



# COLEGIO DE POSTGRADUADOS

---

INSTITUCIÓN DE ENSEÑANZA E INVESTIGACIÓN  
EN CIENCIAS AGRÍCOLAS

**CAMPUS MONTECILLO**

POSTGRADO DE SOCIOECONOMÍA, ESTADÍSTICA E INFORMÁTICA  
ESTADÍSTICA

**Modelación espacio tiempo del ozono urbano en la  
Ciudad de México bajo el enfoque Max-stable**

Lucila Muñiz Merino

**T E S I S**

PRESENTADA COMO REQUISITO PARCIAL PARA  
OBTENER EL GRADO DE:

**MAESTRA EN CIENCIAS**

MONTECILLO, TEXCOCO, EDO. DE MÉXICO  
2013

---

La presente tesis titulada: **MODELACIÓN ESPACIO TIEMPO DEL OZONO URBANO EN LA CIUDAD DE MÉXICO BAJO EL ENFOQUE MAX-STABLE** realizada por la alumna: **LUCILA MUÑIZ MERINO** bajo la dirección del Consejo Particular indicado, ha sido aprobada por el mismo y aceptada como requisito parcial para obtener el grado de:

MAESTRA EN CIENCIAS  
SOCIOECONOMÍA ESTADÍSTICA E INFORMÁTICA  
ESTADÍSTICA

CONSEJO PARTICULAR

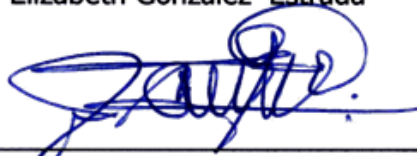
CONSEJERO

  
\_\_\_\_\_  
Dr. Humberto Vaquera Huerta

ASESOR

  
\_\_\_\_\_  
Dra. Elizabeth González Estrada

ASESOR

  
\_\_\_\_\_  
Dr. José A. Villaseñor Alva

ASESOR

  
\_\_\_\_\_  
Dr. Barry C. Arnold

Montecillo, Texcoco, Estado de México, Julio de 2013.

# Modelación espacio tiempo del ozono urbano en la Ciudad de México bajo el enfoque Max-stable

Lucila Muñiz Merino

Colegio de Postgraduados, 2013

## RESUMEN

En este trabajo se aplicó una metodología basada en el proceso max-stable para analizar la tendencia espacio temporal de datos ambientales. Los niveles máximos de ozono del Valle de México, entre los años 1998 a 2010, se modelaron mediante cuatro modelos max-stable: el de Smith el de Schlather, el Gaussiano Geométrico y el de Brown-Resnik. La estimación de parámetros de los cuatro modelos se hizo mediante máxima verosimilitud compuesta por pares. Posteriormente, se seleccionó el mejor modelo, de acuerdo con el criterio de información de verosimilitud de Takeuchy (TIC). La tendencia espacio temporal resultó estadísticamente significativa al fijar el parámetro de suavizamiento del modelo de Schlather y ajustar una regresión lineal a su rango. Lo anterior mostró un cambio en la estructura de dependencia espacio temporal en los niveles máximos de ozono. El ajuste de Modelos Lineales Vectoriales Generalizados (VGLM) confirmó la tendencia del cambio de estructura de la dependencia espacio temporal de los niveles máximos de ozono.

**Palabras clave:** Modelo bivariado max-stable, criterio de Tackeuchy (TIC), análisis de tendencia, verosimilitud compuesta por pares, dependencia extremal.

Lucila Muñiz Merino

Colegio de Postgraduados, 2013

### ABSTRACT

In this research, a methodology based on max-stable process was applied to analyze spatiotemporal trend for environmental data. The maximum ozone levels between 1998 and 2010, in Mexico Valley, were modeled with four max-stable models: Smith's, Schlather's, Gaussian Geometric and Brown-Resnik's. The estimation of the parameters for all models was performed by pairwise-composite maximum likelihood. Then, the best model was selected according to Takeuchi's likelihood information criterion (TIC). The spatiotemporal trend resulted statistically significant when the smoothing parameter of the Schlather's model was held fixed and a linear regression was fitted to its rank. This showed a change in the structure of the spatiotemporal dependence of the maximum ozone levels. Adjustment of Vector Generalized Linear Models (VGLM) confirmed the trend of change in the structure of spatiotemporal dependence of the maximum ozone levels.

**Key words:** bivariate max-stable Model, Takeuchi's criterion (TIC), trend analysis, pairwise-composite likelihood, extreme dependency.

## AGRADECIMIENTOS

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por el apoyo económico brindado durante la realización de mis estudios de postgrado.

Al Colegio de Postgraduados, por haberme brindado la oportunidad de seguir mi formación académica en sus aulas.

A los integrantes de mi Consejo Particular:

Dr. Humberto Vaquera Huerta.

Dra. Elizabeth González Estrada.

Dr. José Aurelio Villaseñor Alva.

Dr. Barry C. Arnold.

A mi hermano:

M.C. Manolo Muñiz Merino

A mis compañeros:

M.C. Enrique Rivera Castillo

M.C. Ivan Hernandez Orzuna

A mis profesores, compañeros de clases y todos aquellos que de alguna u otra manera fueron copartícipes de esta tarea.

## DEDICATORIA

A mis **padres**: Fausto Muñiz Merino† y Rafaela Merino Cortes, gracias por darme la vida, amor, educación y por estar conmigo siempre.

A mis **hermanos y hermanas**: Felipa, Manolo, Claudia, Julia, Martín, Alberto y Dalila, por brindarme su apoyo y confianza.

A mis **sobrinos**: Jasser y Evelyn por mostrarme la felicidad de vivir.

