}{a_n} \stackrel{d}{=} X_1, \forall n \geq 1. \quad (2.16)$$

Este mismo concepto se extiende pero en lugar de una suma de variables aleatorias se toma el máximo de ellas, para luego estudiar la estrecha relación con las *Distribuciones de Valores Extremos (DVE)*.

**Definición 2.2.2** (*Distribuciones Máx-Estables*)

Dada una sucesión de variables aleatorias  $X_1, X_2, \dots, X_n$  i.i.d tal que  $X_1 \sim F$  (no degenerada).  $F$  es una distribución máx-estable, si existen constantes  $a_n > 0$  y  $b_n \in \mathbb{R}$  tales que,

$$\max_{i=1 \dots n} \left\{ \frac{X_i - b_n}{a_n} \right\} \stackrel{d}{=} X_1, \forall n \geq 1. \quad (2.17)$$

Por lo tanto, la definición 2.2.2 mencionada anteriormente nos dice que es posible hacer un cambio de localización y escala a  $M_n = \max_{i=1 \dots n} X_i$ , de modo que, su distribución bajo este cambio, coincida con la distribución de las variables aleatorias originales. Dado que  $F_{M_n}(z) = F^n(z)$  la igualdad 2.17 es equivalente a,

$$F^n(a_n z + b_n) = F(z). \quad (2.18)$$

Lo anterior tiene una relación muy importante con la *Teoría de Valores Extremos*, como se enuncia en el siguiente Teorema.

**Teorema 2.2.1** (*Caracterización de la Distribución de Valores Extremos*)

Una distribución es máx-estable sí y solo sí es una distribución de Valores Extremos Generalizada.

Dada la sucesión  $X_1, X_2, \dots, X_n$  de campos aleatorios se está interesado en caracterizar  $Y(x)_{\{x \in K\}}$  el límite de distribución del siguiente proceso,

$$\max_{i=1\dots n} \left\{ \frac{X_i(x) - b_n(x)}{a_n(x)} \right\}_{x \in K} \xrightarrow{d} \{Y(x)\}_{x \in K}, \quad (2.19)$$

siendo  $Y$  sea un proceso no degenerado, con  $a_n$  y  $b_n$  funciones continuas en un compacto de  $\mathbb{R}^d$  y  $a_n$  una función positiva.

Los resultados presentados por el momento, corresponden únicamente al caso en que se trabaja con variables aleatorias, pero estos resultados se pueden extender para distribuciones multivariadas. Incluso, la propiedad de *máx-estabilidad* puede también ser definida cuando se trabaja con procesos, es aquí que L. de Haan en 1984 (De Haan, 1984), demuestra que los procesos límites coinciden con los *procesos máx-estables*.

**Definición 2.2.3** *Procesos máx-estables:* Sea  $K$  un subconjunto compacto de  $\mathbb{R}^d$  y sea  $Z \in C(K)$  un proceso estocástico con marginales no degeneradas, diremos que  $Z = \{Z(t), t \in T\}$ , es un proceso máx-estable si dadas  $n$  copias independientes de este proceso  $Z_1, \dots, Z_n$ , existes constantes  $a_n > 0$  y  $b_n \in \mathbb{R}$ , tal que,

$$Z \stackrel{d}{=} \max_{i=1\dots n} \left\{ \frac{Z_i - b_n}{a_n} \right\}, \forall n \in \mathbb{N}. \quad (2.20)$$

En general cuando se trabaja en el campo de procesos estocásticos es común considerar el conjunto de índices  $T \subset \mathbb{R}$ , tal que, el conjunto de índices está representando el tiempo en que el proceso evolucionó. En general, este conjunto de índices podría ser cualquier espacio topológico. En este caso, el proceso resultante recibe el nombre de *campo aleatorio*. Adler en 1981, ver (Ehm, 1981), presenta dos posibles definiciones formales de *campo aleatorio*.

Tomando la definición de 2.2.3 y eligiendo las constantes de normalización  $a_n(t) = n$  y  $b_n(t) = 0$ , es posible trabajar con distribuciones marginales de *Fréchet Unitarias*, ver (De Haan, 1984):

$$P(Z(t) \leq z) = e^{-\frac{1}{z}}, z > 0. \quad (2.21)$$

Y, por lo tanto,  $Z$  se define como un *proceso máx-estable simple*, donde:

$$Z \stackrel{d}{=} \max_{i=1 \dots n} \left\{ \frac{Z_i}{n} \right\}, \forall n \in \mathbb{N}. \quad (2.22)$$

De ahora en más se trabajara con distribuciones marginales *Fréchet Unitarias*.

### 2.2.2. Dependencia en procesos máx-estables

En esta sección se presentan algunos estadísticos para medir la dependencia en los *procesos máx-estables*. El objetivo es medir la dependencia a pares del campo aleatorio extremal.

#### 2.2.2.1. Índice Extremal

Una de las medidas de dependencia más básicas es conocida como *Índice Extremal*. A continuación, partiendo del siguiente Teorema y utilizando la definición de *Condición de Dependencia* ( $D(u_n)$ ), se llega al estimador del *Índice Extremal* ( $\hat{\theta}$ ):

**Teorema 2.2.2** *Sea  $\{X_n\}_{n \geq 1}$  una sucesión estacionaria (o estrictamente estacionaria) y  $\{X_n^*\}_{n \geq 1}$  una sucesión de variables aleatorias con la misma distribución  $F$  marginal. Se definió  $M_n = \max_{i=1 \dots n} \{X_n\}_{n \geq 1}$  y  $M_n^* = \max_{i=1 \dots n} \{X_n^*\}_{n \geq 1}$  y  $\theta \in (0, 1)$ . Bajo condiciones adecuadas de regularidad, se tiene:*

$$P\left(\frac{M_n^* - b_n}{a_n} \leq z\right) \rightarrow G_1(z). \quad (2.23)$$

con  $n \rightarrow \infty$ , donde  $a_n > 0$  y  $b_n \in \mathbb{R}$  sucesiones normalizadoras y  $G_1$  es una función de distribución no degenerada, si y solo si, la siguiente expresión es no degenerada:

$$P\left(\frac{M_n - b_n}{a_n} \leq z\right) \rightarrow G_2(z). \quad (2.24)$$

Donde:

$$G_2(z) = G_1(z)^\theta. \quad (2.25)$$

Gran parte de este desarrollo, incluidas algunas demostraciones son estudiadas por Leadbetter Lindgren en 2012, ver (Leadbetter, Lindgren, y Rootzén, 2012).

**Definición 2.2.4** *Condición de Dependencia:* Sea una sucesión estacionaria  $\{X_n\}_{n \geq 1}$ , satisface la condición  $D(u_n)$  si para cualesquiera  $i_1 < \dots < i_p < j_1 < \dots < j_q$ , con  $j_1 - i_p > l_n$ , se cumple que:

$$\begin{aligned} & | P(X_{i_1} \leq u_n, \dots, X_{i_p} \leq u_n, X_{j_1} \leq u_n, \dots, X_{j_q} \leq u_n) - \\ & P(X_{i_1} \leq u_n, \dots, X_{i_p} \leq u_n) P(X_{j_1} \leq u_n, \dots, X_{j_q} \leq u_n) | \leq \alpha(n, l_n). \end{aligned} \quad (2.26)$$

Donde  $\alpha(n, l_n) \rightarrow 0$  para alguna sucesión  $l_n$  tal que  $\frac{l_n}{n} \rightarrow 0$  cuando  $n \rightarrow \infty$ .

Entonces si el máximo de una serie estacionaria converge y se satisface la condición de  $D(u_n)$  apropiada, la distribución límite está relacionada con la distribución límite de una serie independiente según la ecuación 2.25 del Teorema 2.2.2. El parámetro  $\theta$  es conocido como *Índice Extremal*. Si se tiene  $\theta = 1$  se está en presencia

de un proceso completamente independiente, si  $\theta \rightarrow 0$  se tiene un proceso que tiene a dependencia total. Entonces, el efecto de dependencia en series estacionarias consiste en reemplazar a  $G_1$  como distribución límite por  $G_1^\theta$ . Si  $G_1$  corresponde a una *DGVE* con parámetros  $\mu, \sigma, \epsilon \neq 0$ , se contó:

$$G_{\mu, \sigma, \epsilon}(x)^\theta = \exp \left( - \left[ 1 + \epsilon \left( \frac{x - \mu}{\sigma} \right) \right]_+^{\frac{-1}{\epsilon}} \right)^\theta = \exp \left( -\theta \left[ 1 + \epsilon \left( \frac{x - \mu}{\sigma} \right) \right]_+^{\frac{-1}{\epsilon}} \right) = \exp \left( - \left[ 1 + \epsilon \left( \frac{x - \mu^*}{\sigma^*} \right) \right]_+^{\frac{-1}{\epsilon}} \right). \quad (2.27)$$

Donde:  $\mu^* = \mu - \frac{\sigma}{\epsilon}(1 - \theta^{-\epsilon})$  y  $\sigma^* = \sigma\theta^\epsilon$

### 2.2.2.2. Estimación del Índice Extremal

Para lograr una estimación del *Índice Extremal*, es necesario definir la siguiente ecuación:

$$P(M_n \leq u_n) \approx P^\theta(M_n^* \leq u_n) = F^{\theta n}(u_n). \quad (2.28)$$

Partiendo de la ecuación 2.28 se obtiene lo siguiente:

$$\lim_{n \rightarrow \infty} \frac{\log P(M_n \leq u_n)}{n \log F(u_n)} = \theta. \quad (2.29)$$

Esto sucede cuando  $n\bar{F}_n(u_n) \rightarrow \tau > 0$ . Donde  $\bar{F}_n(u_n)$  se define:

$$\bar{F}_n(u_n) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}_{\{X_i \geq u_n\}} = \frac{N}{n}. \quad (2.30)$$

Para lograr el estimador de  $\theta$  es necesario estimar por un lado  $F(u_n)$  y también  $P(M_n \leq u_n)$ . Para estimar  $F(u_n)$  se utiliza el teorema de *Glivenko-Cantelli*, obteniendo:

$$\hat{F}(u_n) = 1 - \frac{N}{n}. \quad (2.31)$$

Donde  $N$  indica el número de excedencias del umbral  $u_n$ . La estimación de  $P(M_n \leq u_n)$  se obtiene a partir de la condición  $D(u_n)$  y pensando  $n$  en  $k$  bloques de tamaño  $r$ , como  $(1 - \frac{K}{k})^k$ , siendo  $K$  el número de bloques con al menos una excedencia. Entonces se obtiene:

$$\hat{\theta} = \frac{k \log(1 - K/k)}{n \log(1 - N/n)} \quad (2.32)$$

### 2.2.2.3. Coeficiente Extremal

Para *procesos máx-estables*, existe una medida de dependencia para la *Teoría de Valores Extremos* en problemas multivariantes y espaciales. En este escenario, para problemas multivariados la medida de dependencia caracteriza la dependencia a nivel bivariado, para todos los pares y todos los órdenes superiores hasta la dimensión de la variable incluida. La función de *coeficiente extremal* se utiliza para establecer una condición necesaria y suficiente para que subconjuntos de medidas de dependencia sean autoconsistentes, es decir garantizar la existencia de una distribución con tal subconjunto de valores para la medida de dependencia. Partiendo de  $X = \{X_1, \dots, X_n\}$  una variable aleatoria  $n$ -dimensional, se puede trabajar con diferentes distribuciones marginales, en este caso *Frechet Unitarias*.

**Definición 2.2.5** *Coeficiente Extremal: Sea un vector  $(X_1, X_2)$  aleatorio bivariado máx-estable con distribuciones marginales idénticas, para todo  $E \subseteq \{1, \dots, n\}$ , existe un número real, con  $1 \leq \theta_E \leq |A|$ , tal que, el máximo normalizado de todas las*

variables indexadas por el conjunto  $A$  converge a una variable distribuida por Frechet con parámetro  $\theta_E$ ; eso es definido:

$$\lim_{n \rightarrow \infty} P \left( \max_{i \in A} \max_{j=1, \dots, n} \left( \frac{X_i^j}{n} \right) < z \right) = \lim_{n \rightarrow \infty} P \left( \max_{j=1, \dots, n} \left( \frac{X_i^j}{n} \right) < z \right)^{\theta_E} = e^{-\frac{\theta_E}{z}}, \quad \forall z > 0. \quad (2.33)$$

Donde  $X_i^j$  es la  $j$ -ésima replica de la variable  $X_i$  y  $\theta_E$  mide la dependencia extrema entre las variables indexadas en el conjunto  $E$  y a  $\theta_E$  se lo conoce como el *coeficiente extremal* de estas variables. La interpretación de este coeficiente es simple, es el número efectivo de variables independientes en el conjunto  $E$ . Se sabe que,  $\theta_i = 1$  y  $\theta_{\{i,j\}}$  mide la dependencia a pares de las variables aleatorias  $X_i$  y  $X_j$ . El *Coficiente Extremal* adopta valores en el intervalo  $[1, 2]$ , donde si  $\theta_{\{i,j\}} = 1$  indica que entre el par de variables existe una dependencia total, mientras si  $\theta_{\{i,j\}} = 2$  existe una independencia total.

La representación (2.33) de  $\theta_E$  surge como una propiedad de las clases completas de distribuciones límite para máximos por componentes de variables i.i.d  $X_1, \dots, X_n$ , esto es posible si se cumple las condiciones mencionadas por Resnick, ver (Resnick, 2008).

$$\lim_{n \rightarrow \infty} P \left( \max_{j=1, \dots, n} \left( \frac{X_i^j}{n} \right) < z_i; \quad i = 1, \dots, k \right) = G_k(z_1, \dots, z_k). \quad (2.34)$$

Donde el límite  $G_k$  pertenece a las clases de distribuciones multivariadas de la *Teoría de Valores Extremos*, asumiendo una distribución marginal *Frechet Unitaria*. La clase completa se caracteriza por la función  $V$ :

$$V(z_1, \dots, z_k) = -\log G(z_1, \dots, z_k) = \int_{S_k} \max_{i \in M_k} \left( \frac{w_i}{z_i} \right) dH_k(w_1, \dots, w_k). \quad (2.35)$$

Donde  $S_k$  es la unidad simple (m-1) dimensional y  $H_k$  es una medida que tiene todas sus esperanzas marginales iguales a 1. Combinando 2.33 con 2.35, se tiene:

$$\theta_E = \int_{S_k} \max_{i \in E} (w_i) dH_k(w_1, \dots, w_k). \quad (2.36)$$

Por lo tanto, si  $X$  sigue una distribución multivariada de *Valores Extremos*,  $\theta_E$  es una medida de dependencia exacta y no se basa en argumentos asintóticos.

**Definición 2.2.6** *Función del Coeficiente Extremal: Sea  $X(x)$  un campo aleatorio estacionario para  $x \in \mathbb{R}^d$  y marginales Frechet Unitarias. Sea  $X_i^j$  réplicas independiente del campo aleatorio  $X(x)$  para  $j = 1, \dots, n$ . Para estudiar la dependencia extrema del proceso  $X(\cdot)$  se enfocó en el máximo por componentes de las  $n$  réplicas. Entonces, De Hann en 1984, ver (De Haan, 1984), define una función de valor real  $\theta(\cdot)$ , tal que, la distribución asintótica del máximo normalizado en un par de ubicaciones separadas por una distancia  $h$  tiene una distribución de Frechet con el parámetro de escala  $\theta(h)$ , es decir, para  $z > 0$ :*

$$\lim_{n \rightarrow \infty} P \left( \max_{j=1, \dots, n} \frac{\max \{X^j(h), X^j(o)\}}{n} \leq z \right) = e^{-\frac{\theta(h)}{z}}, \quad z > 0. \quad (2.37)$$

para todo  $h \in \mathbb{R}^d$ , donde  $o$  denota el origen de coordenadas. Entonces, se denomina  $\theta(\cdot)$  a la función del Coeficiente Extremal. Esta función proporciona suficiente información sobre la dependencia extrema.

El Coeficiente Extremal presenta diferentes propiedades, algunas se enumeran a continuación:

►  $1 \leq \theta(h) \leq 2$ , donde el límite inferior y el límite superior indican dependencia e independencia total en la distancia  $h$  respectivamente.

►  $\theta(0) = 1$ .

► Si  $Y$  es un *proceso máx-estable* la función del *coeficiente extremal* es una medida de dependencia exacta.

► La función  $2 - \theta(h)$  es una función semidefinida positiva.

►  $\theta(\cdot)$  no es diferenciable en el origen, salvo que  $\theta(h) = 1, \forall h$ .

Se presentan algunas funciones del *Coeficiente Extremal* utilizadas en el trabajo, vinculadas con los modelos presentados en 2.2.3

### Modelo Smith

$$\theta(h) = 2\Phi\left(\frac{\sqrt{h^t \Sigma^{-1} h}}{2}\right). \quad (2.38)$$

### Modelo Schlather

$$\theta(h) = 1 + \sqrt{\frac{1 - \rho(h)}{2}}. \quad (2.39)$$

### Modelo de Brown-Resnick

$$\theta(h) = 2\Phi\left(\frac{\sqrt{V(Y(h))}}{2}\right). \quad (2.40)$$

### Modelo Geométrico Gaussiano

$$\theta(h) = 2\Phi\left(\frac{\sqrt{\sigma^2(1 - \rho(h))}}{2}\right). \quad (2.41)$$

### Modelo t-extremal

$$\theta(h) = 2T_{v-1} \left( \sqrt{\frac{v-1}{1-\rho^2(h)}} (1-\rho(h)) - \sqrt{\frac{1-\rho^2(h)}{v+1}} \rho(h) \right). \quad (2.42)$$

#### 2.2.2.4. Madograma y F-Madograma

Uno de los puntos más importantes para estudiar la dependencia espacial entre pares de estaciones meteorológicas y pluviométricas. Una herramienta fue la mencionada anteriormente, el *Coeficiente Extremal*, a la cual se le suma un componente esencial para lograr un mejor entendimiento de la dependencia y que permite tener control del proceso de interpolación que es usado para construir el modelo espacial.

En principio se define el *Variograma* para tener una primera aproximación a una función que da información respecto a las distancias entre estaciones meteorológicas y pluviométricas.

**Definición 2.2.7** *Variograma:* Sea  $Z(x)$  un proceso espacial estacionario, se define *Variograma* a la siguiente función:

$$\gamma(h) = \frac{1}{2} E[(Z(x+h) - Z(x))^2]. \quad (2.43)$$

Este concepto se incorpora en la geoestadística por Cressie en 1993, ver (Cooley, Naveau, y Poncet, 2006).

Sin embargo, el *Variograma* no siempre es adecuado, ya que en la *Teoría de Valores Extremos* es posible que en ocasiones la varianza de las distribuciones de Valores Extremos no exista. Es en este contexto que se opta por el Madograma, ver (Cressie, 1993), el cual evita el uso de momentos de segundo orden.

**Definición 2.2.8** *Madograma:* Sea  $Z(x)$  un proceso espacial estacionario, se define *Madograma* a la siguiente función:

$$v(h) = \frac{1}{2}E[|Z(x+h) - Z(x)|]. \quad (2.44)$$

También se sabe que para *procesos máx-estables* simples, las distribuciones marginales se transforman a *Frechet Unitarias*, y para estas distribuciones se cumple que  $E[Z(x)] = \infty$ ,  $\forall x \in \chi$ , por esto, tanto la función de *Variograma* como el *Madograma* no existen. Para solucionar este inconveniente se incorporan dos funciones alternativas, el *Madograma re-escalado* y el *F-Madograma*.

En este punto es importante destacar que existen relaciones entre los *Madogramas* y el *coeficiente extremal*. Estos avances en la definición de los *Madogramas* que proporcionan definiciones alternativas para solucionar ciertos problemas, así como la relaciones entre estos y el *coeficiente extremal*, son los que permiten incorporarlos a los *procesos máx-estable*, ver (Cooley y cols., 2006).

El *Madograma re-escalado* se define para  $t \in (0, \infty)$  de la siguiente forma:

$$\eta_t(h) = \frac{1}{2}E[|I_{(Z(x+h) \leq t)} - I_{(Z(x) \leq t)}|]. \quad (2.45)$$

### 2.2.2.5. Relación entre el Madograma re-escalado y el coeficiente extremal

Primero, es necesario observar la relación entre el Madograma y el coeficiente extremal obtenida por Cooley, Naveau y Poncet en 2006, ver (Cooley y cols., 2006), a partir de Hosking, Wallis y Wood en 1985, ver (Hosking y cols., 1985).

**Teorema 2.2.3** *Sea  $\theta$  el coeficiente extremal definido en un campo aleatorio máx-estable con distribuciones marginales Frechet Unitarias, el parámetro de forma  $\epsilon$ , el parámetro de escala  $\sigma$  y  $\tau(\cdot)$  la función gamma, entonces se cumple que:*

$$\theta(h) = \begin{cases} u_\beta \left( u + \frac{v(h)}{\tau(1-\epsilon)} \right), & \text{si } \epsilon \neq 0, \epsilon \leq 1 \\ \exp \left( \frac{v(h)}{\sigma} \right), & \text{si } \epsilon = 0. \end{cases} \quad (2.46)$$

Siendo:

$$u_\beta(x) = \left( 1 + \epsilon \left( \frac{x-u}{\sigma} \right) \right). \quad (2.47)$$

Entonces se tiene que para un *campo aleatorio máx-estable* con distribuciones marginales *Frechet Unitarias*:

$$\eta_t(h) = t \left( e^{\frac{-1}{t}} - e^{\frac{-\theta(h)}{t}} \right). \quad (2.48)$$

Y existe un vínculo entre el límite cuando  $t \rightarrow \infty$  de esta última función con la función de *coeficiente extremal* de la siguiente forma:

$$\eta(h) = \lim_{t \rightarrow \infty} \eta_t(h) = \theta(h) - 1. \quad (2.49)$$

Para el cálculo del *Madograma re-escalado* se define el siguiente estadístico:

**Definición 2.2.9** Sea  $\mathcal{N}_h = [(x, y)/d(x, y) = h]$ , es decir el conjunto de los puntos de la muestra que distan  $h$ , entonces:

$$\hat{\eta}_t(h) = \frac{1}{2|\mathcal{N}_h|} \sum_{x, y \in \mathcal{N}_h} |tI_{(Z(x+h) \leq t)} - tI_{(Z(y) \leq t)}|. \quad (2.50)$$

Llegados a este punto se está en condiciones de poder aplicar la teoría correctamente, ya que se llegó a un estadístico que existe siempre y fácilmente puede ser vinculado al *coeficiente extremal*. Por otro lado, se advierte que el *Madograma re-escalado* puede mejorar su eficiencia si se reemplaza por el *F-Madograma*, el cuál asume que las marginales provienen de la *Distribución de Valores Extremos Generalizados*, ver (Cooley y cols., 2006).