# Índice

<b>1. Introducción</b>	<b>1</b>
<b>2. Objetivos</b>	<b>4</b>
<b>3. Teoría de valores extremos</b>	<b>5</b>
3.1. Extremos univariados . . . . .	5
3.1.1. No estacionariedad de extremos por medio de regresión . . . .	10
3.2. Teoría de valores extremos multivariada . . . . .	12
<b>4. Procesos Max-stable para extremos espaciales</b>	<b>14</b>
4.1. Modelo de Smith . . . . .	15
4.2. Modelo de Schlather . . . . .	16
4.2.1. Modelo Gaussiano Geométrico . . . . .	17
4.3. Modelo de Brown Resnik . . . . .	18
4.4. Medidas de dependencia espacial para campos aleatorios Max-stable	19
4.4.1. Coeficiente extremal . . . . .	19
4.4.2. Enfoque basado en Madograma . . . . .	21
4.5. Ajuste de un proceso Max-stable a datos . . . . .	24

## Índice

---

4.5.1. Mínimos cuadrados . . . . .	24
4.5.2. Verosimilitud . . . . .	25
4.5.3. Selección del modelo . . . . .	26
<b>5. Aplicación</b>	<b>28</b>
5.1. Contaminación por ozono . . . . .	28
5.1.1. Descripción de los datos . . . . .	29
5.2. Metodología . . . . .	32
5.3. Resultados . . . . .	33
5.4. Conclusiones . . . . .	64
<b>A. Instrucciones de R</b>	<b>68</b>
<b>B. Tablas de ajuste de modelos de todos los años</b>	<b>72</b>
<b>C. Resultados de prueba Kolmogorov Smirnov</b>	<b>79</b>



# Índice de tablas

5.1. Ubicación de las estaciones meteorológicas . . . . .	30
5.2. Parámetros del modelo de Smith estimados por mínimos cuadrados para 1998 a 2010 . . . . .	44
5.3. Parámetros del modelo de Schlather con correlación exponencial estimados por mínimos cuadrados para 1998 a 2010 . . . . .	45
5.4. Parámetros del modelo de Schlather con correlación Cauchy estimados por mínimos cuadrados para 1998 a 2010. . . . .	45
5.5. Mejores modelos seleccionados de acuerdo al TIC . . . . .	46
5.6. Mejores modelos de Schlather ajustados para todos los años con correlación exponencial. . . . .	47
5.7. Mejores modelos de Schlather ajustados para todos los años con correlación Cauchy. . . . .	49
5.8. Tendencia con regresión Loess . . . . .	55
5.9. Parámetros de la GEV en función de covariables, longitud, latitud y tiempo . . . . .	59
5.10. Parámetros de la GEV en función de covariables, longitud, latitud, tiempo e interacción longitud-tiempo . . . . .	60
5.11. Parámetros de la GEV en función de covariables, longitud, latitud, tiempo e interacción latitud-tiempo . . . . .	61
B.1. Modelos ajustados para el año 1998 . . . . .	72

## Índice de tablas

---

B.2. Modelos ajustados para el año 1999 . . . . .	73
B.3. Modelos ajustados para el año 2000 . . . . .	73
B.4. Modelos ajustados para el año 2001 . . . . .	74
B.5. Modelos ajustados para el año 2002 . . . . .	74
B.6. Modelos ajustados para el año 2003 . . . . .	75
B.7. Modelos ajustados para el año 2004 . . . . .	75
B.8. Modelos ajustados para el año 2005 . . . . .	76
B.9. Modelos ajustados para el año 2006 . . . . .	76
B.10. Modelos ajustados para el año 2007 . . . . .	77
B.11. Modelos ajustados para el año 2008 . . . . .	77
B.12. Modelos ajustados para el año 2009 . . . . .	78
B.13. Modelos ajustados para el año 2010 . . . . .	78
C.1. Estadísticos de Prueba y P-value. Estaciones: LAG, TAC, SAG, AZC. . . . .	79
C.2. Estadísticos de Prueba y P-value. Estaciones: TLA, XAL, MER, PED. . . . .	80
C.3. Estadísticos de Prueba y P-value. Estaciones: CES, PLA, UIZ, TAX. . . . .	80
C.4. Estadísticos de Prueba y P-value. Estaciones: CUA, TAH . . . . .	81

# Índice de figuras

3.1. Función de densidad y distribución Fréchet . . . . .	8
3.2. Función de densidad y distribución Gumbel . . . . .	9
5.1. Mapa de las estaciones meteorológicas en la Ciudad de México . . . .	31
5.2. Valores observados & los del Modelo ajustado en la estación Lagunilla	34
5.3. Series de tiempo de 5 años para todas las estaciones meteorológicas .	35
5.4. Estimador no paramétrico de Smith de la función extremal para el año 1998 . . . . .	36
5.5. Estimador no paramétrico de Schlatner y Tawn de la función extre- mal para el año 1998 . . . . .	37
5.6. Estimador no paramétrico de Smith de la función extremal para el año 2010 . . . . .	38
5.7. Estimador no paramétrico Schlather y Tawn de la función extremal para el año 2010 . . . . .	39
5.8. Función extremal estimada con el f-madograma, para el año 1998 . .	40
5.9. Función extremal estimada con el f-madograma, para el año 2010 . .	41
5.10. Función extremal estimada con el lambda madograma, para el año 1998. . . . .	42
5.11. Función extremal estimada con el lambda madograma, para el año 2010. . . . .	43

## Índice de figuras

---

5.12. Coeficiente extremal con parámetros estimados por máxima verosimilitud para el año 1998. . . . .	48
5.13. Coeficiente extremal con parámetros estimados por máxima verosimilitud para el año 2010. . . . .	50
5.14. Tendencia en el rango de la correlación exponencial . . . . .	51
5.15. Tendencia en el rango de la correlación Cauchy . . . . .	52
5.16. Tendencia en la dependencia extremal de Schlather con corr. exponencial . . . . .	53
5.17. Tendencia en la dependencia extremal de Schlather con corr. cauchy. . . . .	54
5.18. Tendencia del coeficiente extremal a diferentes distancias en el tiempo . . . . .	56
5.19. Tendencia del coeficiente extremal a diferentes distancias en el tiempo . . . . .	57
5.20. Tendencia del coeficiente extremal a diferentes distancias en el tiempo . . . . .	58
5.21. Kriging de los niveles de retorno para el año 1998 . . . . .	62
5.22. Kriging de los niveles de retorno para el año 2010 . . . . .	63

# Capítulo 1

## Introducción

En la actualidad, el cambio climático es un tema de interés debido a sus efectos sobre la sociedad, economía y medio ambiente. Algunos efectos del clima a nivel mundial según reporte de la revista *Tierramérica* [Constanzo \(2009\)](#), elaborado a partir de las opiniones de varios investigadores, son: sequías, inundaciones, derretimiento de glaciares, aumentos de temperatura, plagas agrícolas y enfermedades entre otros. Estos efectos en el medio ambiente son derivados de la variación en frecuencia e intensidad de los eventos climáticos extremos, ([Min et al. \(2011\)](#), [Katz and Brown \(1992\)](#)).

Los sistemas naturales están siendo afectados por cambios de clima regional, mientras que algunos eventos han tendido a incrementarse a través del tiempo y a suceder en prácticamente todo el planeta, otros han disminuido y tienden a presentarse en algunas regiones solamente. Como ejemplos están los incrementos: en la temperatura del planeta, el nivel del mar asociado, de precipitaciones en algunas regiones, en superficies de sequías; así como la disminución en: las extensiones de hielo y nieve, y el nivel de precipitaciones en algunas regiones. Esto según el IV informe del grupo Intergubernamental de Expertos sobre el cambio climático [IPCC \(2007\)](#), (IPCC por sus siglas en inglés). Dicho informe establece que el clima puede variar a escalas espaciales y temporales muy diversas. Las escalas espaciales abarcan desde extensiones locales hasta regionales o continentales y las temporales desde periodos estacionales hasta geológicos.

Las causas del desequilibrio energético en el sistema climático están fuertemente relacionadas a variaciones en la cubierta terrestre, aerosoles en la atmósfera, radiación solar y a las variaciones de concentración de los Gases efecto Invernadero (GEI), estos últimos según el IV informe del IPCC aumentaron en un 70 % entre 1970 y 2004, su incremento en la atmósfera es de aproximadamente 2.5 ppm y su concentración anual gira en torno a 370 ppm, la mayor parte de ellos proviene del suministro de energía, transporte e industria [IPCC \(2007\)](#).

## 1. Introducción

---

Entre los gases de efecto invernadero se encuentra el ozono urbano ( $O_3$ ) [Constanzo \(2009\)](#) el cual provoca daños al medio ambiente y en la salud. En el caso de las plantas, pierden sus hojas, mientras que los seres humanos se ven afectados por enfermedades cardiovasculares. El comportamiento y formación del ozono es complejo, éste no es emitido a la atmósfera directamente, sino que es producido por la interacción meteorológica de óxidos de nitrógeno ( $NO_x$ ) y de compuestos orgánicos volátiles (COVs), algunos estudios muestran que su comportamiento se relaciona directamente con la temperatura, humedad y velocidad del viento.

El ozono es un contaminante importante en las grandes ciudades. Tan solo en la zona Metropolitana de la Ciudad de México, diversos estudios se han enfocado a su análisis, algunos de los cuales muestran que éste presenta tendencia decreciente a través del tiempo y variación en observaciones horarias las cuales incrementan en el transcurso de la mañana alcanzando sus valores máximos entre 1 y 4 de la tarde ([Huerta et al. \(2004\)](#), [Reyes et al. \(2010\)](#)).

Existen dificultades significativas en predecir los niveles de ozono en el ambiente, así como su distribución espacial, comportamiento y tendencia asociada ([Correa et al. \(2000\)](#)). Un inconveniente con este tipo de estudios es que no toman en cuenta la dependencia espacial, incluso los parámetros de la distribución de valores extremos generalizada (GEV) son modelados como una función de localizaciones espaciales y otras covariables para analizar el comportamiento espacial de fenómenos tales como precipitaciones, velocidad del viento, temperaturas extremas, etcetera. Sin embargo, la distribución de valores extremos generalizada (GEV) asume independencia espacial del proceso condicional sobre los parámetros del modelo y esto puede conducir a inferencia y predicción espacial no realista.

Una manera de solucionar este inconveniente es ajustando modelos de extremos espaciales, la literatura presenta un enfoque de muy reciente creación ([de Haan \(1984\)](#)) denominado procesos continuos max-stable, estos proveen un enfoque usual para modelar procesos extremos incorporando la dependencia espacial y temporal. La tendencia espacial y temporal es capturada en los parámetros de la distribución GEV como superficies de respuesta. Varios modelos para procesos max-stable han sido propuestos: [Smith \(1990\)](#) propuso un modelo max-stable con formas determinísticas, [Schlather \(2002\)](#) brindó un modelo basado en un proceso gaussiano, otros modelos planteados incluyen el proceso Brown Resnik [Kabluchko et al. \(2009\)](#). De estas construcciones generales se derivan distintos modelos como casos especiales.

Para el ajuste del modelo [Smith \(1990\)](#) trata un método basado en mínimos cuadrados, [Padoan et al. \(2010\)](#) sugiere estimación de máxima verosimilitud compuesta por pares, recientemente [Genton et al. \(2011\)](#) extiende este enfoque usando verosimilitud compuesta basada en densidades trivariadas. Diferentes estimadores para la función extremal, la cual mide la dependencia entre extremos en diferentes localizaciones, han sido propuestos por: [Smith \(1990\)](#), [Schlather and Tawn \(2003\)](#) y estimadores de madograma por [Cooley et al. \(2006\)](#) y [Naveau et al. \(2009\)](#).

## 1. Introducción

---

Existen en la literatura algunos otros métodos para el modelado espacial de extremos como los modelos de regresión sobre los parámetros de la GEV, entre los que se encuentran los Modelos Lineales Vectoriales Generalizados (VGLM) los cuales son un caso especial de los Modelos Aditivos Vectoriales Generalizados (VGAM) expuestos por [Yee and Stephenson \(2007\)](#) en los que modelan los parámetros de la distribución de valores extremos generalizada en funciones lineales de covariables.

En este trabajo se utiliza el proceso max-stable y un Modelo Lineal Generalizado Vectorial (VGLM) para modelar niveles máximos de un contaminante del aire (ozono), e investigar el comportamiento espacio temporal del ozono urbano en el Valle de México.

El trabajo está organizado como sigue: el capítulo 1 es la introducción de este trabajo, en el capítulo 2 se presentan los objetivos de esta investigación, en el capítulo 3 se incluyen las herramientas básicas de la teoría de valores extremos, en el capítulo 4 se describen los modelos max-stable más importantes y de los cuales se derivan algunos otros, así como las herramientas para llevar a cabo la inferencia y los ajustes de modelos max-stable. En el capítulo 5 se presenta un análisis del comportamiento espacio temporal del ozono urbano en el Valle de México. Finalmente se exponen las conclusiones de este análisis.

# Capítulo 2

## Objetivos

### Objetivo general

- Aplicar una metodología estadística basada en el proceso max-stable para analizar tendencia espacio temporal con datos ambientales.