**Definición 2.2.10** *F-Madograma*: Sea  $F$  la función acumulada de  $Z(x)$  se define *F-Madograma* a la siguiente forma:

$$v_F(h) = \frac{1}{2} E[|F(Z(x+h)) - F(Z(x))|]. \quad (2.51)$$

El *F-Madograma* toma valores de 0, el cual indica dependencia máxima, hasta 1/6 el cual indica independencia total. Además es bien definido ya que  $F(Z(x)) \sim U(0, 1)$  por lo tanto su esperanza es 1/2.

Al igual que el *Madograma re-escalado* posee una relación con el *Coficiente Extremal* fácil de expresar.

Usando la siguiente igualdad  $|b - a| = 2\max(a, b) - b - a$ , se obtuvo:

$$\theta(h) = \frac{1 + 2v_F(h)}{1 - 2v_F(h)}. \quad (2.52)$$

En la práctica, el *F-Madograma* se estimará de la siguiente forma, ver (Ribatet, Dombry, y Oesting, 2016):

Sea  $y_i(x)$ ,  $i = 1, \dots, M$ , son copias *i.i.d* de un *proceso máx-estable* con distribuciones marginales  $F$  entonces:

$$\hat{v}_F(h) = \frac{1}{M} \sum_{i=1}^M \frac{1}{2|\mathcal{N}(h)|} \sum_{(x,x') \in \mathcal{N}(h)} |F(y_i(x)) - F(y_i(x'))|. \quad (2.53)$$

### 2.2.3. Modelos aplicados

La *Teoría de Valor Extremo* se ha desarrollado tradicionalmente para series univariadas, en las que se consideran los extremos de una sola variable, en el trabajo las estaciones meteorológicas y pluviométricas. Los trabajos recientes de Smith (1989), ver (Smith, 1990b) y Davison y Smith (1990), ver (Davison y Smith, 1990), incluyen modelos de regresión que permiten analizar la dependencia entre las estaciones meteorológicas y pluviométricas y las covariables relevantes, pero aún en el contexto de la estimación de valores extremos para una sola serie.

#### Modelo de Smith

Este modelo es una pieza fundamental de este trabajo, brindando herramientas importantes a la hora de modelizar campos aleatorios, al momento de realizar un análisis extremal de los datos de lluvias en el territorio uruguayo presentados más adelante. Lo que presenta *Smith (1990)*, ver (Smith, 1990b), se ve reflejado en el siguiente teorema:

**Teorema 2.2.4** Sean  $\{(u_1, s_1), (u_2, s_2), \dots\}$  puntos de un proceso puntual de Poisson en un espacio  $(0, \infty) \times S$  con medida de intensidad  $\lambda(du, ds) = \frac{du}{u^2} \nu(ds)$ , para alguna medida de intensidad  $\nu$  en  $S$ , donde  $S$  es un conjunto medible arbitrario. Sea  $f$  una función  $S \times T$  tal que  $\int_S f(s, t) \nu(ds) = 1, \forall t \in T$ , entonces el proceso  $Z(t), z \in T$ , definido por:

$$Z(t) = \max_{i=1,2,\dots} [u_i f(s_i, t)]. \quad (2.54)$$

Es un proceso máx-estable, y más aún, todo proceso máx-estable en  $T$  se puede representar de esta manera eligiendo de forma correcta  $S, \nu$  y  $f$ .

La demostración se puede encontrar en (Smith, 1990b).

Mediante el teorema 2.2.4 y tomando  $S = T = \mathbb{R}^d$  y  $f(s, t)$  como función de  $s$  para  $t$  fijo, donde la misma es una densidad normal multivariada con vector de medias  $t$  y matriz de varianza y covarianza  $\Sigma$ :

$$f(s, t) = f_0(s - t) = \frac{1}{(2\pi)^{\frac{d}{2}}} \exp\left\{-\frac{1}{2}(s - t)' \Sigma (s - t)\right\}. \quad (2.55)$$

Por lo tanto, el modelo de Smith para dos elementos del *proceso máx-estable* queda determinado por la siguiente ecuación:

$$P(Z(s) \leq z_1, Z(t) \leq z_2) = \exp\left\{-\frac{1}{z_1} \Phi\left(\frac{a}{2} + \frac{1}{a} \log \frac{z_2}{z_1}\right) - \frac{1}{z_2} \Phi\left(\frac{a}{2} + \frac{1}{a} \log \frac{z_1}{z_2}\right)\right\}. \quad (2.56)$$

Donde:

$a^2 = (s - t)' \Sigma (s - t)$  es la distancia de Mahalanobis entre  $s$  y  $t$  y  $\Phi$  es la función de distribución normal estándar.

El modelo desarrollado por Smith en 1990 presenta formas de construir *procesos máx-estables* con distribuciones bivariadas cuyas expresiones son conocidas.

Este proceso también es conocido como *proceso gaussiano de valores extremos*. Este modelo supone que el proceso en que se estudia es homogéneo en el espacio, es decir, no le da importancia a la ubicación específica de dos puntos, si no a la distancia que los separa. Por esto, en general se usa para modelar un tipo de lluvias llamadas lluvias convectivas, las mismas ocurren en territorios llanos o que tengan pequeñas imperfecciones, son propias de climas tropicales, ya sea cálidos o húmedos y se caracterizan por tener mucha intensidad en su centro e ir disminuyendo conforme aumenta la distancia respecto del mismo.

Como ya se mencionó anteriormente, todo *proceso máx-estable* se puede representar utilizando 2.18 eligiendo de forma correcta los parámetros  $S$ ,  $\nu$  y  $f$ . Entonces,

al utilizar el *modelo de Smith* se tiene que,  $\nu$  es una medida que muestra como se distribuye la tormenta sobre  $S$ ,  $u_i f(s_i, t)$  refleja la cantidad de lluvia de una tormenta de tamaño  $u_i$  en la posición  $t$  que está centrada en  $s_i$ , y la función  $f$ , por ser el *modelo de Smith*, es una función con forma de campana que se determinó al utilizar la *distribución normal bivariada* en la construcción del proceso, la misma muestra la forma de la tormenta que se quiere modelar.

## Modelo de Schlather

Un segundo modelo importante a la hora de modelar las precipitaciones máximas de estaciones meteorológicas y pluviométricas, es el *modelo de Schlather*, ver (Schlather, 2002). El mismo, considera  $Y_i(x) = \sqrt{2\pi} \max\{0, \epsilon_i(x)\}$ , donde  $\epsilon_i$  son copias independientes de un proceso gaussiano estándar con correlación  $\rho$ . Por lo tanto, la función de distribución acumulada bivariada se define de la siguiente forma:

$$P(Z(s) \leq z_1, Z(t) \leq z_2) = \exp \left\{ -\frac{1}{2} \Phi \left( \frac{1}{z_1} + \frac{1}{z_2} \right) \left( 1 + \sqrt{1 - \frac{2\{1 + \rho(h)\}z_1 z_2}{(z_1 + z_2)^2}} \right) \right\}. \quad (2.57)$$

Donde  $h = s - t$ .

A continuación se presentan algunas de las posibles correlaciones  $\rho$  de mayor importancia a la hora de conseguir los resultados, ver (Gaetan y Guyon, 2010):

### Power-Exponencial

$$\rho(h) = \exp \left[ - \left( \frac{h}{c_2} \right)^\nu \right]; c_2 > 0, 0 < \nu < 2. \quad (2.58)$$

### Whittle-Matérn

$$\rho(h) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left( \frac{h}{c_2} \right)^\nu K_\nu \left( \frac{h}{c_2} \right); c_2 > 0, \nu > 0. \quad (2.59)$$

**Cauchy**

$$\rho(h) = \left(1 + \left(\frac{h}{c_2}\right)^2\right)^{-\nu}; \quad c_2 > 0, \nu > 0. \quad (2.60)$$

**Bessel**

$$\rho(h) = \left(\frac{2c_2}{h}\right)^\nu \Gamma(\nu + 1) J_\nu\left(\frac{h}{c_2}\right); \quad c_2 > 0, \nu \geq \frac{d-2}{2}. \quad (2.61)$$

Donde  $h \in \mathbb{R}^+$  es la distancia euclidiana entre la estación  $i$  con la estación  $s$ . Mientras los parámetros  $c_2$  y  $\nu$  son el rango y el el parámetro de suavizado de la función,  $\Gamma$  es la función Gamma y  $J_\nu$  y  $K_\nu$  son la función de *Bessel* y *Bessel modificada* de tercer tipo con orden  $\nu$  y  $d$  es la dimensión de los campos aleatorios. El coeficiente de correlación  $\rho(h)$  puede tomar cualquier valor en  $[-1, 1]$ . La dependencia completa se alcanza cuando  $\rho(h) = 1$  mientras que la independencia ocurre cuando  $\rho(h) = -1$ . Sin embargo, la mayoría de las familias de correlaciones paramétricas no permiten valores negativos por lo que nunca se alcanza la independencia.

De acuerdo a los *procesos gaussianos*, es posible agregar un umbral  $c_1$  y un *efecto nugget*  $\nu$  a estas funciones de correlación.

$$\rho_*(h) = \begin{cases} \nu + c_1, & h = 0 \\ c_1 \rho(h), & h > 0 \end{cases} \quad (2.62)$$

Donde  $\rho(h)$  es una de las funciones de correlación presentadas anteriormente. Sin embargo, como Schlather (2002), ver (Schlather, 2002), considera *procesos gaussianos estándar estacionarios*, los parámetros de umbral y nugget satisfacen,  $\nu = 1 - c_1$ , porque la función de correlación tiene que ser igual a 1 en el origen.

El *efecto nugget* es un fenómeno presente en muchas variables regionalizadas y representa aleatoriedad o ruido a escala corta en la variable regionalizada. Se puede ver en la gráfica del *variograma* como una discontinuidad en el origen de la función. El *efecto nugget* se ve afectado por el volumen de muestreo, disminuyendo en valor a medida que aumenta el volumen. El *efecto nugget* es un término geostatístico utilizado para describir la variabilidad observada entre observaciones que están estrechamente espaciadas. El *efecto nugget* en la variografía se define como la discontinuidad observada en el origen donde a distancias de separación de muestra insignificantes todavía se observa variabilidad entre pares de observaciones. El parámetro de *umbral*, es un valor fijo de última instancia, es el límite del variograma, tendiendo las distancias a infinitos. En la imagen 2.1, se pueden ver estos parámetros, *nugget*, *umbral* y *rango*. Estas correlaciones presentadas, son las que se utilizan a la hora de estimar el *coeficiente extremal*  $\theta(h)$  en la ecuación (2.37).

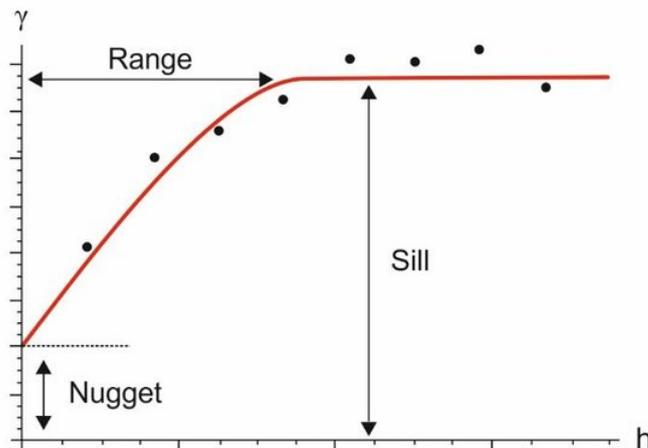


Figura 2.1: Ejemplo de un variograma experimental, que muestra los parámetros nugget, umbral y rango. Fuente: Porras (2014)

## Modelo de Brown-Resnick

Un tercer modelo es el de Brown-Resnick, ver (Brown y Resnick, 1977), donde se considera lo siguiente,  $Y_i(x) = \exp \left\{ \epsilon_i(x) - \frac{\sigma^2(x)}{2} \right\}$ , donde  $\epsilon_i$  al igual que el *modelo*

de *Shlather* son copias independientes de un proceso gaussiano con incrementos estacionarios y tal que  $Var\{\epsilon(x)\} = \sigma^2(x)$  para todo  $x \in \mathcal{X}$ . Su función de distribución viene dada por:

$$P(Z(s) \leq z_1, Z(t) \leq z_2) = \exp \left\{ -\frac{1}{z_1} \Phi \left( \frac{\sqrt{Var\{Y(z_1 - z_2)\}}}{2} \frac{1}{\sqrt{Var\{Y(z_1 - z_2)\}}} \log \frac{z_2}{z_1} \right) - \frac{1}{z_2} \Phi \left( \frac{\sqrt{Var\{Y(z_1 - z_2)\}}}{2} + \frac{1}{\sqrt{Var\{Y(z_1 - z_2)\}}} \log \frac{z_1}{z_2} \right) \right\} \quad (2.63)$$

Aunque el proceso  $Y$  no sea estacionario, el *proceso máx-estables* es estacionario, ver (Kabluchko, Schlather, y De Haan, 2009).

## Modelo t-Extremal

Uno de los últimos modelos, el cual generaliza el *modelo de Schlather*, es conocido como el *modelo t-extremal* propuesto por Opitz (2012), ver (Opitz, 2012):

$$Y_i(x) = c_\nu \max\{0, \epsilon_i(x)\}^\nu; \quad c_\nu = \sqrt{\pi} 2^{-\frac{(\nu-2)}{2}} \Gamma \left( \frac{\nu+1}{2} \right)^{-1}; \quad \nu \geq 1 \quad (2.64)$$

Donde  $\epsilon_i$  es un *proceso gaussiano estándar* con función de correlación  $\rho$  y  $\Gamma$  la función Gamma. La función de distribución bivariada se define de la siguiente forma:

$$P(Z(s) \leq z_1, Z(t) \leq z_2) = \exp \left\{ -\frac{1}{z_1} T_{\nu+1} \left( -\frac{\rho(s-t)}{b} + \frac{1}{b} \left( \frac{z_2}{z_1} \right)^{\frac{1}{\nu}} \right) - \frac{1}{z_2} T_{\nu+1} \left( -\frac{\rho(s-t)}{b} + \frac{1}{b} \left( \frac{z_1}{z_2} \right)^{\frac{1}{\nu}} \right) \right\} \quad (2.65)$$

Donde  $T_\nu$  es la función de distribución acumulada de una variable aleatoria  $t$  de *Student* con  $\nu$  grados de libertad y  $b^2 = \{1 - \rho(s - t)^2/(\nu + 1)\}$ .

## Modelo Geométrico Gaussiano

El último modelo con el que se trabajó es el *modelo Geométrico-Gaussiano* definido por Davison y Gholamrezaee en 2012, ver (Davison y Gholamrezaee, 2012). Se trabaja con el proceso de Schlather y se tiene:

$$Y_i(x) = \exp \left\{ \sigma(x)\epsilon_i(x) - \frac{\sigma^2(x)}{2} \right\},$$

Donde  $\epsilon_i$  es un *proceso gaussiano estándar* y  $\sigma^2(x) = \text{Var}(\epsilon_i(x))$ . La función de distribución bivariada se define de la siguiente forma:

$$P(Z(s) \leq z_1, Z(t) \leq z_2) = \exp \left\{ -\frac{1}{z_1} \Phi \left( \frac{a}{2} + \frac{1}{a} \log \frac{z_2}{z_1} \right) - \frac{1}{z_2} \Phi \left( \frac{a}{2} + \frac{1}{a} \log \frac{z_1}{z_2} \right) \right\} \quad (2.66)$$

Donde  $a^2 = 2\sigma^2(1 - \rho(h))$ .

Se cuenta con una amplia gama de modelos para los *procesos máx-estables*. Se estudian todos los modelos mencionados y mediante criterios de selección de modelos se opta por trabajar con los mejores. Más adelante, se presentan los criterios utilizados para seleccionar los modelos que mejor se comportan (ajustan) para las precipitaciones máximas diarias extremas en todo el territorio uruguayo.

### 2.2.4. Selección del Modelo

En varias aplicaciones, por ejemplo, en *Estadística Espacial* donde se cuenta con modelos considerados complejos, se hace inviable el uso de la función de probabilidad. Existen diferentes métodos para la selección de modelos. Un primer proce-

dimiento, es mediante la adaptación de *criterios de información*, se cuenta con los siguientes, *TIC*, *Criterio de Información de Akaike (AIC)* y *Criterio de Información Bayesiano (BIC)*, ver (Varin y Vidoni, 2005). Por otro lado, la selección se puede realizar mediante pruebas del cociente de verosimilitud, ver (?, ?). Estas técnicas se explican con mayor profundidad en el Anexo, ver 5.2.2

### 2.2.5. Niveles de Retorno

Para lograr diferentes predicciones en este trabajo se utilizaron los llamados *niveles de retorno*, los mismos refieren a umbrales que se espera sean superados cada un período de tiempo, al que se llama *período de retorno*.

Este concepto se utiliza en diferentes áreas, por ejemplo en meteorología. Al momento de trabajar con el *Método de Bloques* en el contexto de la *Teoría de Valores Extremos*, partiendo de la ecuación generalizada de valores extremos (2.9), invirtiendo la misma, se llega a la siguiente expresión:

$$u_p = \begin{cases} \mu - \frac{\sigma}{\epsilon} [1 - -\log(1 - p)^{-\epsilon}], & \text{sí } \epsilon \neq 0 \\ \mu - \sigma [\log(-\log(1 - p))], & \text{sí } \epsilon = 0 \end{cases} \quad (2.67)$$

El valor  $u_p$  es el denominado *nivel de retorno*, con su  $p$  correspondiente, llamado *período de retorno*. Los demás parámetros son los mismos de la ecuación 2.9, correspondiente a forma, posición y escala.

De esta forma se espera que en promedio el umbral  $u_p$  se exceda una vez cada  $1/p$  períodos, en el presente caso, trimestres. Dicho de manera distinta,  $u_p$  sería excedido por el máximo trimestral con probabilidad  $p$ , para cualquier trimestre.

Por ejemplo, en el escenario de precipitaciones máximas, se habla de una precipitación cuyo período de retorno es igual a 25 años, se entiende que dicho evento será igualado una vez cada 25 años, en un transcurso de un gran número de años, por ejemplo, 1000 años.

## CAPÍTULO 2. METODOLOGÍA

---

# Capítulo 3

## Resultados Globales

A lo largo de este capítulo se presentan los diferentes resultados que se obtuvieron durante el trabajo realizado. Una vez llevado a cabo un análisis profundo en el Capítulo 1 sobre la base de datos y depurando los mismos se logró conservar 41 estaciones meteorológicas y pluviométricas repartidas por todo el territorio uruguayo y un total de 147 trimestres desde el año 1982 al año 2020. En primera instancia se presenta la estimación de los parámetros obtenidos aplicando la *Teoría de Valores Extremos* en las diferentes estaciones. Luego de realizar dicha estimación de parámetros y una estandarización marginal del proceso, se procedió a presentar los resultados para la teoría de *procesos máx-estables*. Luego, una vez estudiado los diferentes modelos paramétricos, se seleccionó mediante el criterio de información *TIC* los que mejor ajustaron y en particular aquel que tuvo mejor desempeño en las simulaciones. Finalmente, haciendo uso de estos modelos, se muestran diferentes simulaciones y otros resultados relevantes como la curva de retorno y los niveles de retornos.

### 3.1. Estimación de parámetros para las DGVE

Como se mencionó en el Capítulo 2, existen diferentes técnicas para realizar la estimación de los parámetros para las *DGVE*. Dos de las mismas son la *estimación*

por *Máxima Verosimilitud* y *Métodos de los Momentos Pesados*, si bien se pueden calcular de manera sencilla, no tienen en cuenta la posible dependencia espacial entre los parámetros. Por esto, asumiendo cierta dependencia entre las estaciones dos a dos se utiliza la técnica de *Estimación Conjunta* aplicando un *modelo de regresión lineal*, teniendo en cuenta las latitudes y longitudes de las estaciones.

### 3.1.1. Máxima Verosimilitud y Estimación Conjunta

A la hora de seleccionar entre las técnicas de *Estimación por Máxima Verosimilitud* ó *Métodos de los Momentos Pesados* se optó por utilizar la estimación por *Máxima Verosimilitud*, dado que se contó con las condiciones de regularidad necesarias para que las propiedades asintóticas se cumplan. En este caso, se cuenta con un total de 147 datos por estación, cantidad considerablemente grande de observaciones. Si la muestra hubiese sido más pequeña lo aconsejable habría sido utilizar el *Método de los Momentos Pesados*. En los Cuadros 5.1 y 5.2, ver anexo

cap:anexo, se pueden apreciar las estimaciones de los parámetros de posición, escala y forma por ambas técnicas. Se aprecia que la estimación del parámetro de forma  $\hat{\epsilon} > -0,5$  para las 41 estaciones, por lo tanto, los estimadores de los tres parámetros  $(\hat{\mu}, \hat{\sigma}, \hat{\epsilon})$  presentan las propiedades asintóticas usuales.

Otro método para lograr la estimación de los parámetros, es el método de *Estimación Conjunta*. El mismo presenta cierta ventaja a la hora de estimar los parámetros de cada estación, dado que los mismos se pueden estimar en función de ciertas variables regresoras que caracterizan las estaciones. Por lo tanto, utilizando este método se obtiene una estimación en cualquier punto del territorio uruguayo de forma directa, a diferencia de las dos técnicas antes mencionadas.