### Objetivos particulares

- Modelar la dependencia extremal espacial por medio de los modelos de Smith, de Schlather, Gaussiano Geométrico y de Brown Resnik.
- Analizar la dependencia extremal espacial por medio de los estimadores no paramétricos de Smith y de Schlather y Tawn.
- Analizar la dependencia extremal espacial por medio de los estimadores de madograma, f-madograma y lambda madograma.
- Analizar la tendencia espacio temporal por medio del rango de las funciones de correlación exponencial y Cauchy con el modelo de Schlather, al fijar el parámetro de suavizamiento.
- Validar la metodología empleada usando el modelo VGLM.



# Capítulo 3

## Teoría de valores extremos

El objetivo de la teoría de valores extremos es la extrapolación de información, es decir, teniendo una muestra de una distribución desconocida se quiere estimar la cola de la misma, esto es, el interés se centra en los valores más altos(máximos) o más bajos(mínimos) de las variables en estudio.

Esta teoría tiene aplicaciones en áreas muy diversas como en hidrología, finanzas, oceanografía, el cálculo del valor en riesgo en finanzas; en problemas relacionados con la contaminación ambiental el interés recae en los valores más grandes aquellos que sobrepasan un nivel crítico como es el caso de los valores máximos de ozono.

Los problemas que se tienen en esta teoría son que se pueden tener pocas observaciones en la cola de la distribución.

La teoría de valores extremos se enfoca al estudio de eventos como los máximos y mínimos, algunos resultados importantes han sido derivados para el modelado de los mismos.

### 3.1. Extremos univariados

Un resultado importante en la teoría de valores extremos es el teorema de tipos de extremos, el cual se debe a Fisher y Tippett en 1928 quienes mostraron que las distribuciones límite extremas son únicamente de tres tipos. En 1943 Gnedenko presentó el fundamento riguroso de la teoría de valores extremos. Gnedenko basó los fundamentos matemáticos de la teoría de los valores extremos en la clase de leyes límites de extremos. Von Mises y Jenkinson ([Jenkinson \(1955\)](#), [Von Mises](#)

### 3.1. Extremos univariados

---

(1954)) unificaron a las distribuciones de límites Fréchet, Weibull y Gumbel en una sola distribución con parametrización común la Distribución de Valor Extremo Generalizada (DGVE), debido a esto al teorema de tipos de extremos se le conoce como teorema de Fisher y Tippet y a la (DGVE) como distribución de Von Mises Jenkinson.

El teorema de tipos de extremos establece que para una sucesión de variables aleatorias independientes e idénticamente distribuidas con constantes normalizadas adecuadas, la distribución límite de estos máximos, si existe, sigue una distribución de valor extremo generalizada, esto es:

Sea la sucesión de v.a.'s  $X_1, X_2, \dots$  independientes con función de distribución común  $F(x)$  y máximo  $M_n = \max\{X_1, \dots, X_n\}$  con distribución respectiva

$$P(M_n \leq z) = P(X_1 \leq z, \dots, X_n \leq z) = P(X_1 \leq z) \cdots P(X_n \leq z) = \prod_{i=1}^n P(X_i \leq z) = \{F(z)\}^n$$

En la mayoría de aplicaciones de modelos de valores extremos ambas distribuciones se desconocen. Sin embargo, por la teoría asintótica cuando  $n$  crece, el máximo  $M_n$  se aproxima al punto superior final  $x_F = \sup\{x \in \mathbb{R} : F(x) < 1\}$  de  $F(x)$  y la distribución límite de  $M_n$  es un punto de masa en  $x_F$ . Para obtener un límite no degenerado, de manera similar como en el teorema del límite central, se aplica una normalización lineal, en el límite cuando  $n$  tiende a infinito y para sucesiones de constantes apropiadas  $\{a_n > 0\}$  y  $\{b_n \in \mathbb{R}\}$  se busca la función de distribución del máximo  $G(x)$ .

$$Pr\left(\frac{M_n - b_n}{a_n} \leq x\right) = F^n(a_n x + b_n) \approx G(x) \quad (3.1)$$

Tal distribución se dice que cumple la propiedad de ser max-stable, y de ser no degenerada, puede además decirse que  $F(x)$  se encuentra en el máximo dominio de atracción de  $G(x)$ .

El teorema de tipos de extremos de Fisher y Tippet enunciado en Coles (2001) caracteriza todas las distribuciones límite posibles:

**Teorema 1:** Si existen sucesiones de constantes  $\{a_n > 0\}$  y  $\{b_n\}$  tal que

$$Pr\left(\frac{M_n - b_n}{a_n} \leq x\right) \rightarrow G(x) \quad (3.2)$$

cuando  $n \rightarrow \infty$ , donde  $G$  es una función de distribución no degenerada, en tal caso  $G$  es una de las tres familias

### 3.1. Extremos univariados

$$\text{Gumbel} : G(x) = \exp\{-\exp[-(\frac{x-\mu}{\lambda})]\} \quad -\infty < x < \infty \quad (3.3)$$

$$\text{Fréchet} : G(x) = \begin{cases} 0, & \text{si } x \leq \mu \\ \exp\{-(\frac{x-\mu}{\lambda})^{-\xi}\}, & \text{si } x > \mu \end{cases} \quad (3.4)$$

$$\text{Weibull} : G(x) = \begin{cases} \exp\{-[(\frac{x-\mu}{\lambda})^\xi]\}, & \text{si } x < \mu, \\ 1, & \text{si } x \leq \mu \end{cases} \quad (3.5)$$

Las tres familias se combinan en una sola distribución, la de valor extremo generalizada

$$G(x) = \exp\{-[1 + \xi(\frac{x-\mu}{\lambda})]^{-1/\xi}\} \quad (3.6)$$

definida sobre  $\{x : 1 + \xi(x - \mu)/\lambda > 0\}$ , donde  $-\infty < \mu < \infty$ ,  $\lambda > 0$  y  $-\infty < \xi < \infty$  con función de densidad:

$$g(x) = \lambda\{-[1 + \xi(\frac{x-\mu}{\lambda})]^{-1/\xi-1}\}\exp\{-[1 + \xi(\frac{x-\mu}{\lambda})]^{-1/\xi}\} \quad (3.7)$$

Siempre que  $\{x : 1 + \xi(\frac{x-\mu}{\lambda})\} > 0$

Esta distribución es la única con la propiedad de ser max-stable, esto es para todo  $n = 2, 3, \dots$ , existen constantes  $\{a_n > 0\}$  y  $\{b_n\}$  tal que  $G^n(a_n x + b_n) = G(x)$ .