Al realizar las estimaciones conjuntas se trabajó con las siguientes dos variables regresoras, latitud y longitud y se construyeron distintos modelos, partiendo del completo, donde se utilizaron ambas variables como regresoras de cada parámetro y luego con todos los submodelos posibles. El modelo completo es el siguiente:

$$\begin{cases} \mu(lat, lon) = \alpha_0 + \alpha_1 lat + \alpha_2 lon \\ \sigma(lat, lon) = \beta_0 + \beta_1 lat + \beta_2 lon \\ \epsilon(lat, lon) = \gamma_0 + \gamma_1 lat + \gamma_2 lon \end{cases} \quad (3.1)$$

En el Cuadro 3.1 se presentan los modelos que presentaron un menor TIC. En el modelo 2, los parámetros de posición y escala se estimaron según latitud y longitud, mientras que el parámetro de forma se consideró constante. El modelo 3, el parámetro de posición se estima según latitud y longitud mientras que el parámetro de escala se estimó solo mediante latitud y el de forma se dejó constante. El modelo 4, se estima igual que el modelo 3, pero el parámetro de escala se estima según la variable regresora longitud. El modelo 5, el parámetro de posición se estima solo considerando la variable latitud, mientras que el de escala se consideran ambas variables y el parámetro de forma se deja constante. Por último, el modelo 6 se estima de igual forma que el modelo 5, pero el parámetro de posición se estimó teniendo en cuenta la variable longitud. El modelo que mejor logra una estimación conjunta se selecciona mediante el criterio de información *TIC*. Se estudió el Cuadro 3.1, donde se observó el valor del TIC para cada uno de los seis modelos, y en este caso, el que observó el menor TIC y por ende fue el mejor modelo resultó ser el 5. Se aclara, que a la hora de estudiar los diferentes modelos se consideraron todos los posibles y se presentan solo los seis mejores.

En el Cuadro 3.1 se presentan los modelos que corresponden a los seis valores más pequeños de *TIC*.

Modelos	TIC
Modelo completo	57544.63
Modelo 2	57538.19
Modelo 3	57543.32
Modelo 4	57592.22
Modelo 5	57535.12
Modelo 6	57698.81

Cuadro 3.1: Valor del TIC para los seis mejores modelos utilizando dicho criterio de selección.

En el Cuadro 3.1 se observó que el modelo con menor TIC es el modelo 5, es decir, el mejor utilizando este criterio de selección, como se mencionó con anterioridad, a continuación se muestra su expresión exacta:

$$\begin{cases} \hat{\mu}(x) = 150,675 + 2,973lat \\ \hat{\sigma}(x) = 45,4228 - 0,4521lon + 1,4421lat \\ \hat{\epsilon}(x) = 0,07554 \end{cases} \quad (3.2)$$

Partiendo de este modelo se lograron que las estimaciones de los parámetros de posición, escala y forma para las 41 estaciones meteorológicas y pluviométricas, las mismas se pueden observar en los Cuadros 5.3 y 5.4, ver

cap:anexo. Con el fin de determinar si continuar trabajando con las estimaciones de máxima verosimilitud o las estimaciones conjuntas, se construyeron algunas visualizaciones. Observando la Figura 3.1, donde se muestra la *estimación conjunta* respecto a la *estimación de máxima verosimilitud*, se afirma que la *estimación espacial* presenta un comportamiento correcto, ya que se observa una fuerte dependencia positiva entre las diferentes estimaciones para los parámetros de posición y escala. Por otro lado, estudiando el histograma de la estimación por máxima verosimilitud del parámetro de forma, se observó respecto a línea punteada roja (estimación

### 3.1. Estimación de parámetros para las DGVE

constante mediante estimación conjunta) una variabilidad también baja respecto al valor estimado. A su vez, se ve que en la mayor cantidad de estaciones, las *estimaciones por máxima verosimilitud* son superiores a cero, esto podría ser un indicio de que uno de los posibles escenarios es trabajar con la estimación conjunta de dicho parámetro ( $\hat{\epsilon} = 0,07554$ ), de esta forma se asume un parámetro de forma mayor a 0 para todas las estaciones y se utiliza una distribución *Fréchet* para cada una. Dicha correlación entre las estimaciones, también indica que estaciones meteorológicas y pluviométricas relativamente cercanas presenta un comportamiento (respecto a las precipitaciones extremas) bastante similar.

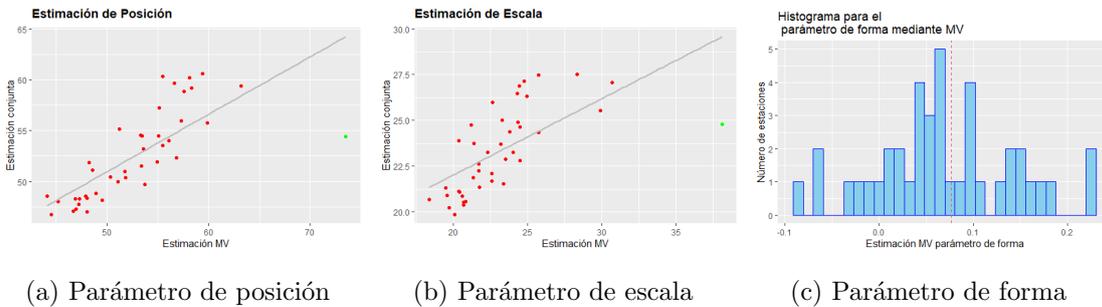


Figura 3.1: Estimación máxima verosímil respecto a estimación conjunta de los parámetros de posición, escala y forma del total de estaciones. Fuente: elaboración propia en R.

Por otro lado, en la Figura 3.1, existe un valor que no sigue las características generales del resto de estaciones, el punto de color verde, posee un comportamiento atípico y considerablemente diferente al resto de las estaciones meteorológicas y pluviométricas, más precisamente en la estimación de los parámetros de posición y escala. El mismo pertenece a la estación meteorológica *Piedras Coloradas*. Este comportamiento se puede deber a que en dicha estación se haya registrado un valor extremo de mayor magnitud en comparación a las demás estaciones. Estudiando la Figura 1.7, se puede observar que para el año 1997 se aprecia un registro atípico en comparación a los máximos anuales de las otras estaciones, el cual corresponde a una precipitación diaria de  $288mm$ , la cual pertenece a la estación *Piedras Coloradas*.

## CAPÍTULO 3. RESULTADOS GLOBALES

---

Si bien es necesario un tratamiento más riguroso, para que los resultados no se vean afectados, se explora la posibilidad de quitarla del análisis. Observando la Figura 3.2, se perciben estaciones meteorológicas y pluviométricas relativamente cercanas, por esta razón, se decidió quitar la estación *Piedras Coloradas*.

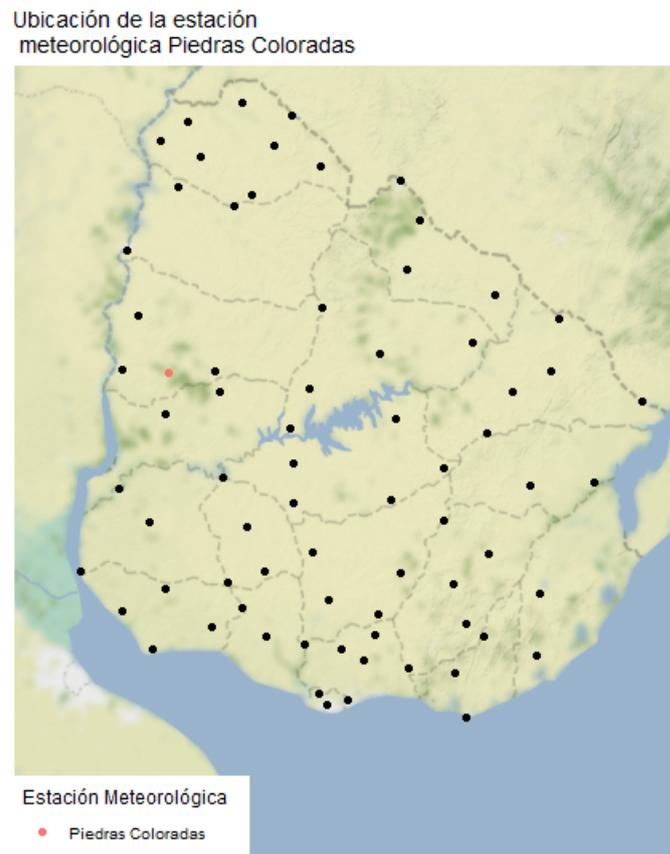


Figura 3.2: Ubicación de la estación meteorológica Piedras Coloradas y las restantes 40 estaciones pluviométricas o meteorológicas. Fuente: elaboración propia en R.

Una vez retirada del análisis se procedió nuevamente a estudiar todos los modelos posibles para aplicar la técnica *Estimación Conjunta*. Nuevamente el modelo que presentó un menor  $TIC$  (53153,3), fue el modelo 5 de la Tabla 3.1. Este modelo estimaba el parámetro de posición solo considerando la variable regresora latitud, mientras, que el parámetro de escala considera ambas variables, latitud y longitud, y por último, el parámetro de forma se deja constante,

$$\begin{cases} \hat{\mu}(x) = 145,437 + 2,822lat \\ \hat{\sigma}(x) = 48,5439 - 0,3381lon + 1,3503lat \\ \hat{\epsilon}(x) = 0,06977 \end{cases} \quad (3.3)$$

La estimación conjunta de los tres parámetros sin considerar la estación meteorológica *Piedras Colorada* se puede ver en los Cuadros 5.5 y 5.6, ver 5. Observando la Figura 3.3, más precisamente la del parámetro de posición y escala, es posible afirmar que al quitar del análisis la estación *Piedras Coloradas* se logró un ajuste considerablemente superior. Esta decisión de quitar dicha estación afirma aún más el hecho de estimar los parámetros mediante estimación espacial.

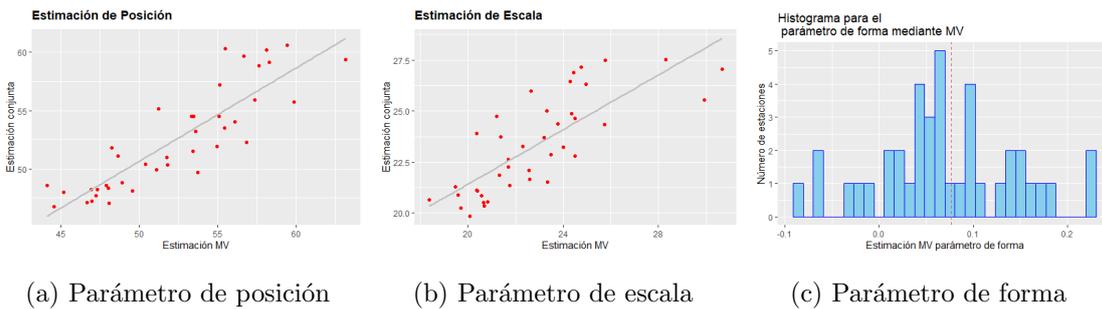


Figura 3.3: Estimación máxima verosímil respecto a estimación conjunta de los parámetros de posición, escala y forma en el total de estaciones sin la estación *Piedras Coloradas*. Fuente: elaboración propia en R.

Una decisión crucial en el modelado es la selección del parámetro de forma, puesto que determina el tipo de distribución a considerar y al ser la estimación cercana a 0 la pregunta es, ¿Es adecuado considerar que dicho parámetro  $\hat{\epsilon} = 0$ ? Este análisis de la estimación del parámetro de forma es de suma importancia a la hora de seleccionar una distribución de la *TVE*. Se recuerda lo siguiente, si dicho parámetro es menor a cero se debe trabajar con una distribución *Weibull*, mientras que si es mayor a cero, con una distribución *Fréchet* y si es igual a cero con una distribución *Gumbel*. Para realizar el estudio se decidió graficar el intervalo de confianza a un 95 % para cada

### CAPÍTULO 3. RESULTADOS GLOBALES

---

*estimación puntual máxima verosímil* del parámetro de forma para las 40 estaciones, observando si los intervalos incluían el cero y si incluían el valor estimado mediante la *estimación conjunta*,  $\hat{\epsilon} = 0,06977$ .

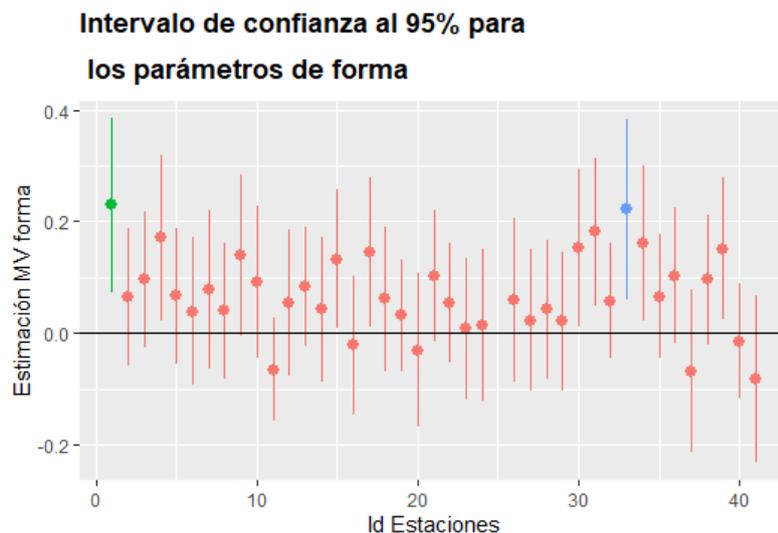


Figura 3.4: Intervalo de confianza para la estimación puntual máximo verosímil del parámetro de forma. Fuente: elaboración propia en R.

Estudiando la Figura 3.4 se observó que de un total de 40 estaciones, en 31 de las mismas los intervalos de confianza incluían al cero, mientras que en 38 estaciones los intervalos incluían el valor estimado mediante *estimación conjunta*  $\hat{\epsilon} = 0,06977$ . A su vez, todas las estaciones considerando el intervalo de confianza al 95 % incluyen valores positivos.

Entonces podemos considerar dos posibles escenarios. Por un lado, asumir que la estimación de dicho parámetro es cero,  $\hat{\epsilon} = 0$ . Por lo tanto, primero se trabajó con una distribución *Gumbel* y luego, se decidió explorar la posibilidad de considerar el parámetro mayor a cero  $\hat{\epsilon} > 0$ , entonces, se optó por la distribución *Fréchet*. Para lograr un mejor ajuste, observando nuevamente la Figura 3.4, la estimación en color verde y azul, son dos estaciones (25 de Agosto y San Antonio) las cuales presentaron la estimación puntual máximo verosímil más grande y a su vez, el intervalo

### 3.1. Estimación de parámetros para las DGVE

de confianza no incluía al cero y tampoco a la estimación obtenida por *estimación espacial*  $\hat{\epsilon} = 0,06977$ . Este comportamiento se consideró atípico y dado que esto podría perjudicar el resultado de la estimación espacial, mediante una decisión arbitraria, se valoró la posibilidad de quitarlas del análisis. Observando la Figura 3.5 se apreció que existían otras estaciones relativamente cerca a estas dos, por lo tanto, ambas estaciones se deciden sacar del estudio.

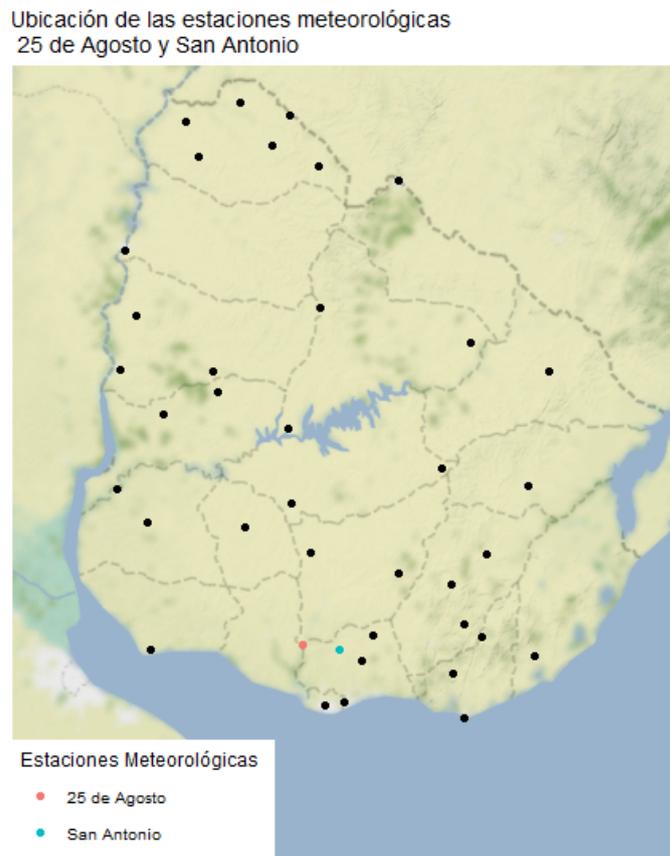


Figura 3.5: Ubicación de las estaciones meteorológicas de 25 de Agosto y San Antonio y las restantes 38 estaciones pluviométricas y meteorológicas. Fuente: elaboración propia en R.

En conclusión, estudiando las estimaciones de los tres parámetros mediante *estimación máxima verosimilitud* y *estimación conjunta*, se apreció que entre los máximos de las precipitaciones de cada estación hay cierta dependencia, entonces, se decidió trabajar con la *estimación conjunta*. También, mediante el estudio del in-

tervalo de confianza del parámetro de forma, se observó que la mayor cantidad de estaciones incluía el valor cero y el valor constante de la estimación conjunta. Por lo tanto primero se tomó para todas las estaciones el valor del parámetro de forma estimado igual a cero, ( $\hat{\epsilon} = 0$ ), y segundo, el valor estimado del parámetro de forma sea mayor a cero ( $\hat{\epsilon} = 0,06977$ ).

Una vez que se contó con las 38 estaciones meteorológicas y pluviométricas con las que se trabajó hasta el final, se volvió a realizar la estimación conjunta. Se consideraron los mismos modelos y el que presentó menor *TIC* fue nuevamente el modelo 5 del Cuadro 3.1. La estimación de los tres parámetros se consigue a partir del siguiente modelo: 3.4.

$$\begin{cases} \hat{\mu}(x) = 144,957 + 2,809lat \\ \hat{\sigma}(x) = 11,7526 - 0,8404ln + 1,0897lat \\ \hat{\epsilon}(x) = 0,06608 \end{cases} \quad (3.4)$$

En los Cuadros 5.7 y 5.8 se pueden apreciar los valores estimados de los tres parámetros mediante estimación espacial para las estaciones seleccionadas en el período 1982-2020. Estas estaciones son las que se utilizan a la hora de aplicar los *procesos máx-estables*. Se aclara, que luego de volver a aplicar la estimación conjunta, se volvió a realizar un estudio profundo sobre el parámetro de forma, y en este escenario no se presentó ningún comportamiento atípico.

Como se mencionó anteriormente, se plantearon dos escenarios posibles, primero considerar  $\hat{\epsilon} = 0$  y segundo  $\hat{\epsilon} = 0,06608$ . En el primer caso, se trabaja con una distribución de la *TVE* del tipo *Gumbel*. Se realizó la estimación espacial mediante la técnica *procesos máx-estables* y se estudió el comportamiento del *coeficiente extremal*. Luego se estudiaron todos los modelos paramétricos de los *procesos máx-estables* y mediante el *TIC* se seleccionó el más eficiente. Por último, se realizaron simulaciones y se compararon con el promedio de las precipitaciones máximas observadas

en el período de 1982-2020 de las 38 estaciones meteorológicas y pluviométricas de Uruguay. En el Anexo B, ver 5.2, se puede observar con más detalles estos resultados. Y mediante los mismos, se concluyó que considerar que el parámetro de forma tome el valor cero ( $\hat{\epsilon} = 0$ ), podría estar influyendo en el desempeño predictivo del modelo. Por esto, es que se decidió trabajar en el segundo escenario, el cual era considerar el parámetro de forma mayor a cero ( $\hat{\epsilon} > 0$ ), obtenido mediante la *estimación conjunta*  $\hat{\epsilon} = 0,06977$ . A continuación, se presentan los resultados en este contexto. Cabe aclarar que la distribución de *Fréchet* tiene cola más pesadas que la Gumbel, asumimos así un escenario menos conservador lo cual es en general deseable en este tipo de problemas.

## 3.2. Estimación espacial

### 3.2.1. Transformación a Fréchet Unitaria

En segunda instancia, se consideró trabajar con la estimación del parámetro de forma que se obtuvo por estimación conjunta,  $\hat{\epsilon} = 0,06608$ , con 38 estaciones. Se recuerda que a cualquiera de las tres distribuciones de la *TVE* (*Weibull*, *Gumbel*, *Fréchet*), se le puede aplicar una transformación y llevarla a una *Fréchet Unitaria*, ver la ecuación 2.10. Esta transformación, se le aplica a cada uno de los datos observados de las precipitaciones máximas en cada uno de los 147 trimestres en las 38 estaciones. Como  $\hat{\epsilon} > 0$ , se debe trabajar con una distribución de la *TVE Fréchet*. En la Tabla 5.10, se puede observar la transformación a *Fréchet Unitaria* para los primeros y últimos cinco trimestres de las 38 estaciones meteorológicas y pluviométricas.

### 3.2.2. Estimación espacial mediante procesos máx-estables

Luego de haber considerado  $\hat{\epsilon} = 0$  y una *DVE Gumbel* se notó que los modelos paramétricos no presentaron un comportamiento similar al promedio de precipita-

## CAPÍTULO 3. RESULTADOS GLOBALES

---

ciones extremas diarias observadas, esto se puede apreciar en Anexo 5, 5.2.3.1. Por lo tanto, se decidió cambiar de escenario y trabajar con la estimación del parámetro de forma obtenido mediante la *estimación conjunta*,  $\hat{\epsilon} = 0,06608$ . Una vez aplicada la *Teoría de Valores Extremos*, se procedió a realizar una estimación espacial mediante la técnica *procesos máx-estables*. En primer lugar, se analizó el comportamiento del *coeficiente extremal*, el cual permitió estudiar la correlaciones dos a dos de las diferentes estaciones. Este estudio del *coeficiente extremal* se realizó para el modelo *t-extremal* con correlaciones *Bessel* y *Power Exponencial*, dado que ambos modelos (como se observa en el Cuadro 3.2) fueron los que presentaron un menor *TIC*. Observando la Figura 3.6, en primer lugar, no se ve mucha variabilidad, lo que da indicio de que las precipitaciones en estaciones cercanas tienen un comportamiento similar, pero mientras las distancias entre estaciones va en aumento, el comportamiento de las precipitaciones no es semejante. La línea violeta representa la estimación no paramétrica del *coeficiente extremal*, la línea azul se obtiene mediante la *función del coeficiente extremal* del modelo *t-extremal* con correlación *Bessel*, mientras que la línea roja, es la *función del coeficiente extremal* del modelo *t-extremal* con correlación *Power Exponencial*. Observando tanto la línea roja como la azul se pudo afirmar que ambos modelos tienen un ajuste considerablemente correcto. También analizando la estimación del Madograma y *F-Madograma* para las 38 estaciones meteorológicas y pluviométricas en la Figura 3.7 se apreció un comportamiento similar al que brindó el *coeficiente extremal*, estaciones cercanas tienen un comportamiento semejante, mientras que a mayor distancia entre estaciones el comportamiento de las mismas se asemeja más a una situación de independencia. En función del *coeficiente extremal*, no es muy claro el modelo *máx-estable* a utilizar, pero observando la Figura 3.6, se puede observar que el modelo *t-extremal* con correlaciones *Bessel* o *Power Exponencial* puede lograr un comportamiento bastante correcto. Por otro lado, sabemos que aquellos modelos que presentan una función del coeficiente extremal acotada por una constante pequeña, por ejemplo 2, no se pueden utilizar, como es el caso del modelo de *Shalther*, ya que los datos en algunos casos superan este valor, como

se observa en la figura 3.6. Otro ejemplo similar, es lo que sucede con el modelo de *Smith*, donde la matriz de varianza y covarianza estimada del mismo es:

$$\Sigma = 0,00001 \begin{pmatrix} 2,225 & -1,907 & 1,627 \\ -1,907 & 1,723 & -1,561 \\ 1,627 & -1,561 & 1,523 \end{pmatrix} \quad (3.5)$$

El determinante de dicha matriz es cercano a cero, lo que produce problemas de invertibilidad, por lo tanto, el modelo de *Smith* presenta problemas numéricos y también se descarta.

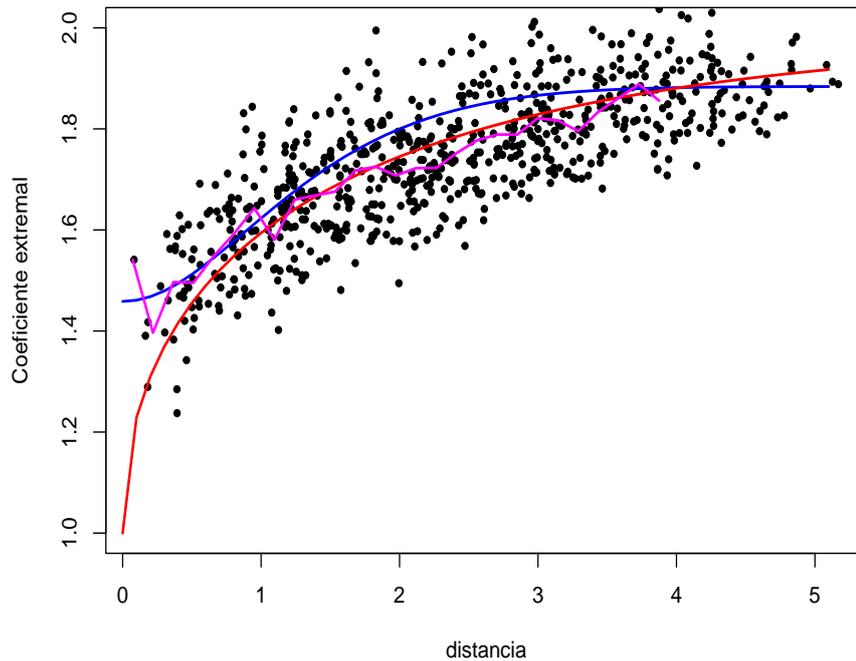


Figura 3.6: Estimación del coeficiente extremal del proceso t-extremal no paramétrica (violeta), paramétrica con Bessel (azul) y paramétrica con Power Exponencial (violeta), luego de transformar los datos a una Fréchet Unitaria del período 1982-2020 de las 38 estaciones, considerando una *DVE* Fréchet. Fuente: elaboración propia en R.

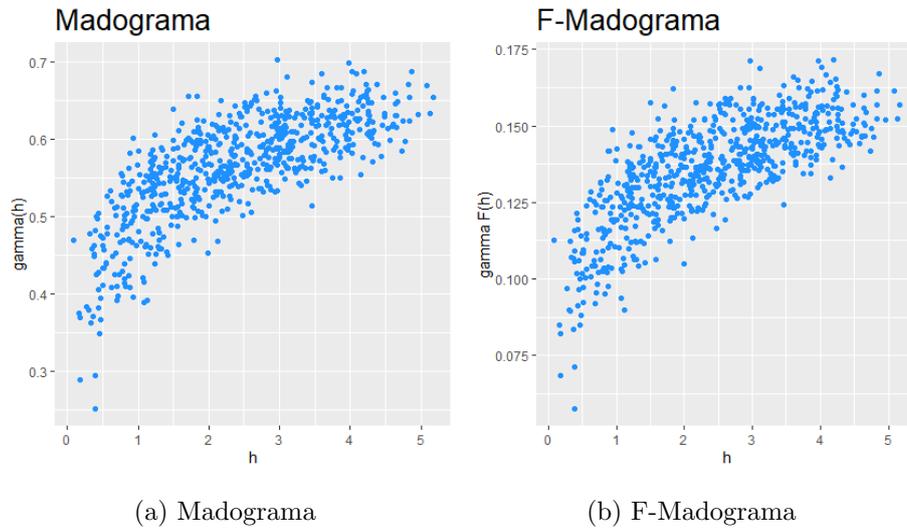


Figura 3.7: Estimación del Madograma y F-Madograma para las 38 estaciones meteorológicas ó pluviométricas. Fuente: elaboración propia en R.

### 3.2.3. Selección de Modelos

A la hora de estudiar los distintos modelos paramétricos se consideraron todos los posibles, con sus diferentes correlaciones, y el proceso de selección del modelo más adecuado se realizó utilizando el criterio de información  $TIC$ . En el Cuadro 3.2, se observa que el modelo con menor  $TIC$  es el  $t$ -*extremal*, con correlaciones *Whittle-Mattern*, *Power Exponencial* y *Bessel* al igual que cuando se trabajó con una *DVE Gumbel*. A su vez, se destaca que estos tres modelos con menor  $TIC$  son todos  $t$ -*extremal*.

Modelo	TIC
Geo-Gauss W-M	874256.60
Geo-Gauss Pow-Exp	874245.60
t-extremal W-M	874217.60
t-extremal Pow-Exp	874213.90
t-extremal Bessel	874219.70

Cuadro 3.2: Valor del TIC para los cinco modelos con menor TIC en estimación espacial mediante procesos máx-estables para el período 1982-2020 en 38 estaciones meteorológicas y pluviométricas considerando una *DVE Fréchet*.

En el Cuadro 3.3 se presenta la estimaciones de los diferentes parámetros del modelo *t-extremal* con correlaciones *Power Exponencial*, *Bessel* y *Whittle Mattern* para los *procesos máx-estables*, ver Figura 3.6.

Modelos	Parámetros Estimados			
	Nugget	Rango	Suavizado	Grados de Libertad
t-extremal W-M	0.007531	88.849695	0.452012	81
t-extremal Bessel	0.24978	0.09416	146.24461	4
t-extremal Pow-Exp	0.008749	74.233148	0.917324	71

Cuadro 3.3: Estimación de los parámetros para los mejores tres modelos de los procesos máx-estables.

### 3.3. Simulaciones

En el ámbito del estudio de simulación propuesto, se tomó una grilla de 30x30 equiespaciada, situada en las coordenadas  $-35,00^\circ 0'0''$  a  $-30,00^\circ 0'0''$  latitud y  $-58,50^\circ 0'0''$  a  $-53,00^\circ 0'0''$  longitud, donde se aseguró que la misma incluía todo el territorio uruguayo como se exhibe en la Figura 3.8.

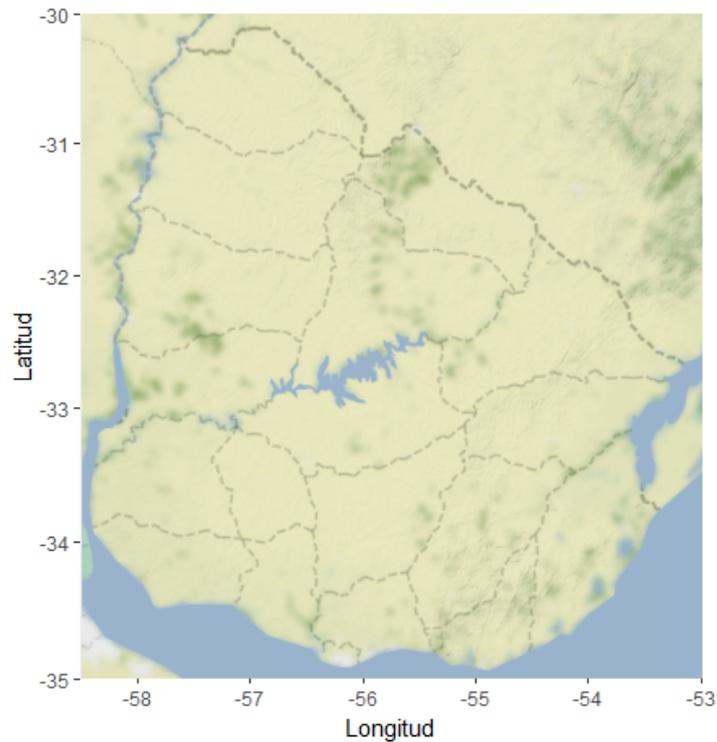


Figura 3.8: Región del territorio donde se realizaron las simulaciones. Fuente: elaboración propia en R.

Se simularon 147 matrices de 30x30, esta elección fue dada ya que luego de seleccionadas las 38 estaciones y realizada la depuración de datos correcta, en el período 1982-2020 se logró conservar 147 trimestres (bloques). Se procedió a escoger de cada una de las 147 matrices, el punto que más cercano estaba de la grilla equiespaciada a las coordenadas de cada una de las 38 estaciones, para esto se utilizó la distancia euclidiana considerando las medidas en longitud y latitud. Al ser dichos valores simulados *Fréchet Unitarios* se transformó los mismos, de manera que sean *Fréchet*.

De esta forma se obtuvo para cada estación meteorológica y pluviométrica 147 valores simulados (uno por cada bloque), para los tres modelos anteriormente mencionados. Para estudiar que tan bien ajustaban estos modelos, se decidió seleccionar de los datos observados, de cada estación, el promedio de las precipitaciones extremas registradas en los 147 trimestres para cada estación, y lo mismo, se realizó

con los valores simulados, de los 147 valores simulados por estación se calculó el promedio. Estos dos vectores se compararon, el vector observado y el simulado del promedio de los máximos de las precipitaciones en el período 1982-2020 (147 trimestres) en el territorio uruguayo. En el Anexo, ver 5, se presenta parte del código en R a la hora de realizar las simulaciones y las visualización. El mismo muestra una simulación mediante el modelo *t-extremal* con correlación *Bessel*, este mismo código con mínimas modificaciones es el utilizado para los otros dos modelos, *t-extremal* con correlaciones *Whittle-Matérn* y *Power Exponencial*.

En la Figura 3.9 se observan los resultados obtenidos con una de las tantas simulaciones realizadas para los tres mejores modelos. Donde el que mejor ajusta comparando el promedio de las precipitaciones máximas por trimestres observadas, contra el promedio de las precipitaciones máximas simuladas, es el modelo *t-extremal* con correlación *Bessel*. Los otros dos modelos, *t-extremal* con correlaciones *Whittle-Matérn* y *Power Exponencial*, lograron un ajuste en líneas generales correcto. Pero ambos modelos captaron valores de promedios de precipitaciones demasiado altos, sobrestiman en promedio valores grandes de precipitaciones diarias en comparación a lo observado. También el modelo *t-extremal* con correlación *Bessel* a la hora de simular sobrestima en promedio valores grandes de precipitaciones en comparación a lo observado pero en forma más leve. Por otro lado, las diferentes simulaciones con el modelo *t-extremal* con correlación *Whittle-Matérn* presentaban una variabilidad del promedio de las precipitaciones máximas por trimestres superior a los datos observados. Mientras que con el modelo *t-extremal* con correlación *Power Exponencial* sucedió lo contrario, la variabilidad del promedio de las precipitaciones máximas por trimestres era inferior con respecto a los datos observados. Este análisis se realizó mediante una cantidad considerablemente grande de simulaciones y, en todas las simulaciones, el modelo que mejor ajusto fue el *t-extremal* con correlación *Bessel*.

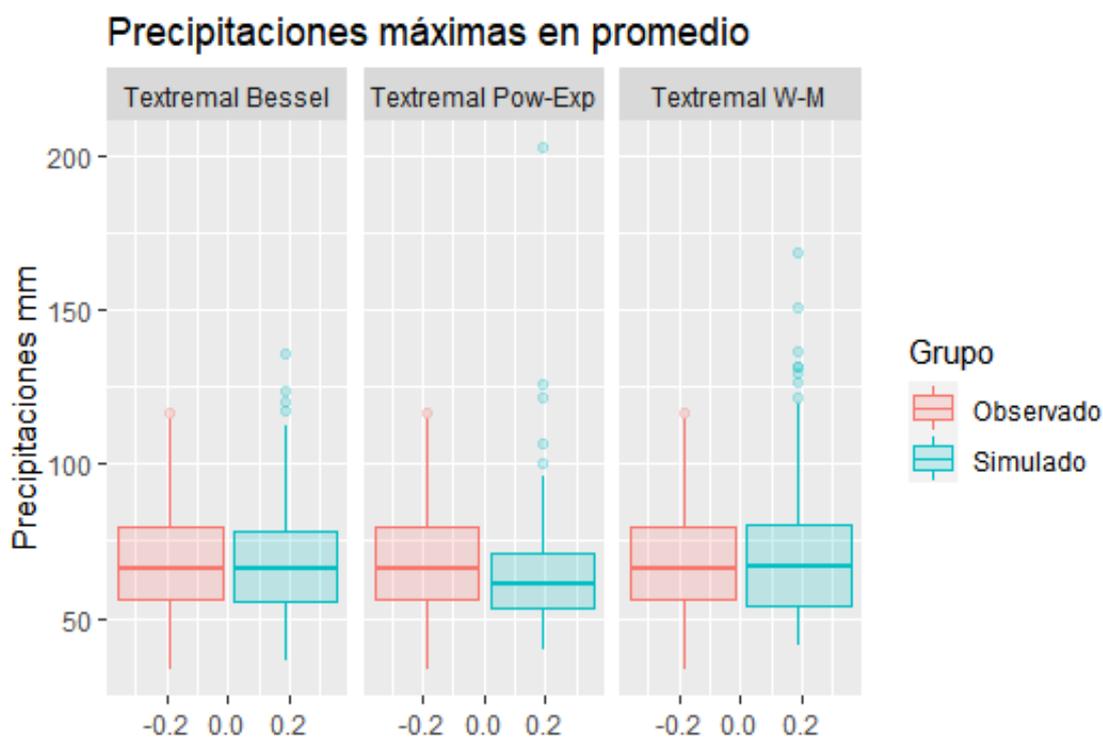


Figura 3.9: Comparación del promedio de precipitaciones máximas simulados mediante el modelo  $t$ -extremal con correlaciones *Bessel*, *Whittle-Matérn* y *Power Exponencial* contra el promedio de los máximos de las precipitaciones observadas. Fuente: elaboración propia en R.

En la Figura 3.10, se aprecian algunas de las diferentes simulaciones realizadas con el mejor modelo comparando el promedio de los máximos predichos y los observados en los 147 trimestres del período 1982-2020. Se observa que en todas las simulaciones el modelo sobrestima valores más altos del promedio de precipitaciones diarias en comparación al promedio de precipitaciones diarias extremas observadas, por ejemplo en la simulación 6 el modelo simuló un valor atípico de una precipitación diaria en un trimestre superior a  $200\text{mm}$ . Sin embargo en general las cajas concentran los valores de manera parecida, es decir, tienen valores similares del primer, segundo y tercer cuartil.

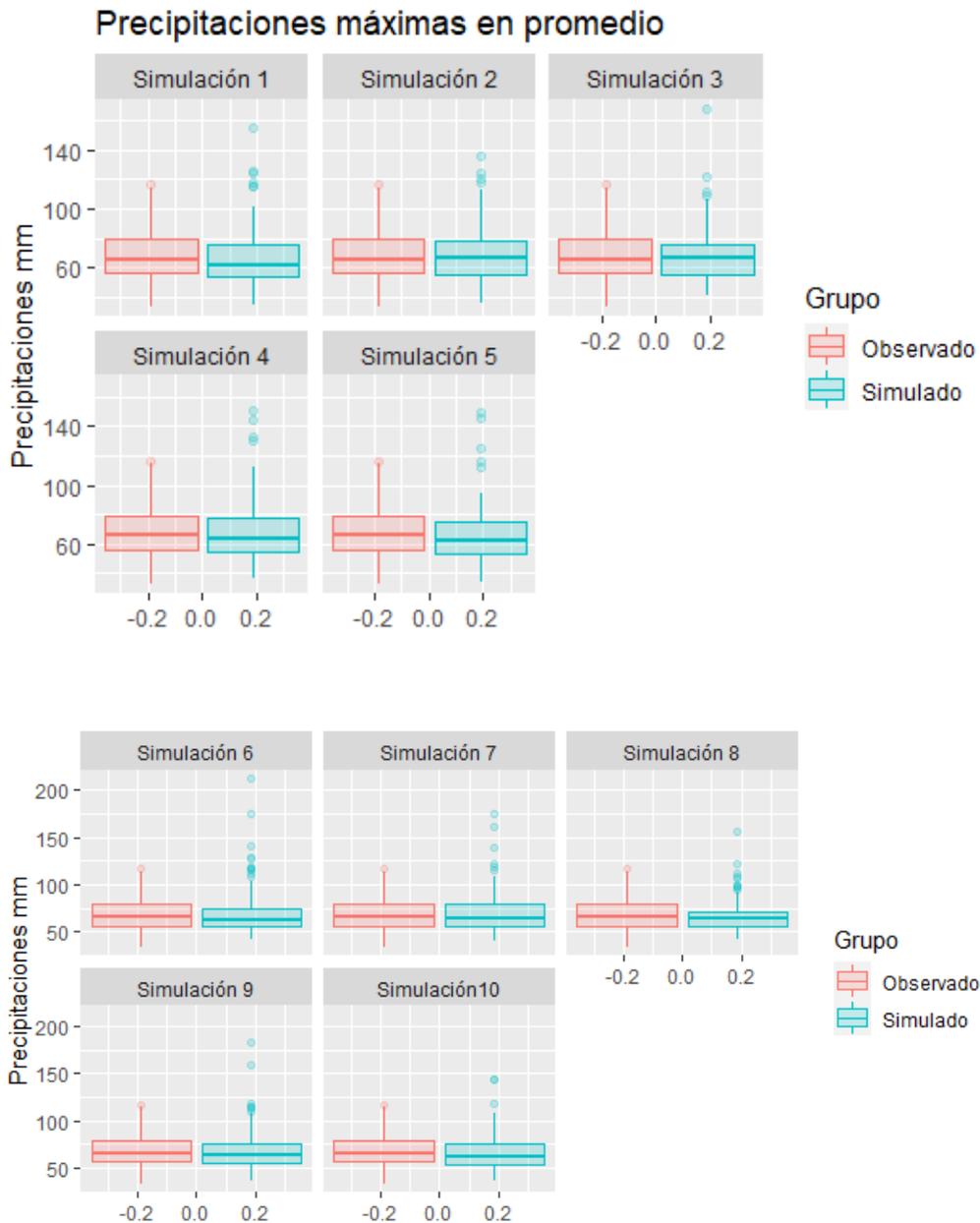


Figura 3.10: Comparación entre valores simulados del promedio de las precipitaciones máximas por trimestres considerando un modelo  $t$ -*extremal* con correlación *Bessel* y el promedio de las precipitaciones máximas observadas por trimestres. Fuente: elaboración propia en R.

En las Figuras 3.11 y 3.12 se observan gráficos similares a los anteriores, esta vez comparando cada estación por separado para el mejor modelo. Se destaca que

### CAPÍTULO 3. RESULTADOS GLOBALES

---

el modelo logró en líneas generales un ajuste adecuado en todas las estaciones meteorológicas y pluviométricas con las que se contaba. Para la estación de Barriga Negra se observó que el modelo subestimó levemente las precipitaciones máximas en el rango de lluvias entre el segundo y tercer cuartil, pero los valores más altos de lluvias si los predijo adecuadamente. En la mayoría de las estaciones el modelo captó los máximos más grandes de las precipitaciones, a su vez, la variabilidad que se aprecia en las simulaciones es semejante a los datos observados.