La (DGVE) provee soporte para modelar máximos de bloques. De sus parámetros  $\mu, \lambda$  y  $\xi$ , el de forma determina la cola superior de la densidad, cuanto mayor sea el valor más gruesa será la cola, esto es para  $\xi < 0$  la distribución es una Weibull con cola acotada, para  $\xi = 0$  una Gumbel con cola moderada y no acotada y  $\xi > 0$  corresponde a una Fréchet con cola pesada y momentos infinitos para todos los ordenes más grandes que  $1/\xi$  (por ejemplo la varianza no existe para  $\xi > 0.5$ , y la media no existe para  $\xi > 1$ ).

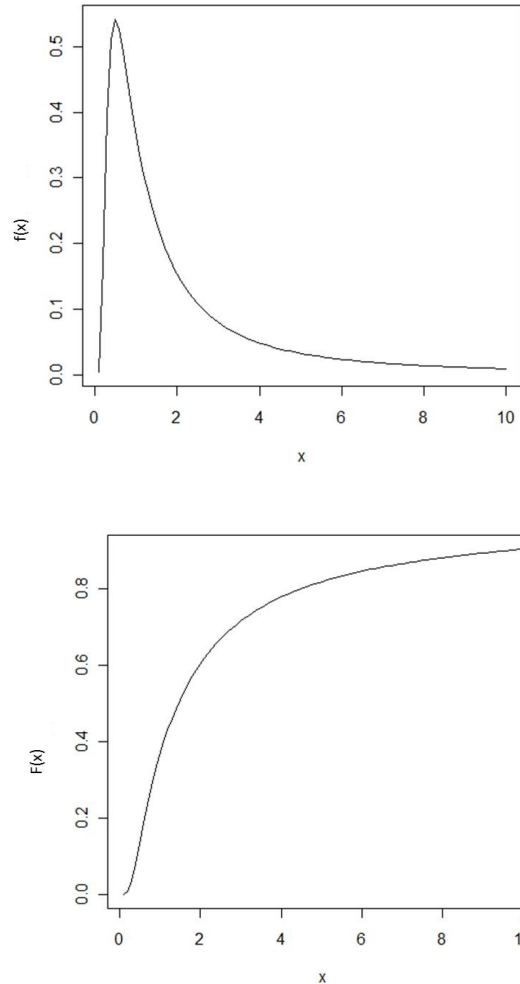
La distribución Fréchet en la estadística multivariada y espacial con procesos max-stable juega un importante rol, pues es más simple desarrollar expresiones matemáticas equivalentes que resultan de la variable transformada Davison et al. (2012),  $Y = \{1 + \frac{\xi(x-\mu)}{\lambda}\}^{-1/\xi}$ , con función de distribución unitaria Fréchet,  $\exp(-1/y)$ ,  $y > 0$ . En este caso la propiedad de max estabilidad es  $mY \stackrel{D}{=} \max(Y_1, \dots, Y_m)$ , donde  $Y, Y_1, \dots, Y_m$  representan mutuamente variables aleatorias independientes unitarias Fréchet y  $\stackrel{D}{=}$  denota igualdad en distribución. Distribuciones para las que la Fréchet es límite son la Cauchy y Pareto. Ésta es una distribución muy utilizada en aplicaciones financieras, la Fig. 3.1 muestra la densidad y distribución de la

### 3.1. Extremos univariados

---

misma.

Figura 3.1: Función de densidad y distribución Fréchet



El modelo Gumbel es utilizado para análisis de eventos extremos en precipitación, finanzas, ciencias ambientales, para predecir terremotos, etc. Una simplificación a la familia Gumbel se realiza con el argumento de que hay muchas distribuciones para los  $X_i$  para los cuales la distribución Gumbel es un límite apropiado para  $M_n$ , ejemplos de tales distribuciones son la normal, lognormal, exponencial y gamma. La reducción a esta familia puede ser aceptada por una prueba de razón de verosimilitud.

El contraste de hipótesis es:

$H_0$ : El modelo es Gumbel, es decir  $\xi = 0$ .

### 3.1. Extremos univariados

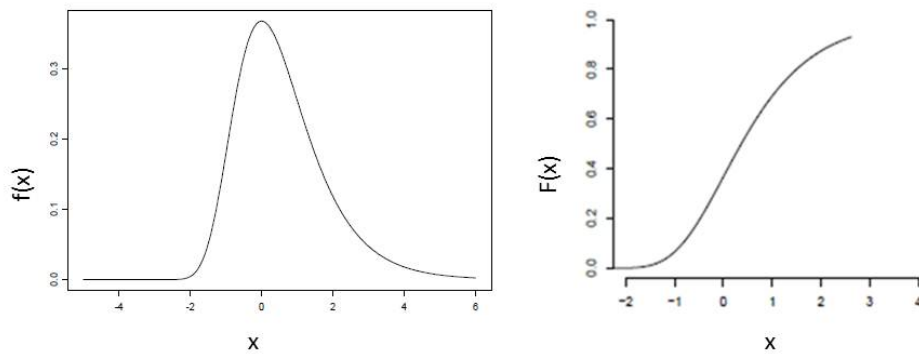
---

$H_1$ : El modelo es GEV, es decir  $\xi \neq 0$ .

se utiliza la estadística de prueba  $D = 2(\log(\text{ModeloGEV}) - \log(\text{ModeloGumbel})) \sim X^2_{p_{GEV}-p_{Gumbel}}$ . La regla de decisión es rechazar la hipótesis nula si D es mayor que un valor de tablas de la distribución chi cuadrada  $X^2_{p_{GEV}-p_{Gumbel}}$ , donde p es el número de parámetros de cada modelo.

En la Fig. 3.2 se presentan la densidad y distribución de la Gumbel.

**Figura 3.2:** Función de densidad y distribución Gumbel



La Weibull es límite de la distribución uniforme, ésta modela tasas de fallo en sistemas y es utilizada para análisis de fiabilidad.