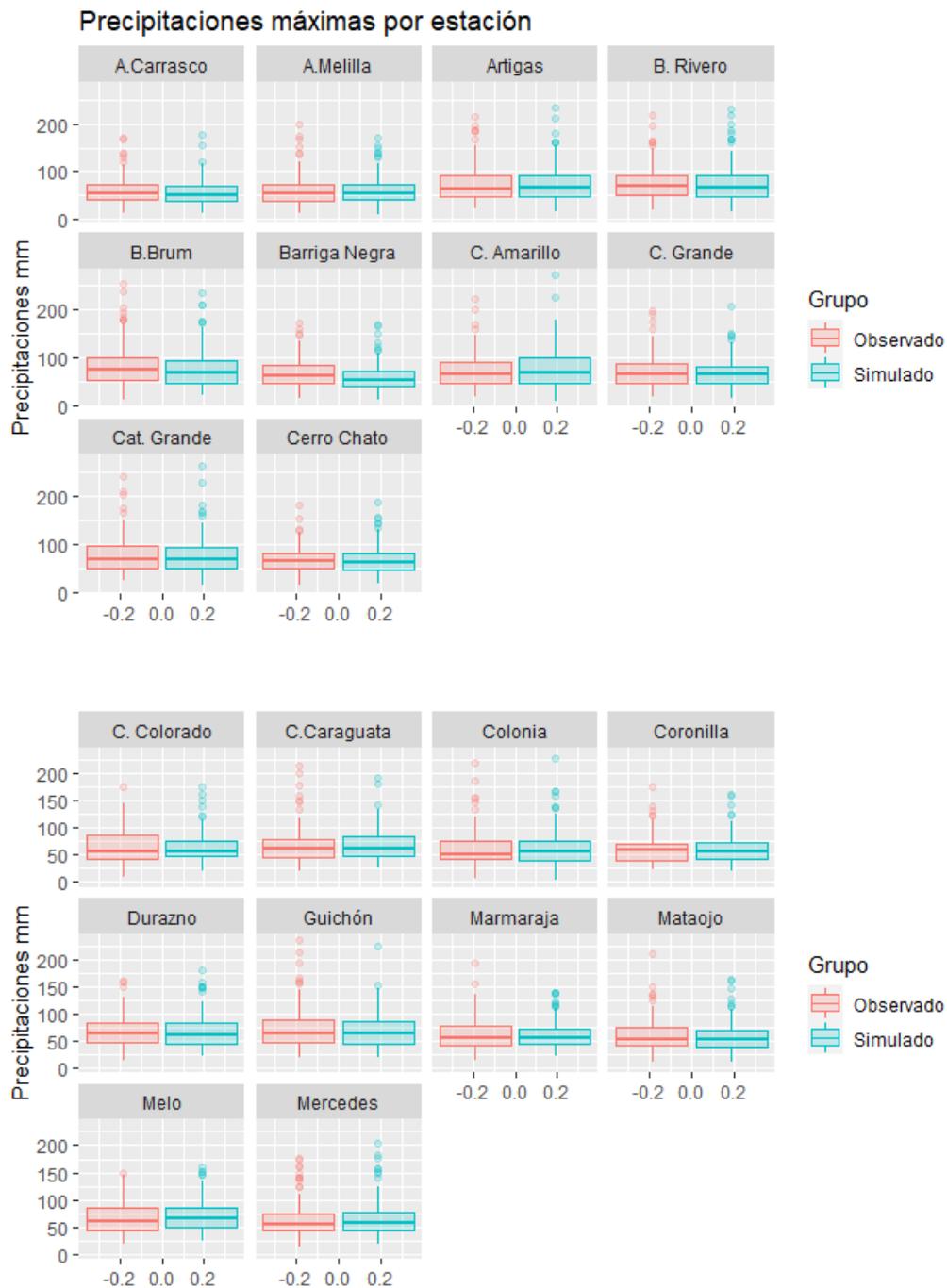


Figura 3.11: Comparación entre valores simulados del promedio de las precipitaciones máximas por trimestres considerando un modelo  $t$ -extremal con correlación *Bessel* y el promedio de las precipitaciones máximas observadas por trimestres estación por estación. Fuente: elaboración propia en R.

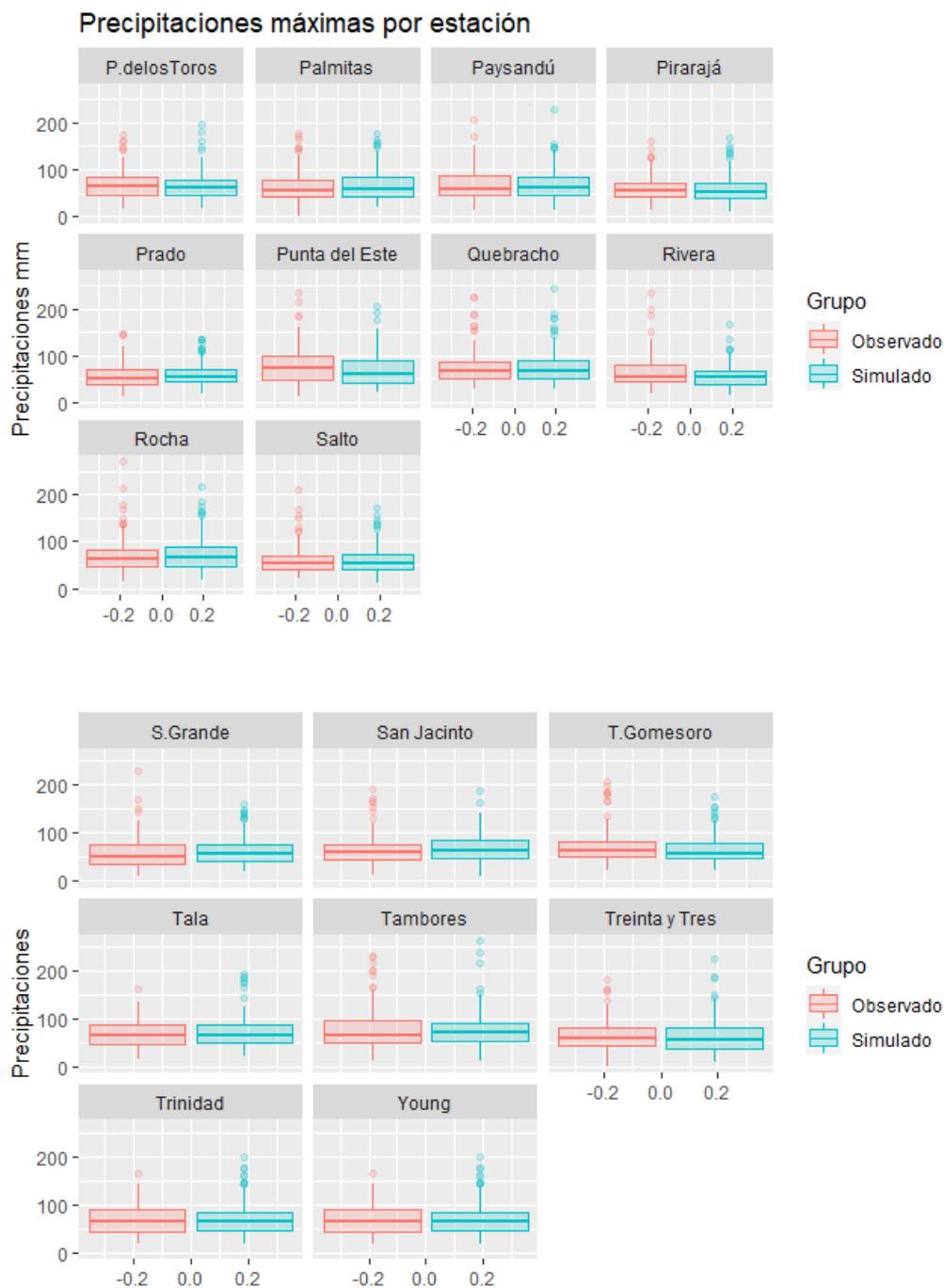


Figura 3.12: Comparación entre valores simulados del promedio de las precipitaciones máximas por trimestres considerando un modelo  $t$ -extremal con correlación *Bessel* y el promedio de las precipitaciones máximas observadas por trimestres estación por estación. Fuente: elaboración propia en R.

Luego se realizaron diferentes mapas de calor en todo el territorio uruguayo utilizando el modelo *t-extremal* con correlación *Bessel*. Se seleccionó la grilla de coordenadas de tamaño  $30 \times 30$ , de  $-35,00^\circ 0'0''$  a  $-30,00^\circ 0'0''$  en latitud y de  $-58,50^\circ 0'0''$  a  $-53,00^\circ 0'0''$  en longitud. Esta grilla incluía todo el territorio uruguayo. De los 900 valores simulados se filtraron solo aquellos que pertenecían a Uruguay. A continuación se exhiben algunos mapas de calor.

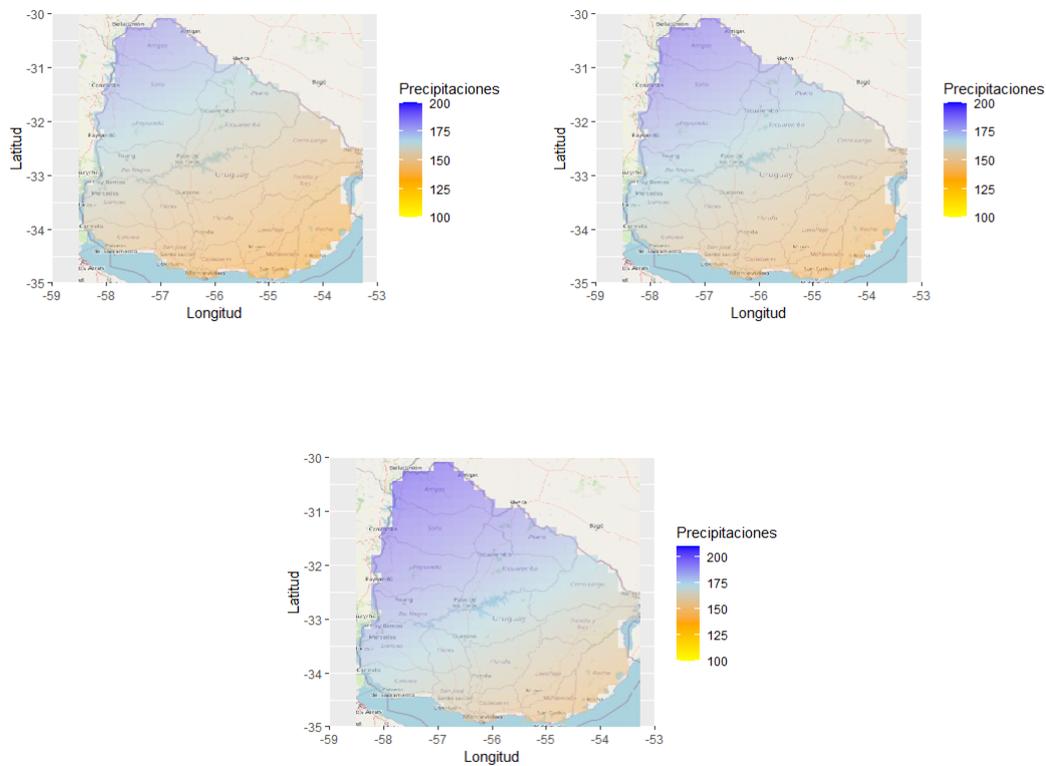


Figura 3.13: Mapas de calor de las precipitaciones máximas simuladas con un modelo *t-extremal* con correlación *Bessel* en Uruguay. Fuente: elaboración propia en R.

En los mapas de calor, Figura 3.13, se aprecia que las precipitaciones máximas diarias predichas tienden a crecer desde el sureste hacia el noroeste. Estos mapas de calor son basados en los máximos trimestrales. Se aclara que al utilizar el método de bloques, los máximos simulados no tienen porqué registrarse en el mismo día

### CAPÍTULO 3. RESULTADOS GLOBALES

---

del trimestre. En el norte del país es donde se registran las precipitaciones diarias más altas, además, este comportamiento presenta un correlato similar en las observaciones. La precipitación máxima registrada en el período 1982-2020, se captó en la estación meteorológica de Salto ubicada en el norte de Uruguay. Por otro lado, se recuerda que la técnica de estimación que se utilizó fue *estimación conjunta*, mediante el modelo 3.4. Como se puede ver, el coeficiente estimado  $\hat{\beta}_1$ , asociado a la variable regresora latitud para la estimación del parámetro de posición toma un valor positivo (2,809), esto nos indica que las precipitaciones máximas diarias serán mayores a medida que se tiende a desplazar al norte. Por otra parte, analizando la estimación del parámetro de escala, se ve que el coeficiente estimado  $\hat{\beta}_1$  asociado a la variable regresora latitud es negativo (-0,8404), lo que produce que cuanto más al norte de Uruguay, las precipitaciones máximas diarias tiendan a tener una dispersión mayor y el coeficiente estimado  $\hat{\beta}_2$  asociado a la variable regresora longitud es positivo (1,0897), entonces a medida que se desplaza hacia el sureste del territorio uruguayo, las precipitaciones máximas diarias tiende a tener una dispersión menor. Este patrón que se observa en los mapas de calor, en la Figura 3.13 es similar al presentado en el Capítulo 1, en la Figura 1.2, en la cual se apreciaba que las precipitaciones acumuladas en Uruguay tienden a crecer a medida que se desplaza hacia el norte del país.

Observando el patrón anteriormente mencionado se decidió efectuar un análisis de las precipitaciones máximas comparativo entre el sur y el norte de Uruguay. Tomando el modelo t-extremal con correlación *Bessel*. Se dividió el territorio uruguayo de manera de separar las simulaciones que se encontraban en el norte del Río Negro, y por otro lado, las ubicadas en el sur del Río Negro. En el norte del Río Negro se contó con un total de 17 estaciones meteorológicas y pluviométricas mientras que en el sur del Río Negro se computan 21 estaciones.

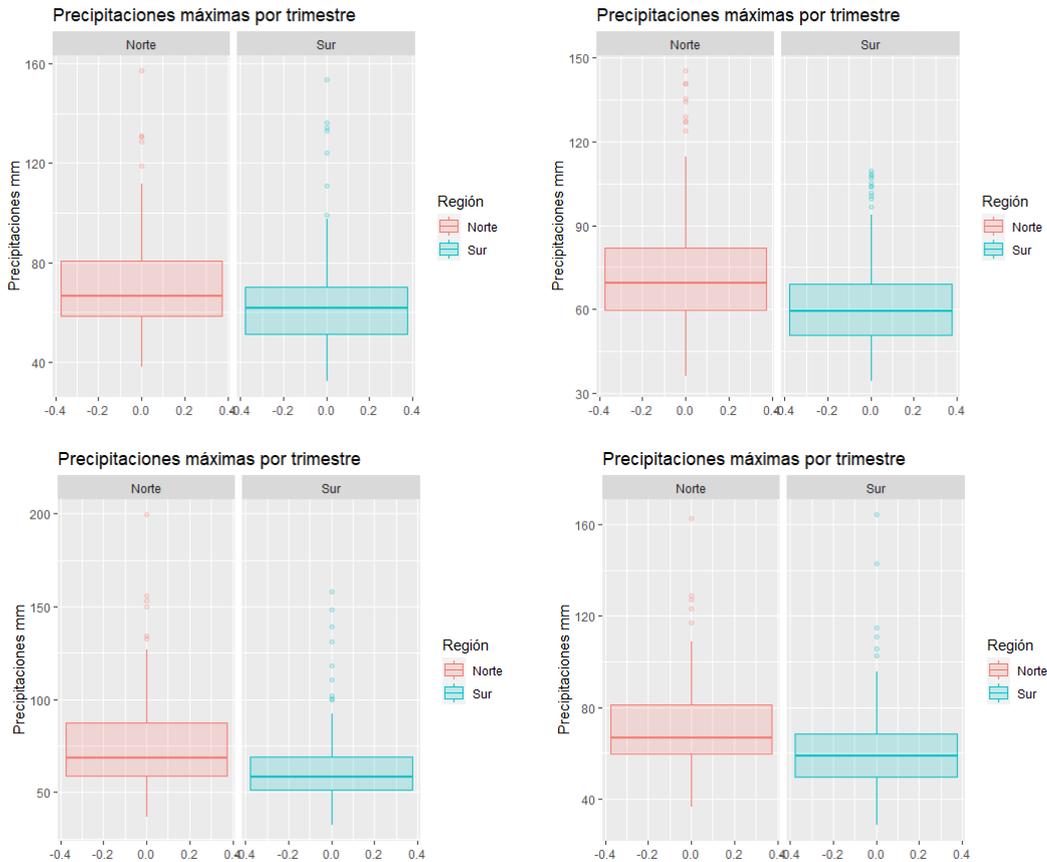


Figura 3.14: Comparación del promedio de las precipitaciones máximas simuladas entre el norte y sur de Uruguay. Fuente: elaboración propia en R.

En la Figura 3.14 se puede observar que en el norte de Uruguay las precipitaciones diarias máximas en promedio tienden a ser mayores que en el sur. También en el norte presentan una mayor variabilidad respecto al sur y, por otra parte, el promedio de las precipitaciones máximas es mayor en el norte que en el sur. Por último, tanto en el norte como el sur la distribución del promedio de precipitaciones máximas presenta un leve sesgo positivo. Esta comparación respalda a lo mostrado en los mapas de calor en la Figura 3.13.

Una vez realizadas diferentes simulaciones para el territorio uruguayo y un análisis comparativo entre el sur y el norte del país se decidió estudiar las probabilidades de que las precipitaciones extremas diarias superen un umbral en el próximo año y

## CAPÍTULO 3. RESULTADOS GLOBALES

---

en los próximos 5 y 10 años.

Se decidió aproximar estas probabilidades mediante simulaciones obtenidas con el modelo *t-extremal* con correlación *Bessel* el cual fue el que mejor se comportó, y también se presentan las probabilidades obtenidas con simulaciones de un modelo *t-extremal* con correlación *Power Exponencial*, el cual también había logrado un ajuste considerablemente correcto.

Para poder estimar las probabilidades de que las lluvias diarias extremas superen un umbral en los diferentes períodos de tiempos mencionados se procedió de la siguiente manera. Recordando que los bloques que se consideraron en el trabajo fueron bloques trimestrales. Si se realizan 4 simulaciones sería 1 año, entonces, se efectúan 40 simulaciones lo que equivale a 10 años, este procedimiento se repite mil veces. Cada 40 simulaciones sobre todo el territorio uruguayo se recoge el máximo, el cual es la precipitación máxima extrema diaria registrada en 10 años. Como el proceso se repitió mil veces, se obtuvieron 1000 máximos extremos diarias de precipitaciones cada 10 años. Una vez que se tienen estos mil valores se procede a calcular las probabilidades de que las precipitaciones máximas extremas diarias superen los umbrales de  $150mm$ ,  $200mm$ ,  $250mm$ . En el caso de conseguir las probabilidades para un período de 5 años el procedimiento es similar al del período de 10 años, con la excepción, de que en vez de realizar 40 simulaciones se realizan 20 simulaciones lo que equivale a 5 años y se consigue el máximo cada 20 simulaciones, y una vez más el proceso se repite mil veces. Por último, se realiza lo mismo cuando se consideró el período de 1 año, y en este caso se realizan 4 simulaciones, se consigue el máximo de las 4 simulaciones y también se repite el proceso mil veces.

En el Cuadro 3.4 se puede observar las estimaciones de las probabilidades realizadas mediante simulaciones de superar los umbrales que se mencionaron anteriormente,

	Modelos					
	t-extremal Bessel			t-extremal Pow-Exp		
Probabilidad	1 Año	5 Años	10 Años	1 Año	5 Años	10 Años
$P(\text{superar } 200\text{mm})$	0,456	0,949	0,998	0,641	0,996	1,000
$P(\text{superar } 250\text{mm})$	0,160	0,558	0,795	0,244	0,746	0,940
$P(\text{superar } 300\text{mm})$	0,051	0,243	0,388	0,081	0,346	0,564

Cuadro 3.4: Probabilidades de superar un umbral de precipitaciones diarias máximas para un período de 1 año, 5 años y 10 años con el modelo *t-extremal* con correlaciones *Bessel* y *Power Exponencial*

En la Tabla 3.4 se observa que la probabilidad de superar una precipitación de  $200\text{mm}$  en 1 año, 5 años y 10 años son, 0,456, 0,949 y 0,998. La probabilidad de superar  $250\text{mm}$  en 1 año, 5 años y 10 años son 0,160, 0,558 y 0,795, mientras que la probabilidad de superar  $300\text{mm}$  en los próximos 1 año, 5 años y 10 años son 0,051, 0,243 y 0,388 para el modelo *t-extremal* con correlación *Bessel*. Se apreció que a medida que el umbral es más grande las probabilidades de superar dicho umbral decrecen, y a su vez, mientras el período estudiado de superar el umbral va en aumento las probabilidades tienden a crecer. Se realizó un estudio con la distribución empírica, solo considerando un período de un año. Se recuerda que se contaba con un total de 147 trimestres en un período desde 1982 hasta 2020, donde la mayoría de los años contaba con los cuatro trimestres pero hubo casos en que algunos trimestres se tuvieron que quitar del análisis. En total se tenía un total de 38 años, en cada uno de los años se tomó el máximo de las precipitaciones diarias, por lo tanto, se tiene 38 máximos de precipitaciones diarias, de los mismo un 48,7% superó un umbral de  $200\text{mm}$ , muy similar a lo logrado con los datos simulados, mientras que un 0,051% supera el umbral de  $250\text{mm}$ , y el umbral de  $300\text{mm}$  no es superado por ninguna de las precipitaciones máximas anuales. Se recuerda que la precipitación máxima registrada en este período de 38 años se dio en el año 1997

donde se dio un acontecimiento de una precipitación diaria de  $269,7mm$ .

### 3.4. Niveles de Retornos estimados

A continuación, se presentan la curva de retorno y los niveles de retornos estimados de las precipitaciones extremas trimestrales mediante el modelo *t-extremal* con correlación *Bessel*. A la hora de estimar la curva y los niveles de retorno se optó por tomar una grilla de coordenadas que incluía todo el departamento de Montevideo, con el objetivo de estimar tanto la curva como los niveles de retornos en esta grilla. Dentro de la misma se contaba con las estaciones pluviométricas *Aeropuerto Carrasco* y *Aeropuerto Melilla*. En la Figura 3.15 se aprecia el espacio geográfico donde se estimaron dichos *niveles de retornos* junto con las dos estaciones mencionadas. Se recuerda que mediante los *procesos máx-estables* se logra modelar las precipitaciones extremas para todo los puntos de la grilla.

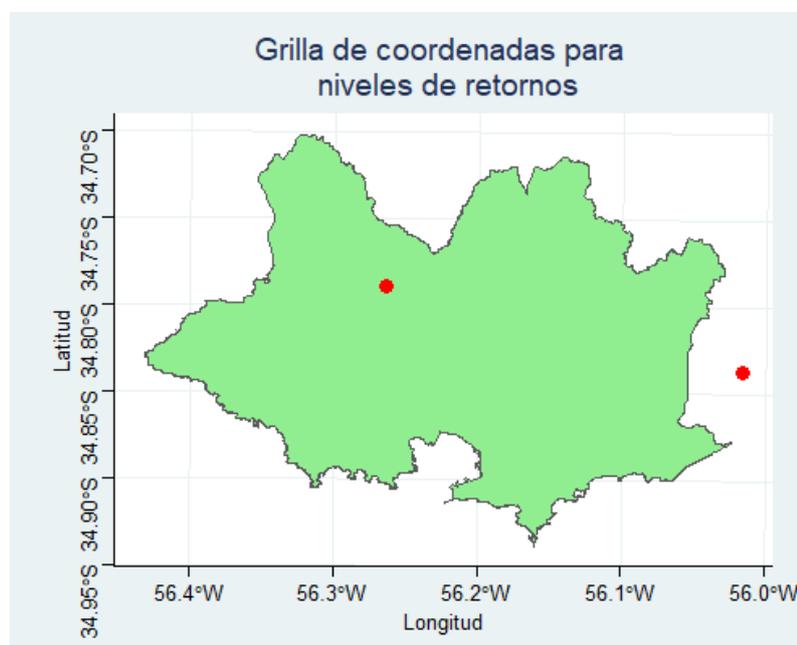


Figura 3.15: Grilla de coordenadas con la ubicación de las estaciones pluviométricas de *Aeropuerto Carrasco* y *Aeropuerto Melilla* donde se estimaron los diferentes *niveles de retornos*. Fuente: elaboración propia en R.