### Estimación de parámetros

Los parámetros de la distribución GEV pueden obtenerse por el método de máxima verosimilitud o por métodos no paramétricos. El logaritmo de la función de verosimilitud es el siguiente:

### 3.1. Extremos univariados

---

$$\ell = -n \log \lambda - \left(\frac{1}{\xi} + 1\right) \sum_{i=1}^n \log \left[1 + \xi \left(\frac{x_i - \mu}{\lambda}\right)\right] - \sum_{i=1}^n \left[1 + \xi \left(\frac{x_i - \mu}{\lambda}\right)\right]^{-\frac{1}{\xi}} \quad (3.8)$$

En el método de máxima verosimilitud se establece que los parámetros  $(\hat{\mu}, \hat{\lambda}, \hat{\xi})$  se deben escoger de tal forma que maximicen  $\ell$ .

la log-verosimilitud para estimar los parámetros de la Gumbel es:

$$\ell = -n \log \lambda - \sum_{i=1}^n \left(\frac{x_i - \mu}{\lambda}\right) - \sum_{i=1}^n \exp\left(-\frac{x_i - \mu}{\lambda}\right)$$

El cálculo de la estimación de parámetros puede hacerse numéricamente mediante computadora, con el paquete en ismev de R, el cual estima los parámetros por máxima verosimilitud.

## Estimación de cuantiles

Un objetivo en aplicaciones es la estimación del cuantil de la distribución de Z, es de interés, por ejemplo, para la prevención de inundaciones la cantidad de lluvia en un área dada en un día que excede una vez en 100 años, o el cuantil 100-anual del total de lluvia en un área en un día dado, es decir el nivel de retorno  $z_p$  asociado con el periodo de retorno  $1/p$ , el cual se obtiene de invertir la ecuación de la función de distribución GEV. Definiendo  $x_p = -\log(1 - p)$ , se tiene

$$z_p = \begin{cases} \mu - \frac{\lambda}{\xi} [1 - x_p^{-\xi}], & \text{para } \xi \neq 0, \\ \mu - \lambda \log(x_p), & \text{para } \xi = 0, \end{cases} \quad (3.9)$$

$z_p$  es el nivel de retorno y si las observaciones son máximos anuales y T representa los años, entonces se espera que  $z_p$  nivel de retorno T anual sea excedido en promedio una vez cada  $T=1/p$  años, es decir,  $z_p$  es excedido por el máximo anual en cualquier año en particular con probabilidad p.

### 3.1.1. No estacionariedad de extremos por medio de regresión

Para reducir el número de parámetros se consideran superficies de respuesta sobre los parámetros de la GEV, lo que en extremos univariados se conoce como no estacionariedad, esta superficie puede ser modelada mediante una tendencia lineal (forma paramétrica) o por medio de funciones de suavizamiento (forma no paramétrica). Por ejemplo, una tendencia lineal en cada uno de los parámetros

### 3.1. Extremos univariados

de la distribución de valor extremo generalizada se estimaría suponiendo que se tienen  $n$  máximos muestrales  $y_1 \cdots y_n$  con correspondientes vectores de covariables  $x_1, \cdots, x_n$ , los modelos de regresión general toman la forma:

$$y_i|x_i \text{ GEV}(\mu(x_i), \lambda(x_i), \xi(x_i))$$

donde,  $\mu(x_i) = g(X\beta)_i$ ,  $g$  es una función liga,  $\beta$  es un vector de coeficientes de regresión y  $X$  es una matriz diseño asociada con los  $x_i$ 's. Estructuras similares pueden ser impuestas sobre  $\lambda(x_i)$  y  $\xi(x_i)$ .

La log verosimilitud para estos modelos es:

$$\ell = - \sum_{i=1}^n \{ \log \lambda(x_i) + (1 + 1/\xi(x_i)) \log [1 + \xi(x_i) (\frac{y_{x_i} - \mu(x_i)}{\lambda(x_i)})] + [1 + \xi(x_i) (\frac{y_{x_i} - \mu(x_i)}{\lambda(x_i)})]^{-1/\xi(x_i)} \}$$

Uno de los modelos tipo regresión son los VGAM tratados por [Yee and Stephenson \(2007\)](#), éstos incluyen a los Lineales Vectoriales Generalizados (VGLM) con los cuales es posible determinar si existe relación entre los parámetros de la distribución GEV con covariables, en este caso longitud, latitud y tiempo. Los VGAM y VGLM permiten modelar los parámetros de la función GEV como funciones lineales o de suavizamiento de sus covariables.

Suponga que la respuesta observada  $\mathbf{y}$  es un vector  $q$ -dimensional. Los VGLMs son definidos como un modelo para el cual la distribución condicional de  $\mathbf{Y}$  dada la explicativa  $\mathbf{x}$  es de la forma:

$$f(\mathbf{y}|\mathbf{x}; \mathbf{B}) = h(\mathbf{y}, \eta_1, \dots, \eta_M)$$

para alguna función conocida  $h(\cdot)$ , donde  $\mathbf{B} = (\beta_1, \beta_2, \dots, \beta_M)$  es una matriz de orden  $p \times M$  de los coeficientes de regresión desconocidos, y el  $j$ -ésimo predictor lineal es  $\eta_j = \eta_j(\mathbf{x}) = \beta_j^T \mathbf{x} = \beta_{(j)1}x_1 + \cdots + \beta_{(j)p}x_p = \sum_{k=1}^p \beta_{(j)k}x_k$ ,  $j = 1, \dots, M$

donde  $\mathbf{x} = (x_1, \dots, x_p)^T$  con  $x_1 = 1$  si hay intercepto. La ecuación anterior indica que todos los parámetros pueden ser potencialmente modelados como funciones de las covariables  $\mathbf{x}$ , en este caso las covariables corresponden a longitud, latitud y tiempo.

Los VGAMs proporcionan extensiones de modelos aditivos a los VGLMs, es decir, ahora la ecuación es:

$$\eta_j(\mathbf{x}) = \beta_{(j)1} + f_{(j)2}(x_2) + \cdots + f_{(j)p}(x_p) = \beta_{(j)1} + \sum_{k=2}^p f_{(j)k}(x_k), \quad j = 1, \dots, M$$

una suma de funciones suavizadas de las covariables individuales. Las  $\eta_j$  son referidos como predictores aditivos.  $\mathbf{f}_k = (f_{(1)k}(x_k), \dots, f_{(M)k}(x_k))$  es centrada por unicidad, y  $f$ 's se estiman de manera simultánea utilizando vectores de suavizamiento.