En la siguiente Figura se presenta la *curva de retorno* estimada mediante un modelo *t-extremal* con correlación *Bessel* para el espacio geográfico exhibido en la Figura 3.15. Esta curva se estimó para un período continuo desde 1 trimestre hasta 200 trimestres, donde este último equivale a estimar dicha curva para un período de 50 años. Se puede apreciar que en los primeros 20 trimestres el valor esperado de las precipitaciones extremas crece de forma abrupta, y luego, a medida que se va moviendo en el tiempo, el valor esperado de precipitaciones extremas tiende a crecer pero de forma más leve.

Por último, en la Figura 3.16 se exhiben los niveles de retornos estimados para las precipitaciones extremas en la grilla de coordenadas anteriormente mencionada. Estos *niveles de retorno esperados*, se estimaron para un período de 20, 40, 60, 80 y 200 trimestres, lo que es equivalente a 5, 10, 15, 20 y 50 años. Estos niveles, como se mencionó en 2.2.5 son los percentiles 0, 95 %, 0, 975 %, 0, 983 %, 0, 9875 % y 0,995 %, del modelo *t-extremal* con correlación *Bessel*.

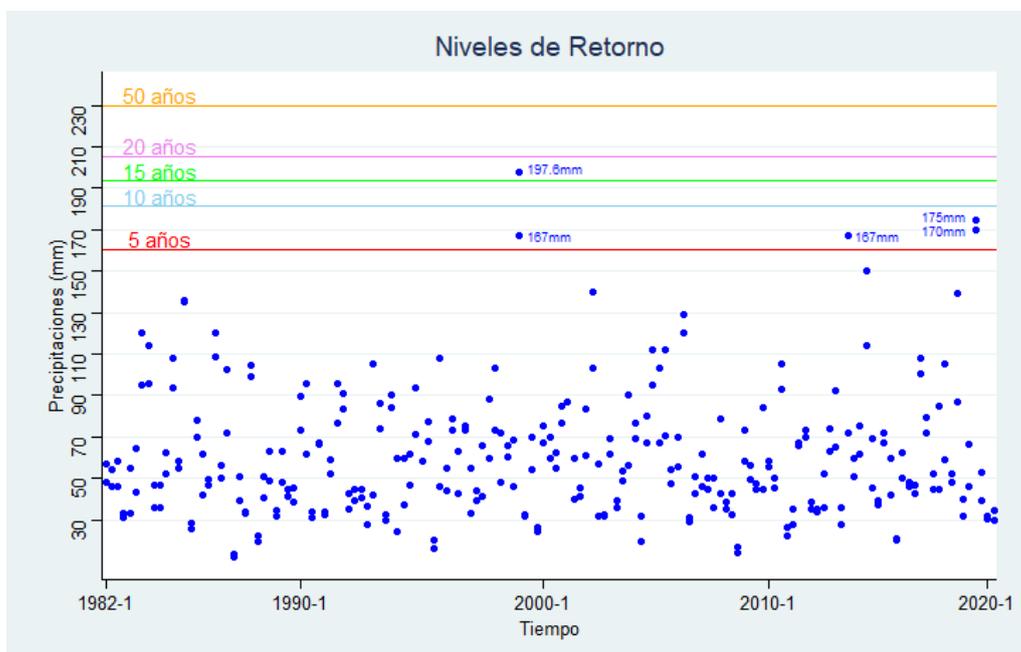


Figura 3.16: Niveles de retornos estimados para las precipitaciones extremas para los períodos de tiempo de 5, 10, 15, 20 y 50 años. Fuente: elaboración propia en R.

### CAPÍTULO 3. RESULTADOS GLOBALES

---

Observando la Figura 3.16 se llega a que los *niveles de retornos* estimados para los tiempos de 5, 10, 15, 20 y 50 años son 160,23mm, 181,74mm, 193,44mm, 205,25mm y 230mm. Estos valores de lluvias extremas diarias se esperan que sucedan cada 5, 10, 15, 20 y 50 años. El valor máximo registrado de las precipitaciones extremas diarias en las dos estaciones pluviométricas que se encontraba en la grilla utilizada en el período de 1982-2020, se dio en el trimestre 1 del año 1999 en la estación *Aeropuerto Melilla* y se computó una precipitación diaria de 197mm. Este valor se encontró por encima del *nivel de retorno* correspondiente a 15 años, por lo tanto, esta lluvia extrema ocurrida en el año 1999 se espera que ocurran dentro de la grilla de coordenada cada 15 años. Mientras que en ese mismo año y en el mismo trimestre pero en estación *Aeropuerto Carrasco* se registró una precipitación de 167mm, este registro también ocurrió en el trimestre 2 del año 2017, y otros registros como el que se dio en el trimestre 4 del año 2018, donde en la estación *Aeropuerto Carrasco* se registró una precipitación de 170mm, mientras que en la estación *Aeropuerto Melilla* se registró un valor de 175mm. Estos últimos cuatro registros se encontraron por encima del *nivel de retorno* de 5 años, por lo tanto, precipitaciones de esta magnitud para la grilla de coordenadas utilizada para calcular dichos niveles se esperan ocurran cada un período de 5 años.

# Capítulo 4

## Conclusiones

A lo largo del trabajo se estudiaron diversas herramientas estadísticas y matemáticas, comenzando por la *Teoría de Valores Extremos*, una herramienta que permite la identificación y modelización de los máximos de una variable aleatoria. Se exploraron diferentes alternativas para la estimación de los parámetros, siendo la más importante la *estimación conjunta*, la cual asume cierta relación entre los parámetros de una estación y otra. También se estudiaron diversas técnicas que permiten estudiar la dependencia de valores extremos en un territorio, en este caso las lluvias diarias extremas en Uruguay. Por último, se aplicaron las diferentes técnicas que brindan los *procesos máx-estables* que parte de la *Teoría de Valores Extremos* y permite la modelización del campo de lluvias extremas en todo el territorio uruguayo.

Mediante las diferentes técnicas presentadas se logró cumplir el objetivo principal, modelar las precipitaciones diarias extremas en el territorio uruguayo, y se consiguieron resultados muy reveladores. Se observó que las precipitaciones máximas en Uruguay en líneas generales son homogéneas, y no son escasas, ya que cada cierto tiempo se registra algún hecho extremo. Por lo tanto, estas herramientas para un país como Uruguay, en el cual, gran parte de su economía se basa en el sector agricultor y turístico son de suma importancia. También se logró observar y detectar el comportamiento de las precipitaciones en forma de gradiente desde el Sureste

## CAPÍTULO 4. CONCLUSIONES

---

al Norte del territorio uruguayo, donde este comportamiento tanto en registro de precipitaciones diarias extremas como en registros de precipitaciones acumuladas anuales tiende a crecer desde el Sur al Norte del país y el modelo que mejor ajustó logró captar dicho comportamiento.

El modelo que ajustó de forma más adecuada a las precipitaciones máximas diarias extremas, fue el modelo *t-extremal* con correlación *Bessel*. Incluso los tres modelos que mejor TIC reportaron fueron todos *t-extremal*, por lo tanto, es una modelización que podría representar de buena forma las precipitaciones máximas en Uruguay.

Al analizar el parámetro de forma  $\epsilon$  se observó que se podría considerar que  $\hat{\epsilon} = 0$ , por lo tanto, una distribución marginal para cada estación del tipo *Gumbel* era factible, sin embargo, la mejor opción fue considerar que  $\hat{\epsilon} > 0$ , en particular se enfocó por la estimación arrojada por la *estimación conjunta*. Eligiendo la estimación del parámetro de forma por lo tanto utilizando distribuciones marginales *Frechet* se logró un ajuste mucho más preciso del modelo *t-extremal* con correlación *Bessel* a las observaciones de las precipitaciones máximas en el territorio uruguayo. Esto se debe a que la distribución *Frechet* posee colas más pesadas que la *Gumbel*, y se observó que esta última subestimaba lo observado.

Mediante otras técnicas estudiadas, partiendo de la estimación conjunta de los parámetros y luego aplicando herramientas como el *Madograma* y el *F-Madograma* se concluyó que las precipitaciones diarias extremas presentan cierta dependencia entre estaciones, la cual decrece con la distancia que las separa, y en términos generales son bastante homogéneas, ya que no se observaron cambios bruscos, sin embargo sí se observó un crecimiento de las lluvias máximas hacia el norte, es decir un aumento en el sentido de la latitud. Por esto se estudió el comportamiento de estos eventos comparando Norte y Sur del país y se logró apreciar una cierta diferencia entre estas zonas del país. Donde en el norte se logró predecir lluvias diarias extremas superiores al sur de manera clara.

---

Por lo tanto, un futuro trabajo puede ser aplicar estas técnicas por un lado para las estaciones que se ubican en el sur del Río Negro y, por otro lado, para las estaciones que se encuentran al norte del Río Negro.

Otro futuro trabajo podría ser un análisis considerando dos etapas del año, por un lado, considerando los trimestres de las estaciones otoño e invierno, y por otro lado, los trimestres de las estaciones primavera y verano. También se podría explorar la aplicación de otras técnicas, como las “Cópula extremas”.

## CAPÍTULO 4. CONCLUSIONES

---

# Capítulo 5

## Anexo

### 5.1. Anexo A - Código en el software R

**Función para calcular porcentajes de NA en las estaciones por trimestres**

```
FuncionNAs <- function(estacion=estacion) {  
cont = 1 trimestreNAs = matrix(rep(0,27456),nrow = 6864, ncol = 4)  
for (e in 1:44) {  
  for (a in 1:39) {  
    for (q in 1:4) {  
      valor <- sum(is.na(dplyr::filter(datos,  
Estación == estacion [e] & year(Fecha)==1981+a,  
quarter(Fecha)==q)))/max(row(dplyr::filter(datos, Estación==estacion[e] &  
year(Fecha)==1981+a, quarter(Fecha)==q)))  
trimestreNAs[cont,4]<-round(valor,3)  
trimestreNAs[cont,3]<-q  
trimestreNAs[cont,2]<-1981+a  
trimestreNAs[cont,1]<- estacion[e]  
cont<- cont+1 }  
  } }  
}
```

```
return(trimestreNAs) }
```

**Función para calcular los máximos de los 147 trimestres de las 41 estaciones meteorológicas y pluviométricas** `maximo <- function(estacion=estacion)`

```
maximo <- matrix(rep(0, 6396), nrow = 41, ncol = 156)
for (e in 1:41) {
cont<-1
for (a in 1:39) {
for (q in 1:4) {
valor <- filter(datos.corregidos, Estación == estacion_1 [e] &
year(Fecha)==1981+a & quarter(Fecha)==q) v_max <- max(valor [,3])
maximo[e,cont]<- v_max cont <- cont+1 }
} }
return(maximo)
```

**Estimación Conjunta de las 38 estaciones meteorológicas ó pluviométricas**

```
tmaximo_1 <- t(maximo_1) trasponemos la matrix de maximos
form.loc_1 <- as.formula(loc ~ lon + lat)
form.scale <- as.formula(scale ~ lon + lat)
form.shape_1 <- as.formula(shape ~ lon + lat)
covariables_1.1 <- as.numeric(coordenadas.corregidos_2[, 1])
covariables_2.1 <- as.numeric(coordenadas.corregidos_2[, 2])
covariables_1 <- cbind(covariables_1.1, covariables_2.1)
colnames(covariables_1) <- c("lat", "lon")
form.shape_2.1 <- as.formula(shape ~ 1)
form.scale_3.1 <- as.formula(scale ~ lat)
form.scale_4.1 <- as.formula(scale ~ lon)
form.loc_2.1 <- as.formula(loc ~ lat)
```

```

form.loc_3.1 <- as.formula(loc ~ lon)
a_1.1j
-
fitspatgev(data=as.matrix(tmaximo_1),
  covariables = covariables_1, form.loc_1, form.scale_1, form.shape_1)
a_2.1 <- fitspatgev(data=as.matrix(tmaximo_1),covariables = covariables_1,
  form.loc_1, form.scale_1, form.shape_2.1)
a_3.1 <- fitspatgev(data=as.matrix(tmaximo_1),covariables = covariables_1,
form.loc_1, form.scale_3.1, form.shape_2.1)
a_4.1 <- fitspatgev(data=as.matrix(tmaximo_1),covariables = covariables_1,
form.loc_1, form.scale_4.1, form.shape_2.1)
a_5.1 <- fitspatgev(data=as.matrix(tmaximo_1),covariables = covariables_1,
  form.loc_2.1, form.scale_1, form.shape_2.1)
a_6.1 <- fitspatgev(data=as.matrix(tmaximo_1),covariables = covariables_1,
  form.loc_3.1, form.scale_1, form.shape_2.1)

```

### Simulación con modelo *Textremal* con correlación *Bessel*

```

y_mapa<- seq(from=-35,to=-30,length = 30)
x_mapa<- seq(from=-58.5,to=-53,length = 30)
coord_mapa<- cbind(x_mapa, y_mapa)
coord_mapa<- as.data.frame(coord_mapa)

# Calculamos la distancia euclidiana de las 38 estaciones con la grilla de 30x30
de coordenadas para simular:

dist_euclidiana_mapa<-array(0,c(30,30,38))
coord_mapa<-as.data.frame(coord_mapa)
coord_corregida[17,1]<- coord_corregida[17,1]+0.01
coord_corregida[12,2]<- coord_corregida[12,2]+0.01
coord_corregida[16,2]<-coord_corregida[16,2] - 0,01

```

```
for (k in 1:38) {
for (j in 1:30) {
for (i in 1:30) {
  dist_euclidiana_mapa[i,j,k] <- sqrt((coord_corregida$x_corregida[k] -
coord_mapa$x_mapa[j])^ 2 +
  (coord_corregida$y_corregida[k] - coord_mapa$y_mapa[i])^ 2)
}
}
}

# Saco la posición de la mínima distancia euclidiana de cada estación en com-
paración a la grilla 30x30:

pos_mapa <- matrix(0, 38, 2)
a=1
for (j in 1:2) {
for (i in 1:38) {
print(a)
pos_mapa[i,j] <- which(dist_euclidiana_mapa[,i] == min(dist_euclidiana_mapa[,i]),
arr.ind = TRUE) [,j]
colnames(pos_mapa) <- c("fila", "col")
a=a+1
}
}

pos_mapa <- as.data.frame(pos_mapa)

# Consigo valores simulados para el modelo Textremal con correlación Bessel:

Simulacion_mapa_BE_1.1 <- rmaxstab(147, coord_mapa, cov.mod = "tbessel",
sill = 1, nugget=0.24978,
range = 0.09416, smooth = 146.24461,
grid = TRUE, DoF=3.54467, control = list(nlines = 1))
```

# Consigo los valores simulados de la grilla que están mas cerca para la simulación anterior:

```
maximo_mapa_BE_1_1 <- matrix(0,nrow = 38,ncol = 147)
for (i in 1:147) {
for (j in 1:38) {
  maximo_mapa_BE_1_1[j,i] <- Simulacion_mapa_BE_1_1[pos_mapa_fila[j],
pos_mapa$col[j],i]
} }

```

#Llevo la simulación (*Frechet Unitaria*) a una *Frechet*:

```
frechet_transf_mapa_BE_1_1 <- matrix(0,ncol = 147,nrow = 38)
for (i in 1:38) {
frechet_transf_mapa_BE_1_1[i,] <- frech2gev(maximo_mapa_BE_1_1[i,],
estimacion_conjunta_1_1[i,2],
estimacion_conjunta_1_1[i,3],
0.08)
}

```

# Sacamos los máximos de la matriz simulada con el modelo *Textremal* con correlación *Bessel* que ya llevamos a *Frechet*:

```
frechet_transf_mapa_BE_1_1 <- apply(frechet_transf_mapa_BE_1_1,2,max)

```

# Se realiza un gráfico de caja para comprar el promedio de precipitaciones máximas conseguido mediante la simulación contra los valores máximos observados:

```
histograma_mapa_BE_1_1 <- as.data.frame(cbind(c(frechet_transf_mapa_BE_1_1,
maxtrimestre),
c(rep("Simulado",147),rep("Observado",147))))
histograma_mapa_BE_1_1$V1 <- as.numeric(histograma_mapa_BE_1_1$V1)

```

## CAPÍTULO 5. ANEXO

---

```
names(histograma_mapa_BE_1_1) <- c("Precipitaciones.mm", "Grupo")
```

```
# Gráfico de cajas simulado contra observado:
```

```
ggplot(data = histograma_mapa_BE_1_1, aes(y=Precipitaciones.mm, fill=Grupo, color=Grupo)) +  
geom_boxplot(alpha = 0.2) +  
facet_wrap( Grupo) +  
labs(title = "Precipitaciones máximas por trimestre",  
x="",  
y="Precipitaciones mm")
```

## 5.2. Anexo B - Metodología y Resultados

### 5.2.1. Métodos de los Momentos Pesados

Sea  $X \sim F$ , se define el *Método de los Momentos Pesados* de la siguiente forma,

$$M_{p,r,s} = E[X^p F(X)^r (1 - F(X))^s] \quad (5.1)$$

Donde  $p, r, s$  son reales no negativos.

Este método se considera de utilidad potencial para derivar expresiones explícitas para los parámetros de distribuciones  $F$  cuyas formas inversas las cuáles se denotan de la siguiente forma  $x = x(F)$  pueden definirse explícitamente. Si esto sucede se puede definir a la expresión  $M_{p,r,s} = E[X^p F(X)^r (1 - F(X))^s]$ ,

$$M_{p,r,s} = \int_0^1 x(F)^p F^r (1 - F)^s dF \quad (5.2)$$

Si  $r = s = 0$  y  $p$  un número real no negativo, entonces,  $M_{p,0,0}$ , representa el momento convencional sobre el origen de orden  $p$ . Si  $M_{p,0,0}$  existe y  $X$  es una función continua de  $F$ , entonces,  $M_{p,r,s}$  existe para todos los números reales no negativos  $r$  y  $s$ . Si  $r$  y  $s$  son reales no negativos, entonces, para estimar los parámetros de  $F$  se parte de,

$$M_{1,r,s} = E[X F(X)^r (1 - F(X))^s] \quad (5.3)$$

Denotamos a  $\omega_r(\theta) = M_{1,r,0} = E[X G_\theta^r(X)]$ ,  $r \in \mathbb{N}_0$  donde  $G_\theta$  es la DGVE y  $X \sim G_\theta$ . Consideramos  $\epsilon < 1$ , dado que, en caso contrario  $\overline{G}_\theta$  es de variación regular con índice  $1/\epsilon$  y si esto sucede tenemos  $\omega_0 = \infty$ . Se tiene el estimador plug-in de  $\omega_r$ ,

$$\hat{\omega}_r(\theta) = \int_{-\infty}^{+\infty} x G_{\theta}^r(x) dF_n(x) \quad (5.4)$$

A la hora de estimar  $\theta$  se le resuelven las siguientes ecuaciones,  $\omega_r(\theta) = \hat{\omega}_r(\theta)$ ,  $r = 0, 1, 2$ .

Dado que  $G_{\theta}(X_{1:n}, \dots, X_{n:n}) \stackrel{d}{=} (U_{1:n}, \dots, U_{1:n})$ , donde estos son los estadísticos de orden de una sucesión *i.i.d* de variables uniformes, y se puede escribir de la siguiente forma,

$$\hat{\omega}_r(\theta) = \frac{1}{n} \sum_{j=1}^n X_{j:n} U_{j:n}^r, \quad r = 0, 1, 2 \quad (5.5)$$

Si tenemos  $\epsilon < 1$  o  $\epsilon \neq 0$ , se cumple,

$$\omega_r(\theta) = \frac{1}{1+r} \left( \mu - \frac{\sigma}{\epsilon} (1 - \Gamma(1 - \epsilon)(1+r)^{\epsilon}) \right) \quad (5.6)$$

Utilizando las ecuaciones 5.5 y 5.6 se logra el estimador de los momentos pesados para  $\hat{\theta}$ , ver (Greenwood, Landwehr, Matalas, y Wallis, 1979).

Este método conviene utilizarlo, cuando las condiciones de regularidad para las propiedades asintóticas a la hora de aplicar la *Estimación Máxima Verosímil* no se cumplan. En general, se aplica esta técnica cuanto no contamos con una cantidad de observaciones lo suficientemente grande.

### 5.2.2. Selección de Modelos

A continuación se presentan los diferentes criterios de selección de modelos.

A la hora de trabajar con el *criterio de información*, es necesario introducir el concepto de *verosimilitud compuesta*:

**Definición 5.2.1** *Verosimilitud Compuesta:* Sea  $f(x, \theta)$ ,  $x \in \mathcal{X}$ ,  $\theta \in \Theta$ , un modelo estadístico paramétrico, con  $\mathcal{X} \subseteq \mathbb{R}^n$ ,  $\Theta \subseteq \mathbb{R}^d$ ,  $n \geq 1$ ,  $d \geq 1$ . Considere un conjunto de eventos  $\{\mathcal{A}_i : \mathcal{A}_i \subseteq \mathcal{F}, i \in I\}$ , donde  $I \subseteq \mathbb{N}$  y  $\mathcal{F}$  es una  $\sigma$ -álgebra sobre  $\mathcal{X}$ . Entonces, una función de verosimilitud compuesta es una función de  $\theta$  definida de la siguiente forma:

$$CL_f(\theta, x) = \prod_{i \in I} f(x \in \mathcal{A}_i; \theta)^{\omega_i} \quad (5.7)$$

Donde  $f(x \in \mathcal{A}_i; \theta) = f(\{x_j \in x : x_j \in \mathcal{A}_i\}, \theta)$ , con  $x = (x_1, \dots, x_n)$ , mientras  $\omega_i$ ,  $i \in I$ , es un conjunto de pesos adecuados.