## 3.2. Teoría de valores extremos multivariada

En la práctica, tal vez se quiera limitar el efecto de una covariable a que sea el mismo para algunos de los  $\eta_j$  y no tenga efecto en otras. Por ejemplo, en los VGAMs se quisiera tener:

$$\begin{aligned}\eta_1 &= \beta_{(1)1} + f_{(1)2}(x_2) + f_{(1)3}(x_3) \\ \eta_2 &= \beta_{(2)1} + f_{(1)2}(x_2),\end{aligned}$$

de modo que  $f_{(1)2} \equiv f_{(2)2}$  y  $f_{(2)3} \equiv 0$ . Para los VGAMs, podemos representar estos modelos usando:

$$\eta(\mathbf{x}) = \beta_{(1)} + \sum_{k=2}^p f_k(x_k) = \mathbf{H}_1 \beta_{(1)}^* + \sum_{k=2}^p \mathbf{H}_k f_k^*(x_k) \quad (3.10)$$

donde  $\mathbf{H}_1, \mathbf{H}_2, \dots, \mathbf{H}_p$  se conocen como matrices de restricción de rango completo por columnas,  $f_k^*$  es un vector que contiene al posible conjunto reducido de funciones de los componentes y  $\beta_{(1)}^*$  es el vector de interceptos desconocido. Cuando no hay restricciones,  $\mathbf{H}_1 = \mathbf{H}_2 = \dots = \mathbf{H}_p = \mathbf{I}_M$  y  $\beta_{(1)}^* = \beta_{(1)}$ . Al igual que  $f_k, f_k^*$  es centrada por unicidad. En los VGLMs,  $f_k$  es lineal, por lo tanto

$$\mathbf{B}^T = \left( \mathbf{H}_1 \beta_{(1)}^* \quad \mathbf{H}_2 \beta_{(2)}^* \quad \dots \quad \mathbf{H}_p \beta_{(p)}^* \right)$$

Los VGLMs son estimados usualmente por máxima verosimilitud, por Fisher Scoring o Newton-Rapson.

## 3.2. Teoría de valores extremos multivariada

La teoría de valores extremos multivariada, fue creada para estudiar la distribución conjunta en varias series, se aplica para análisis de supervivencia, estimación de dependencia en extremos de altura de olas, contaminación ambiental, etc.

Para la distribución multivariada de extremos de forma equivalente con la distribución univariada, si  $(X_{i,1}, \dots, X_{i,p}), i = 1, \dots, n$  son vectores aleatorios p-variados iid con distribución conjunta común F y

$$M_n = (M_{n,1}, \dots, M_{n,p}) = (\max_{1 \leq i \leq n} X_{i,1}, \dots, \max_{1 \leq i \leq n} X_{i,p})$$

es el vector de máximos de cada componente, entonces por definición:

**Definición:** si existen sucesiones de constantes  $a(n, j) > 0, b(n, j), j = 1, \dots, p$  tal que cuando  $n \rightarrow \infty$ ,

$$P((M_{n,1} - b_{n,1})/a_{n,1} \leq x_1, \dots, (M_{n,p} - b_{n,p})/a_{n,p} \leq x_p) = F^n(a_{n,1}x_1 + b_{n,1}, \dots, a_{n,p}x_p + b_{n,p}) \rightarrow$$



### 3.2. Teoría de valores extremos multivariada

---

$$G(x_1, \dots, x_p)$$

para alguna distribución p-variada G no degenerada, la distribución G es llamada una distribución max-stable multivariada p-dimensional.

Dos características importantes de esta distribución son:

1. Cada distribución marginal de extremos multivariada sigue una de las tres distribuciones de valor extremo univariadas (Gumbel, Fréchet, Weibull).
2. Para cualesquiera funciones marginales monótonas crecientes, el máximo de las marginales es igual a la función del máximo para cada vector de variables aleatorias.

La distribución límite como representación de Pickands 1981 se escribe como:  $G(x) = \exp\{-V(x)\}$ ,  $V(x)$  es llamada la medida exponente y es la integral:

$$V(x) = \int_{S_p} \max\left(\frac{w_1}{x_1}, \dots, \frac{w_p}{x_p}\right) dH(w)$$

donde  $S_p = \{(w_1, \dots, w_p) : w_j \geq 0, \sum_{j=1}^p w_j \leq 1\}$ , para cualquier  $j \in \{1, \dots, p\}$

$V(\infty, \infty, \dots, j, \dots, \infty) = \int_{S_p} \max\left(\frac{w_1}{x_1}, \dots, \frac{w_p}{x_p}\right) dH(w) = \frac{1}{x_j} \int_j dH(w)$ , para asegurar que las distribuciones marginales son Fréchet unitarias,  $H(w)$  tiene que satisfacer  $\int_{S_p} w_j dH(w) = 1$  para cada  $j$  de 1 a  $p$ . A la densidad correspondiente  $h(w)$  se le denomina densidad espectral. A diferencia de los extremos univariados, no hay forma paramétrica simple para la distribución límite multivariada,  $V$  puede tomar cualquier forma.

Estas distribuciones tienen varias limitaciones: algunas de ellas son que están sobrep parametrizadas por lo que a lo que inferencia estadística se refiere existe problema en la derivación de la verosimilitud, debido a que la suma de términos se vuelve grande dependiendo del número de variables; la densidad espectral en algunas también esta sobrep parametrizada por lo que existen inconvenientes en la integración para obtener la medida exponente. La integral para obtener la medida exponente es difícil de evaluar, lo cual es una desventaja; y respecto al contexto espacial la estructura de dependencia de una distribución multivariada de extremos no está relacionada con la dependencia entre sitios.

En [Kotz and Nadarajah \(2000\)](#) se tratan algunas distribuciones multivariadas, y [Smith \(1990\)](#) provee ejemplos de algunas familias de distribuciones como casos especiales de un proceso max-stable.

# Capítulo 4

## Procesos Max-stable para extremos espaciales

Un objetivo común del análisis espacial es cuantificar y caracterizar el comportamiento de fenómenos tales como niveles de precipitación, velocidad del viento, temperaturas diarias. Se ha desarrollado toda una teoría para el modelado espacial, aunque, ésta no es necesariamente ideal para aspectos extremos dado que su enfoque es sobre el nivel medio del proceso. Sin embargo, una teoría que ha emergido como una herramienta para el modelado espacial de valores extremos son los procesos estocásticos estacionarios max-stable ([de Haan \(1984\)](#), [de Haan and Pickands \(1986\)](#)), que cumplen una propiedad importante para el modelado espacial, la dependencia de la cola entre las variables localizadas en el plano decrece monótonicamente y continuamente con la distancia, además tienen la estructura de dependencia de las variables aleatorias representada por los parámetros del modelo. Una dificultad que presentan es obtener distribuciones marginales para dimensiones mayores que dos, lo único con lo que se cuenta es con la forma cerrada bivariada, esto representa una complejidad para la inferencia. Sin embargo, un procedimiento de verosimilitud compuesta por pares ha sido validado por simulación para el modelo [Smith \(1990\)](#) por [Padoan et al. \(2010\)](#).