Esta idea de la utilidad que brinda utilizar la probabilidad compuesta surge naturalmente en un marco de función de estimación, ver (Heyde, 1997). De hecho, dado el conjunto de eventos realizados  $\{\mathcal{A}_i \subseteq \mathcal{F}, i \in I\}$ , el estimador de *máxima verosimilitud compuesta* suele definirse como una solución de la ecuación de verosimilitud compuesta.

$$\nabla \log CL_f(\theta, x) = 0 \quad (5.8)$$

Donde  $\nabla \log CL_f(\theta, x) = \sum_{i \in I} \nabla \log f(x \in \mathcal{A}_i; \theta)^{\omega_i}$  es la *función score compuesta*. Se usará la notación  $\nabla h(\theta)$  para el vector columna de las primeras derivadas parciales de la función  $h(\theta)$ , mientras  $\nabla^2 h(\theta)$  es la matriz simétrica de segundas derivadas. Dado que la *función score compuesta* es una combinación lineal de funciones de estimación no sesgadas, entonces en condiciones de regularidad adecuadas, el *estimador de máxima verosimilitud compuesta*  $\hat{\theta}_{MCL}(X)$  es consistente y se distribuye asintóticamente Normal:

$$\hat{\theta}_{MCL}(X) \xrightarrow{d} \mathcal{N}(\theta, H(\theta)^{-1} J(\theta) H(\theta)^{-T}) \quad (5.9)$$

Donde:  $J(\theta) = \text{var}_{f(x)}\{\nabla \log CL(\theta, X)\}$  y  $H(\theta) = E_{f(x)}\{\nabla^2 \log CL(\theta, X)\}$ .

Aplicando los métodos de verosimilitud compuesta es posible realizar una selección de modelos. A continuación, se presenta un procedimiento de selección de modelo predictivo basado en la siguiente generalización de la *divergencia de Kullback-Leibler*.

**Definición 5.2.2** *Divergencia de Kullback-Leibler: Dada dos funciones de densidad  $g(z)$  y  $f(z)$  para una variable aleatoria  $Z$  la información compuesta asociada a Kullback-Leiber se define por la cantidad no negativa:*

$$D_p(f_\theta, g) = \sum_{i=1}^{k-1} \sum_{j=i+1}^k \omega_{i,j} E \left( \log \frac{g(Z_i, Z_j)}{f(Z_i, Z_j, \theta)} \right) \quad (5.10)$$

Donde  $Z = (Z_1, \dots, Z_k) \sim g$ . La ecuación 5.10 es una combinación lineal de las divergencias de *Kullback-Leibler*, correspondiente a cada objeto de probabilidad que forma la función de probabilidad compuesta. Además, 5.10 es una cantidad clave para evaluar las propiedades de la inferencia de *verosimilitud compuesta* y los procedimientos de selección de modelos, bajo un modelo estadístico posiblemente mal especificado. Se puede mostrar que el modelo que minimiza la divergencia de *Kullback-Leibler* compuesta es el mismo modelo que minimiza:

$$TIC(f_\theta) = -2\mathcal{L}_p(\hat{\theta}_p, z) + 2 \left( \hat{J}(\theta_0) \hat{H}(\theta_0)^{-1} \right)^t \quad (5.11)$$

Donde  $\hat{J}(\theta_0)$  y  $\hat{H}(\theta_0)$  son estimadores consistentes de las matrices  $J(\theta_0)$  y  $H(\theta_0)$  y la ecuación 5.11 es una generalización del *Criterio de Información de Takeuchi* (*TIC*).

### 5.2.3. Resultados

#### Estimación de los parámetros mediante Máxima Verosimilitud y Método de los Momentos Pesados

	Máxima Verosimilitud			Momentos Pesados			
	Estación	Posición	Escala	Forma	Posición	Escala	Forma
25 de Agosto	48.01	19.46	0.23		48.61	20.70	0.16
Aeropuerto Carrasco	46.68	20.67	0.06		46.45	20.56	0.08
Aeropuerto Melilla	46.96	21.20	0.10		46.77	21.12	0.11
Artigas	55.49	24.75	0.17		55.85	25.71	0.13
Baltasar Brum	63.18	30.67	0.07		63.02	30.77	0.07
Barriga Negra	53.74	21.74	0.04		53.79	22.34	0.02
Bernabe Rivera	59.45	25.76	0.08		59.74	26.75	0.05
Cañada Grande (P. Mellizos)	56.10	23.77	0.04		56.16	24.18	0.03
Catalán Grande	58.31	24.29	0.14		58.71	25.40	0.10
Cerro Amarillo	56.64	24.43	0.09		56.86	25.28	0.06
Cerro Chato	56.86	21.69	-0.07		56.88	21.45	-0.07
Cerro Colorado	51.10	22.59	0.05		50.92	22.72	0.06
Colonia	46.90	21.32	0.08		46.43	20.51	0.13
Coronilla	47.88	18.38	0.04		48.39	19.30	-0.01
Cuchilla Caraguata (N)	51.22	20.36	0.13		51.16	20.42	0.13
Durazno	53.40	23.50	-0.02		53.24	23.73	-0.02
Guichón	55.06	24.50	0.14		55.06	24.80	0.13
Marmaraja	48.91	20.57	0.06		49.07	21.19	0.04

Cuadro 5.1: Estimación de los parámetros de la DVEG por el método de máxima verosimilitud y por el método de los momentos pesados en el período 1982-2020

CAPÍTULO 5. ANEXO

Estación	Máxima Verosimilitud			Momentos Pesados		
	Posición	Escala	Forma	Posición	Escala	Forma
Mataojo	47.23	20.69	0.03	46.79	19.75	0.08
Melo	53.47	22.30	-0.03	53.53	23.04	-0.05
Mercedes	48.23	21.38	0.10	47.74	20.51	0.14
Palmitas	48.61	24.00	0.05	47.77	22.53	0.12
Paso de los Toros	53.60	23.21	0.01	53.88	24.01	-0.02
Paysandú	53.31	23.30	0.01	53.26	23.67	0.01
Piedras Coloradas	73.53	38.11	0.00	74.24	40.13	-0.04
Pirarajá	51.81	23.35	0.06	52.26	24.64	0.01
Prado	48.06	20.82	0.02	47.95	21.04	0.02
Punta del Este	44.53	20.08	0.04	44.30	20.02	0.06
Quebracho	59.89	29.93	0.02	60.02	30.61	0.00
Rivera	57.62	22.65	0.15	57.92	23.36	0.12
Rocha	49.56	19.71	0.18	49.56	19.98	0.17
Salto	55.15	24.95	0.06	54.94	24.14	0.08
San Antonio	47.32	20.41	0.22	47.93	21.69	0.15
San Jacinto	45.17	19.57	0.16	45.25	19.87	0.14
Sarandí Grande	50.37	21.70	0.07	50.11	21.10	0.09
Tala	44.10	20.37	0.10	44.21	20.79	0.09
Tambores	57.35	24.36	-0.07	57.30	25.25	-0.08
Tomás Gomensoro	58.11	28.32	0.10	57.50	27.24	0.14
Treinta y Tres	54.97	22.57	0.15	54.70	22.13	0.17
Trinidad	51.75	24.50	-0.01	50.96	23.17	0.05
Young	55.46	25.73	-0.08	55.57	26.94	-0.10

Cuadro 5.2: Estimación de los parámetros de la DVEG por el método de máxima verosimilitud y por el método de los momentos pesados en el período 1982-2020

**Estimación Conjunta de las 41 estaciones meteorológicas ó pluviométricas iniciales**

Estación	Posición	Escala	Forma
25 de Agosto	48.37	21.29	0.08
Aeropuerto Carrasco	47.10	20.51	0.08
Aeropuerto Melilla	47.24	24.76	0.08
Artigas	60.33	27.15	0.08
Baltasar Brum	59.37	27.05	0.08
Barriga Negra	49.71	21.34	0.08
Bernabe Rivera	60.62	27.49	0.08
Cañada Grande	54.02	24.38	0.08
Catalan Grande	59.17	26.47	0.08
Cerro Amarillo	59.64	26.88	0.08
Cerro Chato	52.30	22.63	0.08
Cerro Colorado	49.95	21.67	0.08
Colonia	48.26	21.86	0.08
Coronilla	48.55	20.66	0.08
Cuchilla Caraguata Norte	55.15	23.90	0.08
Durazno	51.53	22.87	0.08
Guichón	54.50	24.63	0.08
Marmaraja	48.82	20.86	0.08
Mataojo	47.72	20.37	0.08
Melo	54.50	23.27	0.08
Mercedes	51.82	23.72	0.08

Cuadro 5.3: Estimación de los parámetros de la DVEG por el método de estimación conjunta para las primeras 21 estaciones meteorológicas ó pluviométricas para el periodo 1982-2020.

CAPÍTULO 5. ANEXO

---

Estación	Posición	Escala	Forma
Palmitas	51.08	23.24	0.08
Paso de los Toros	53.19	23.69	0.08
Paysandu	54.53	25.02	0.08
Piedras Coloradas	54.44	24.79	0.08
Piraraja	50.37	21.52	0.08
Prado	47.04	20.56	0.08
Punta del Este	46.74	19.85	0.08
Quebracho	55.75	25.55	0.08
Rivera	58.84	25.99	0.08
Rocha	48.14	20.24	0.08
Salto	57.23	26.31	0.08
San Antonio	48.26	21.09	0.08
San Jacinto	48.02	20.89	0.08
Sarandi Grande	50.43	22.26	0.08
Tala	48.58	21.11	0.08
Tambores	55.93	24.89	0.08
Tomas Gomensoro	60.21	27.51	0.08
Treinta y Tres	51.91	22.10	0.08
Trinidad	50.99	22.80	0.08
Young	53.52	24.35	0.08

Cuadro 5.4: Estimación de los parámetros de la DVEG por el método de estimación conjunta para las últimas 20 estaciones meteorológicas ó pluviométricas para el periodo 1982-2020.

**Estimación Conjunta de las 40 estaciones meteorológicas ó pluviométricas sin considerar Piedras Coloradas**

Estación	Posición	Escala	Forma
25 de Agosto	48.33	21.15	0.07
Aeropuerto Carrasco	47.12	20.44	0.07
Aeropuerto Melilla	47.26	23.63	0.07
Artigas	59.68	26.61	0.07
Baltasar Brum	58.77	26.46	0.07
Barriga Negra	49.60	21.30	0.07
Bernabe Rivera	59.96	26.90	0.07
Cañada Grande	53.69	23.97	0.07
Catalan Grande	58.58	26.00	0.07
Cerro Amarillo	59.03	26.35	0.07
Cerro Chato	52.06	22.51	0.07
Cerro Colorado	49.83	21.57	0.07
Colonia	48.22	21.55	0.07
Coronilla	48.50	20.69	0.07
Cuchilla Caraguata Norte	54.77	23.72	0.07
Durazno	51.32	22.61	0.07
Guichón	54.15	24.20	0.07
Marmaraja	48.76	20.86	0.07
Mataojo	47.71	20.39	0.07
Melo	54.15	23.18	0.07
Mercedes	51.61	23.28	0.07

Cuadro 5.5: Estimación de los parámetros de la DVEG por el método de estimación conjunta para las primeras 21 estaciones meteorológicas ó pluviométricas para el periodo 1982-2020.

CAPÍTULO 5. ANEXO

---

Estación	Posición	Escala	Forma
Palmitas	50.90	22.85	0.07
Paso de los Toros	52.90	23.38	0.07
Paysandu	54.17	24.50	0.07
Piraraja	50.22	21.50	0.07
Prado	47.06	20.47	0.07
Punta del Este	46.78	19.92	0.07
Quebracho	55.33	25.00	0.07
Rivera	58.27	25.61	0.07
Rocha	48.11	20.33	0.07
Salto	56.74	25.71	0.07
San Antonio	48.22	20.98	0.07
San Jacinto	47.99	20.81	0.07
Sarandi Grande	50.28	22.05	0.07
Tala	48.53	21.03	0.07
Tambores	55.50	24.52	0.07
Tomas Gomensoro	59.56	26.87	0.07
Treinta y Tres	51.69	22.07	0.07
Trinidad	50.82	22.51	0.07
Young	53.21	23.90	0.07

Cuadro 5.6: Estimación de los parámetros de la DVEG por el método de estimación conjunta para las últimas 20 estaciones meteorológicas ó pluviométricas para el periodo 1982-2020.

**Estimación Conjunta de los parámetros de las 38 estaciones finales meteorológicas ó pluviométricas**

Estación	Posición	Escala	Forma
Aeropuerto Carrasco	47.09	20.87	0.06608
Aeropuerto Melilla	48.13	21.38	0.06608
Artigas	59.59	26.13	0.06608
Baltasar Brum	58.69	26.46	0.06608
Barriga Negra	49.56	21.02	0.06608
Bernabe Rivera	59.87	26.61	0.06608
Cañada Grande	53.64	24.36	0.06608
Catalan Grande	58.50	25.48	0.06608
Cerro Amarillo	58.95	26.00	0.06608
Cerro Chato	52.01	22.03	0.06608
Cerro Colorado	49.79	21.50	0.06608
Colonia	48.19	22.75	0.06608
Coronilla	48.47	20.37	0.06608
Cuchilla Caraguata Norte	54.70	22.87	0.06608
Durazno	51.28	22.89	0.06608
Guichón	54.09	24.57	0.06608
Marmaraja	48.72	20.60	0.06608
Mataojo	47.68	20.27	0.06608
Melo	54.09	22.04	0.06608

Cuadro 5.7: Estimación de los parámetros de la DVEG por el método de estimación conjunta para las primeras 19 estaciones meteorológicas ó pluviométricas finales para el periodo 1982-2020.

CAPÍTULO 5. ANEXO

---

Estación	Posición	Escala	Forma
Mercedes	51.56	24.31	0.06608
Palmitas	50.86	23.81	0.06608
Paso de los Toros	52.85	23.52	0.06608
Paysandu	54.11	25.28	0.06608
Piraraja	50.18	21.00	0.06608
Prado	47.04	21.00	0.06608
Punta del Este	46.75	19.84	0.06608
Quebracho	55.27	25.61	0.06608
Rivera	58.19	24.77	0.06608
Rocha	48.07	19.81	0.06608
Salto	56.67	26.23	0.06608
San Jacinto	48.19	21.33	0.06608
Sarandi Grande	50.24	22.34	0.06608
Tala	48.50	21.19	0.06608
Tambores	55.43	24.29	0.06608
Tomas Gomensoro	59.48	26.87	0.06608
Treinta y Tres	51.64	21.25	0.06608
Trinidad	50.77	23.04	0.06608
Young	53.16	24.58	0.06608

Cuadro 5.8: Estimación de los parámetros de la DVEG por el método de estimación conjunta para las últimas 19 estaciones meteorológicas ó pluviométricas finales para el período 1982-2020.

### 5.2.3.1. Gumbel

#### Transformación a Frechet Unitaria

Vale recordar que cualquiera de las tres distribuciones de la *TVE* (Weibull, Frechet y Gumbel), se podía aplicar una transformación, ver 2.10 y lograr una distribución *Frechet Unitaria*. Se decidió trabajar en primer instancia, con el parámetro de forma asumido,  $\hat{\epsilon} = 0$ , por lo tanto, la distribución seleccionada de la *TVE* es la *Gumbel*.

#### Estimación espacial mediante procesos máx-estables

Una vez aplicada la *Teoría de Valores Extremos*, se procedió a realizar una estimación espacial mediante la técnica *procesos máx-estables*. En primer lugar, se analizó el comportamiento del coeficiente extremal, el cual nos permitió estudiar la correlaciones dos a dos de las diferentes estaciones. Observando la Figura 5.1, en primer lugar no se ve mucha variabilidad, lo que da indicio de que las precipitaciones en estaciones cercanas tienen un comportamiento similar, pero mientras las distancias entre estaciones va en aumento, el comportamiento de las precipitaciones no son semejantes. También analizando la Figura 5.2 se apreció un comportamiento similar al que brinda el coeficiente extremal, estaciones cercanas tienen un comportamiento similar, mientras que a mayor distancia entre estaciones el comportamiento de las mismas son más independiente. En función del coeficiente extremal, no es muy claro el modelo *máx-estable* a utilizar. Sin embargo, sabemos que aquellos modelos que presentan una función del coeficiente extremal acotada por una constante pequeño, por ejemplo 2, no se pueden utilizar, como es el caso del modelo de *Shalther*.

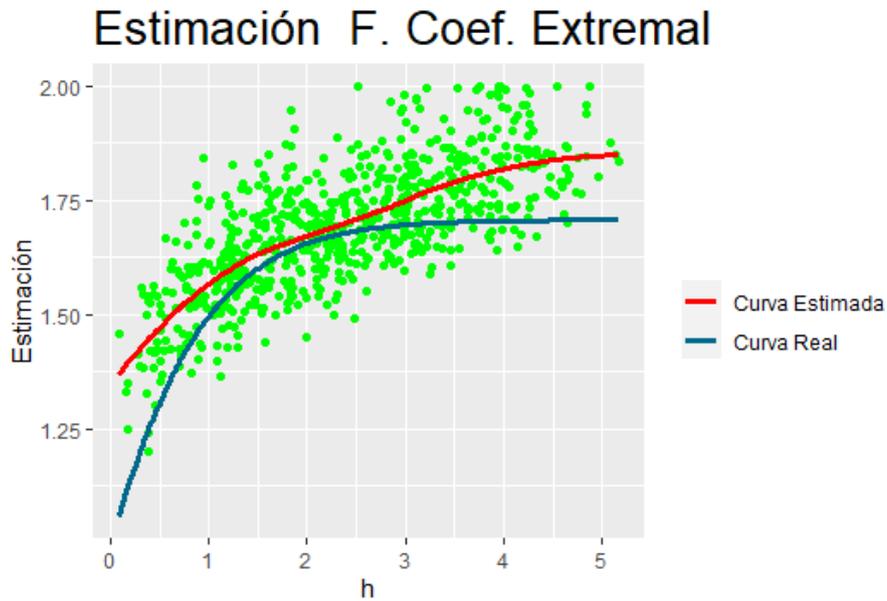


Figura 5.1: Coeficiente extremal estimada considerando una DVE del tipo Gumbel.

Fuente: elaboración propia en R.

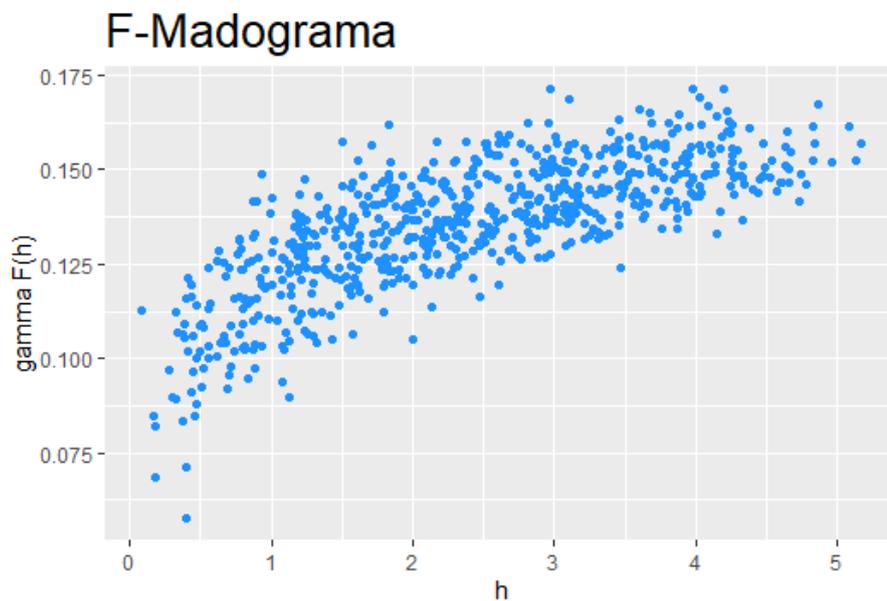


Figura 5.2: F-madograma estimada considerando una DVE del tipo Gumbel. Fuente:

elaboración propia en R.

El primer escenario con el que se trabajó, fue considerar la estimación del parámetro de forma igual a cero,  $\hat{\epsilon} = 0$ . Por lo tanto, se consideró una distribución de la *TVE*

*Gumbel*. Se realizaron los mismos estudios con respecto a considerar la estimación del parámetro de forma mayor a cero,  $\hat{\epsilon} > 0$ . Primero se analizó el comportamiento del *coeficiente extremal*, se estudiaron el *f-madograma* y *Madograma*, se seleccionaron mediante el criterio de información *TIC* los mejores de modelos paramétricos de los *procesos máx-estables* y por último, se simuló y evaluó el comportamiento de los diferentes modelos. A continuación se presentan los resultados más relevantes, que permiten observar que considerar la estimación del parámetro de forma igual a cero no da resultados considerablemente aceptables.

Primero, se estudió la matriz de varianza y covarianza del modelo de *Smith*:

$$\Sigma = \begin{pmatrix} 8,988e - 06 & -7,860e - 06 & 6,766e - 06 \\ -7,860e - 06 & 7,383e - 06 & -6,875e - 06 \\ 6,766e - 06 & -6,875e - 06 & 6,029e - 06 \end{pmatrix} \quad (5.12)$$

El determinante de dicha matriz es cercano a cero, lo que produce problemas de invertibilidad, por lo tanto, el modelo de *Smith* presenta problemas numéricos.

### Selección de Modelos

Se estudiaron todos los modelos posibles con sus diferentes correlaciones para lograr una estimación espacial mediante *procesos máx-estables*. A la hora de seleccionar el modelo nuevamente se utilizó el criterio de información *TIC*. En la Tabla 5.9, el modelo *T-extremal* con correlaciones *Power-Exponencial* y *Bessel* son los que presentaron menor *TIC*.

Modelo	TIC
Geo-Gauss W-M	916868,00
Geo-Gauss Pow-Exp	916882,10
t-Extremal W-M	916820,40
t-Extremal Pow-Exp	916817,20
t-Extremal Bessel	916509,80

Cuadro 5.9: Valor del TIC para los seis modelos con menor TIC en estimación espacial mediante procesos máx-estables para el período 1982-2020 en 38 estaciones climatológicas y pluviométricas considerando una *DVE Gumbel*

### Simulación

Luego de estudiar cuales son los modelos más interesantes para poder realizar estimaciones espaciales de las precipitaciones en todo el territorio uruguayo, se procedió a simular. Se recuerda que el valor del parámetro de forma estimado  $\hat{\epsilon}$  es cero, por lo tanto, se trabajó con una *DVE* del tipo *Gumbel*. A la hora de simular se tomó la misma grilla de coordenadas que se trabajó considerando la estimación del parámetro de forma mayor a cero, y se realiza él mismo procedimiento de simulación.

Observando la Tabla 5.9 el modelo que mejor se ajustó a los datos observados fue el *t-Extremal* con diferentes correlaciones, estas fueron: *Bessel*, *Power-Exponencial* y *Whittle-Mattern*.

Los valores simulados eran *Frechet Unitarios* y se debió realizar una transformación para poder llevarlos a una *Gumbel*.

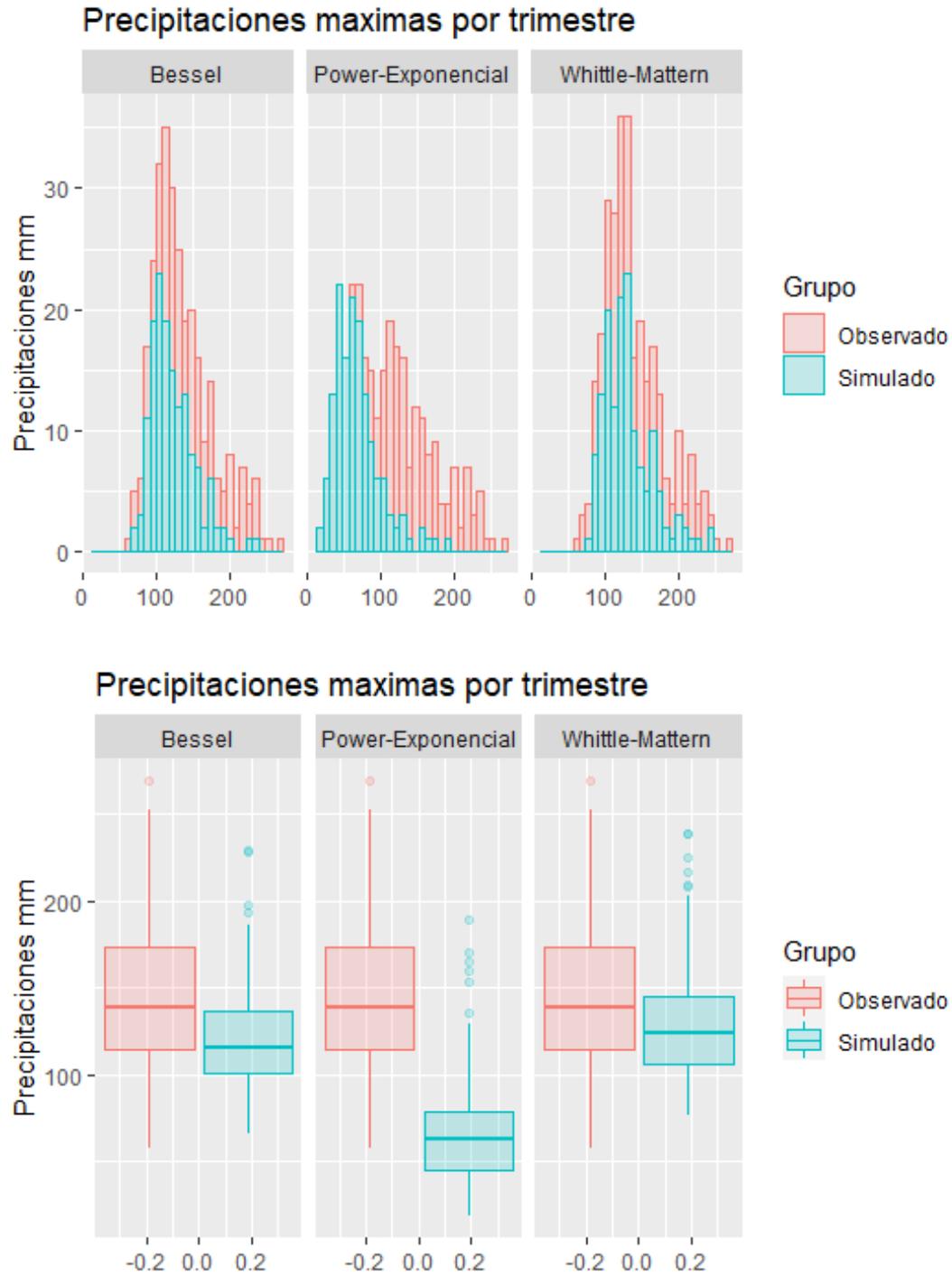


Figura 5.3: Comparación entre los valores promedios de los máximos observados en los 147 trimestres en el período 1982-2020 de las 38 estaciones contra los valores simulados, considerando un modelo t-extremal con correlaciones: Bessel, Whittle-Mattern y Power Exponencial

Analizando la Figura 5.3 se apreció que los modelos que mejor ajustan según el promedio de los máximos de las precipitaciones observadas en las 38 estaciones durante los 147 trimestres, son los modelos *T-extremal* con correlaciones *Bessel* y *Whittle-Mattemn*. Ambos modelos, no logran captar la precipitación máxima que se registró en Uruguay en el período 1982-2020. Por otro lado, los datos observados presentan una mayor variabilidad en las precipitaciones máximas en comparación a los datos que se consiguen con ambos modelos. Se apreció que los modelos paramétricos no presentaron un comportamiento correcto.

## 5.2.4. Frechet

## Transformación Frechet Unitaria

	1982-1	1982-2	1982-3	1982-4	1983-1	2018-3	2018-4	2019-1	2020-1	2020-4
Aeropuerto Carrasco	1.63	0.96	0.96	0.51	0.51	2.56	400.77	0.67	0.44	0.55
Aeropuerto Melilla	1.04	1.34	1.56	0.52	1.37	0.95	174.26	1.26	0.53	0.49
Artigas	9.49	1.09	2.32	4.71	98.77	3.46	4.33	15.71	1.14	0.70
Baltasar Brum	0.71	0.59	0.71	4.49	1.43	1.99	7.82	5.82	1.65	0.59
Barriga Negra	0.44	338.23	1.62	0.84	5.23	1.50	19.40	120.65	0.68	0.88
Bernabe Rivera	0.73	0.57	0.94	1.26	142.65	0.76	1.31	300.68	1.31	1.13
Cañada Grande (P. Mellizos)	1.93	4.03	3.87	6.33	1.18	2.18	4.03	264.25	8.78	0.33
Catalán Grande	2.65	1.16	1.03	3.87	79.58	1.16	3.87	4.50	0.71	0.59
Cerro Amarillo	1.58	2.13	2.38	1.22	60.66	2.57	0.94	8.77	0.78	1.22
Cerro Chato	0.69	1.32	9.40	0.79	0.77	1.29	0.90	10.26	0.90	0.19
Cerro Colorado	1.00	9.18	2.77	0.80	8.37	0.87	10.07	20.12	0.80	0.53
Colonia	6.22	0.67	7.43	1.36	0.51	2.06	7.40	3.40	0.75	1.06
Coronilla	0.60	0.93	13.29	4.16	0.52	1.37	5.84	3.60	0.93	0.41
Cuchilla Caraguata (N)	1.08	2.83	2.02	0.71	1.51	46.64	0.74	1.51	1.64	0.23
Durazno	2.65	4.59	2.51	0.43	0.86	1.73	30.91	9.33	1.04	0.23
Guichón	4.14	2.82	0.90	0.77	1.73	37.85	32.18	34.90	0.74	0.83
Marmaraja	0.33	9.15	0.66	0.96	0.60	4.68	1.34	4.46	0.66	0.26
Mataojo	0.18	4.88	3.46	0.68	1.66	0.83	56.82	1.06	0.28	0.17
Melo	2.20	3.40	1.02	4.40	1.67	3.26	0.63	0.62	2.52	0.47
Mercedes	42.74	9.41	1.29	0.56	1.28	1.01	4.69	1.54	0.82	0.31
Palmitas	2.26	19.41	1.47	0.62	12.62	0.68	59.41	17.81	0.57	0.50
Paso de los Toros	0.77	1.65	4.08	0.48	1.03	1.55	6.36	90.81	0.72	0.89
Paysandú	5.00	3.66	1.08	1.22	1.21	0.69	8.21	17.75	0.74	0.94
Pirarajá	0.49	1.03	2.49	1.49	1.24	0.56	2.61	6.61	0.56	0.68
Prado	1.53	3.47	4.05	0.71	1.05	2.71	29.05	0.54	0.45	0.49
Punta del Este	1.97	4.93	16.18	0.92	0.89	2.45	134.27	0.41	0.47	0.96
Quebracho	1114.49	529.78	5.88	3.27	16.91	4.65	5.65	3.02	1.18	0.35
Rivera	0.43	2.84	1.60	9.34	57.88	1.27	2.07	2.17	0.45	0.43
Rocha	3.25	7.31	2.39	0.57	0.95	2.95	0.63	44.17	1.19	0.23
Salto	0.66	0.71	0.52	3.23	2.80	0.75	2.64	5.96	1.59	0.98
San Jacinto	0.38	1.40	6.16	0.35	1.77	0.54	2.25	9.04	0.26	0.26
Sarandí Grande	0.72	553.27	4.32	0.98	14.54	3.02	14.54	1.76	1.17	0.28
Tala	0.26	0.70	1.36	0.43	3.66	0.97	4.43	5.61	1.07	0.56
Tambores	3.49	1.44	8.78	1.50	14.80	6.63	2.74	1.44	0.96	0.82
Tomás Gomensoro	0.69	0.89	1.19	0.92	1.43	0.92	0.99	9.80	0.80	0.89
Treinta y Tres	3.39	2.67	12.09	0.77	1.46	1.10	0.85	2.77	0.65	0.26
Trinidad	3.60	15.31	0.62	0.65	1.26	0.53	10.69	1.45	2.68	0.26
Young	2.11	1.74	0.73	0.55	3.25	2.62	9.06	9.64	9.92	0.40

Cuadro 5.10: Transformación a Frechet Unitaria de los datos observados para los primeros y últimos trimestres de las 38 estaciones meteorológicas ó pluviométricas.

### 5.2.5. Gráfica, Riesgos de la actualidad

Esta figura fue tomada de la tesis de Florencia Santiñaque, ver (Santiñaque Mesones, 2020)

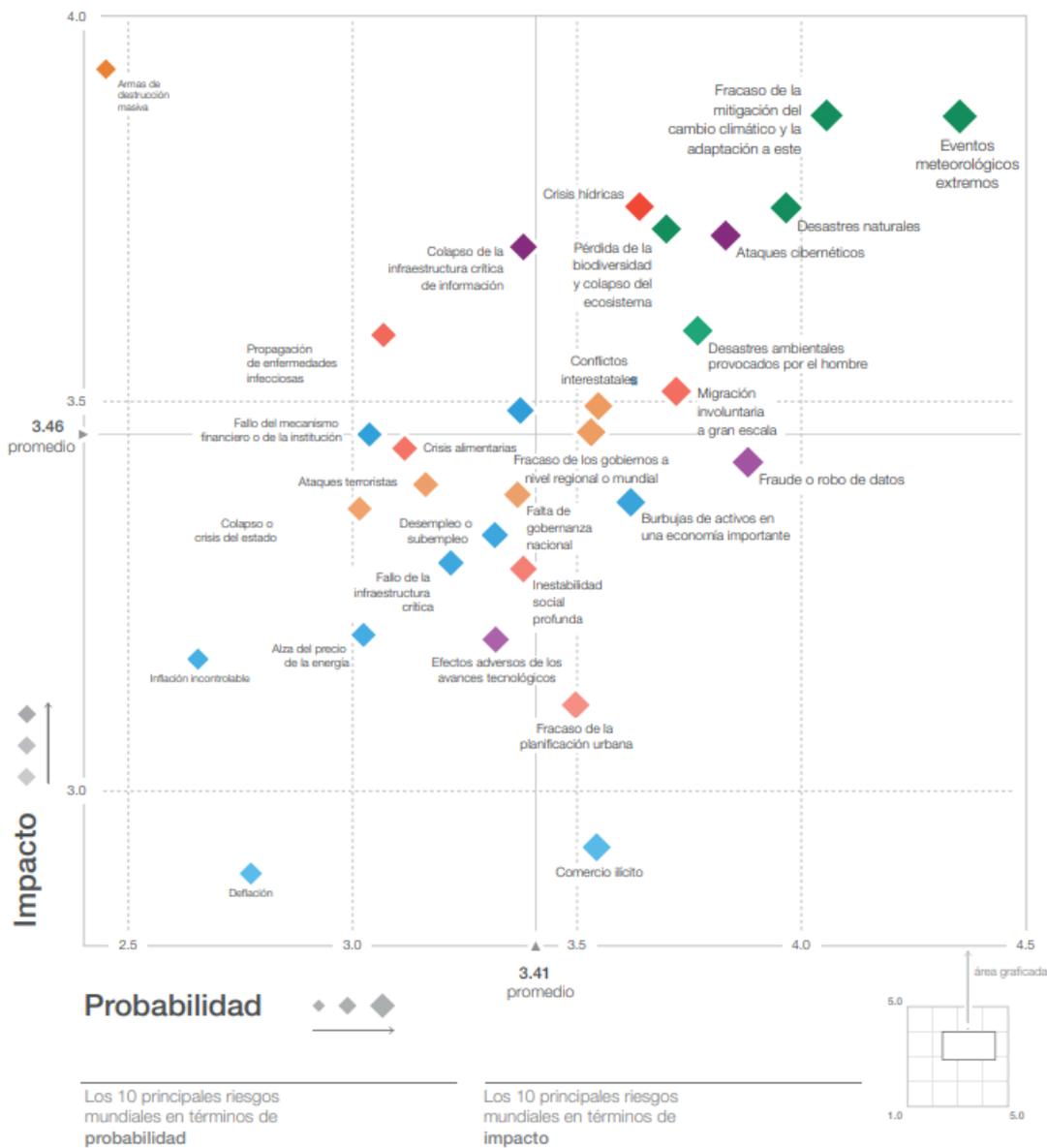


Figura 5.4: Capacidad destructiva y ocurrencia de eventos mundiales. Fuente: Informe de riesgos mundiales 2019. Fuente: World Economic Forum.

---

## Referencias

- Alves, M. I. F., y Neves, C. (2011). Extreme value distributions. *International encyclopedia of statistical science*, 2, 493–496.
- Aparicio-Effen, M., Arana, J., Ivar, Cortez, P., Coronel, G., Pastén, M., Nagy, G., ... Bidegain, M. (2016). Introducing hydro-climatic extremes and human impacts in Bolivia, Paraguay and Uruguay. En *Climate change and health* (pp. 449–473). Springer.
- Badii, M., Guillen, A., Lugo Serrato, O., y Aguilar Garnica, J. (2014). Non-parametric correlation and its application in scientific research. *International Journal of Good Conscience*, 9(2), 31–40.
- Balkema, A. A., y De Haan, L. (1978). Limit distributions for order statistics. *Theory of Probability & Its Applications*, 23(1), 77–92.
- Beirlant, J., Goegebeur, Y., Segers, J., y Teugels, J. L. (2004). *Statistics of extremes: theory and applications* (Vol. 558). John Wiley & Sons.
- Brown, B. M., y Resnick, S. I. (1977). Extreme values of independent stochastic processes. *Journal of Applied Probability*, 14(4), 732–739.
- Cardarello, M., y Luraghi, L. (2019). Análisis de valores extremos: una aplicación a temperaturas mínimas en Uruguay. Trabajo final de grado. Universidad de la República (Uruguay). Facultad de Ciencias Económicas y de Administración. <https://www.colibri.udelar.edu.uy/jspui/handle/20.500.12008/31396>.
- Cooley, D., Naveau, P., y Poncet, P. (2006). Variograms for spatial max-stable random fields. En *Dependence in probability and statistics* (pp. 373–390). Springer.
- Cressie, N. (1993). *Statistics for spatial data* New York. Wiley-Interscience.
- Davison, A. C., y Gholamrezaee, M. M. (2012). Geostatistics of extremes. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 468(2138), 581–608.

- Davison, A. C., y Smith, R. L. (1990). Models for exceedances over high thresholds. *Journal of the Royal Statistical Society: Series B (Methodological)*, 52(3), 393–425.
- De Haan, L. (1984). A spectral representation for max-stable processes. *The Annals of Probability*, 1194–1204.
- De Haan, L., y Ferreira, A. (2006). *Extreme value theory: an introduction* (Vol. 21). Springer.
- Dutfoy, A., Parey, S., y Roche, N. (2014). Multivariate extreme value theory-a tutorial with applications to hydrology and meteorology. *Dependence Modeling*, 2(1).
- Ehm, W. (1981). Sample function properties of multi-parameter stable processes. *Journal of Probability Theory*, 56(2), 195–228.
- Embrechts, P., Klüppelberg, C., y Mikosch, T. (2013). *Modelling extremal events: for insurance and finance* (Vol. 33). Springer Science & Business Media.
- Estenssoro Saavedra, F. (2010). Crisis ambiental y cambio climático en la política global: un tema crecientemente complejo para América Latina. *Universum (Talca)*, 25(2), 57–77.
- Fisher, R. A., y Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. En *Mathematical proceedings of the Cambridge philosophical society* (Vol. 24, pp. 180–190).
- Gaetan, C., y Guyon, X. (2010). *Spatial statistics and modeling* (Vol. 90). Springer.
- Gnedenko, B. (1943). Sur la distribution limite du terme maximum d’une serie aleatoire. *Annals of mathematics*, 423–453.
- González, D. S. (2011). Precipitaciones extremas y sus implicaciones en procesos de remoción en masa en la planificación urbana de tampico, méxico. *Cuadernos Geográficos*(48), 135–159.
- González, E. M., y Macías, E. (2011). *Análisis de máximos para datos espaciales de lluvias* (Tesis Doctoral no publicada). Centro de Investigación en Matemáticas.

- Greenwood, J. A., Landwehr, J. M., Matalas, N. C., y Wallis, J. R. (1979). Probability weighted moments: definition and relation to parameters of several distributions expressible in inverse form. *Water resources research*, 15(5), 1049–1054.
- Heyde, C. C. (1997). *Quasi-likelihood and its application: a general approach to optimal parameter estimation*. Springer.
- Hosking, J. R. M., Wallis, J. R., y Wood, E. F. (1985). Estimation of the generalized extreme-value distribution by the method of probability-weighted moments. *Technometrics*, 27(3), 251–261.
- Inda, H., y Mazzeo, N. (2012). Clima de cambios: nuevos desafíos de adaptación en Uruguay. En anuario OPYPA, MGAP, Montevideo.
- Jenkinson, A. F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society*, 81(348), 158–171.
- Kabluchko, Z., Schlather, M., y De Haan, L. (2009). Stationary max-stable fields associated to negative definite functions. *The Annals of Probability*, 37(5), 2042–2065.
- Khan, S., Kuhn, G., Ganguly, A. R., Erickson III, D. J., y Ostrouchov, G. (2007). Spatio-temporal variability of daily and weekly precipitation extremes in south america. *Water Resources Research*, 43(11).
- Leadbetter, M. R., Lindgren, G., y Rootzén, H. (2012). *Extremes and related properties of random sequences and processes*. Springer Science & Business Media.
- Moreno Romero, L. (2013). Precipitaciones máximas en el estado de Guanajuato, México. Tesis de maestría. Universidad de la República (Uruguay). Facultad de Ingeniería. <https://www.colibri.udelar.edu.uy/jspui/handle/20.500.12008/24127>.
- Opitz, T. (2012). A spectral construction of the extremal-t process. *arXiv preprint arXiv:1207.2296*.

- Ortega Sanchez, J. (2010). Valores Extremos. [Notas curso]. Centro de Investigación en Matemáticas (Cimat), Guanajuato, Mexico.
- Padoan, S. A., Ribatet, M., y Sisson, S. A. (2010). Likelihood-based inference for max-stable processes. *Journal of the American Statistical Association*, 105(489), 263–277.
- Resnick, S. I. (2008). *Extreme values, regular variation, and point processes* (Vol. 4). Springer Science & Business Media.
- Ribatet, M., Dombry, C., y Oesting, M. (2016). Spatial extremes and max-stable processes. *Extreme Value Modeling and Risk Analysis: Methods and Applications*, 179–194.
- Salado, A. I. A., Huerta, H. V., Guzmán, M. E. R., Lazalde, J. R. V., y Salado, C. A. A. (2013). Estimación del valor en riesgo en la bolsa mexicana de valores usando modelos de heteroscedasticidad condicional y teoría de valores extremos. *Economía Mexicana. Nueva Época*, 22(1), 177–205.
- Santiñaque Mesones, F. (2020). Análisis estadístico de las precipitaciones anuales extremas en Uruguay. Tesis de Maestría. Universidad de la República (Uruguay). Facultad de Ingeniería. <https://www.colibri.udelar.edu.uy/jspui/handle/20.500.12008/24577>.
- Schlather, M. (2002). Models for stationary max-stable random fields. *Extremes*, 5(1), 33–44.
- Smith, R. L. (1990a). Extreme value theory. *Handbook of applicable mathematics*, 7, 437–471.
- Smith, R. L. (1990b). Max-stable processes and spatial extremes. *Unpublished manuscript*, 205, 1–32.
- Varin, C., y Vidoni, P. (2005). A note on composite likelihood inference and model selection. *Biometrika*, 92(3), 519–528.
- Villacis Rivadeneira, E., y Marrero de León, N. (2017). Precipitaciones extremas en la ciudad de Quito, provincia de Pichincha-Ecuador. *Ingeniería Hidráulica y Ambiental*, 38(2), 102–113.

Zhou, C. (2010). The extent of the maximum likelihood estimator for the extreme value index. *Journal of Multivariate Analysis*, 101(4), 971–983.