[de Haan \(1984\)](#) define un proceso max-stable de la siguiente forma:

**Definición:** Sea  $X$  un conjunto de índices y sea  $\{Y_s(x)\}_{x \in X, s = 1, \dots, n}$  son  $n$  replicas independientes de un proceso estocástico continuo. Asumiendo que existen sucesiones de funciones continuas  $a_n(x) > 0$  y  $b_n(x) \in \mathbb{R}$ , tales que

$$Y(x) = \lim_{n \rightarrow \infty} \frac{\max_{s=1, \dots, n} Y_s(x) - b_n(x)}{a_n(x)}, x \in X$$

## 4.1. Modelo de Smith

---

Si este límite existe, el proceso  $Y(x)$  es un proceso max-stable.

La formulación max-stable puede verse como una extensión de la propiedad de máxima estabilidad de la GEV a un proceso continuo. Dos propiedades se siguen de la definición (de [Haan and Resnick \(1977\)](#)):

1. Las distribuciones marginales unidimensionales son de la clase de distribuciones de valor extremo generalizada (Gumbel, Weibull, Fréchet).
2. Para cualquier  $k = 2, 3, \dots$ , las distribuciones marginales  $k$ -dimensionales son de la clase de distribuciones de valor extremo multivariadas.

Sin pérdida de generalidad si,  $Y(x)$  se asume estacionario y  $a_n(x) = n$  y  $b_n(x) = 0 \forall x$ , entonces el proceso correspondiente tiene marginales unitarias Fréchet.

Considerando la representación espectral de un proceso max-stable,  $\Pi$  es un proceso Poisson sobre  $R_+$  de intensidad  $s^{-2}ds$  y  $\{Y_s(x) : x \in X, s \in R_+\}$  es una colección de procesos aleatorios no negativos independientes e idénticamente distribuidos con medida  $\nu$  sobre  $R_+^x$  satisface que  $E\{Y_s(x)\} = 1$  para todo  $x$ . [Schlather \(2002\)](#) muestra que

$$Z(x) = \max_{s \in \Pi} s Y_s(x), x \in X \quad (4.1)$$

es max-stable con distribuciones marginales unitarias Fréchet.

Enseguida son expuestos algunos modelos de procesos max-stable en la literatura como son el de [Smith \(1990\)](#), el de [Schlather \(2002\)](#), el gaussiano geométrico, así como el de Brown Resnik ([Brown and Resnik \(1977\)](#); [Kablichko et al. \(2009\)](#)).

## 4.1. Modelo de Smith

El modelo de tormentas aleatorias fue considerado primero por [Smith \(1990\)](#). Más recientemente [Padoan et al. \(2010\)](#) aplicó el enfoque de verosimilitud compuesta en esta formulación para modelar extremos espaciales. En este ajuste, sea  $X = R^d$  y  $Y_s(x) = f(x - X_s)$ , donde  $f$  es una función de densidad sobre  $X$  y  $X_s$  un punto de un proceso Poisson de tasa unitaria en  $X$ .

La interpretación física de este proceso es: si  $s$  es la magnitud de la tormenta en la localización  $x$ , centrada en  $X_s$  y de forma  $f$ , entonces  $sY_s(x)$  es el impacto de la

## 4.2. Modelo de Schlather

tormenta en  $X$ , por tanto  $Z(x)$  es el impacto de la tormenta más grande observada en  $X$ .

[Smith \(1990\)](#) mostró que si  $f$  es seleccionada a que sea una normal multivariada con matriz de covarianza  $\Sigma$ , la distribución marginal bivariada es:

$$Pr\{Z(x_i) \leq z_i, Z(x_j) \leq z_j\} = \exp\left\{-\frac{1}{z_i}\Phi\left(\frac{a(h)}{2} + \frac{1}{a(h)}\log\frac{z_j}{z_i}\right) - \frac{1}{z_j}\Phi\left(\frac{a(h)}{2} + \frac{1}{a(h)}\log\frac{z_i}{z_j}\right)\right\} \quad (4.2)$$

donde  $\Phi$  es la distribución normal estándar y  $\theta(h) = 2\Phi(\sqrt{h^T \Sigma^{-1} h}/2)$  es el coeficiente extremal por pares.  $a(h)^2 = h^T \Sigma^{-1} h$  es la distancia de Mahalanobis, los casos límite son cuando  $a \rightarrow 0^+$  y  $a \rightarrow +\infty$  corresponden respectivamente a dependencia perfecta  $\theta(h) = 1$ , e independencia  $\theta(h) = 2$ .

La matriz de varianzas covarianzas para este modelo es:

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}.$$

Por tomar las derivadas parciales de segundo orden la función de densidad bivariada puede ser obtenida de la forma:

$$f(z_i, z_j) = \exp\left\{-\frac{\Phi(w(h))}{z_i} - \frac{\Phi(v(h))}{z_j}\right\} \left\{ \left( \frac{\Phi(v(h))}{z_i^2} + \frac{\varphi(w(h))}{a(h)z_i^2} - \frac{\varphi(v(h))}{a(h)z_i z_j} \right) \left( \frac{\Phi(v(h))}{z_j^2} + \frac{\varphi(v(h))}{a(h)z_j^2} - \frac{\varphi(w(h))}{a(h)z_i z_j} \right) + \left( \frac{v(h)\varphi(w(h))}{a(h)^2 z_i^2 z_j} + \frac{w(h)\varphi(v(h))}{a(h)^2 z_i z_j^2} \right) \right\}$$

$\varphi$  es la función de densidad Gaussiana estándar.  $w(h) = a(h)/2 + \log(z_j/z_i)/a(h)$  y  $v(h) = a(h) - w(h)$

Más modelos considerando la distribución t de Student y la Exponencial fueron derivados por [de Haan and Pereyra \(2006\)](#), quienes también estimaron los parámetros de forma no paramétrica.

## 4.2. Modelo de Schlather

El proceso aleatorio sugerido por [Schlather \(2002\)](#) considerando  $Y_s(x) = \max\{0, \sqrt{2\Pi}\xi(x)\}$ , donde  $\{\xi(x)\}$  es un proceso Gaussiano estacionario con media cero, varianza uni-

## 4.2. Modelo de Schlather

taria y función de correlación  $\rho(h)$ . Este proceso es llamado Schlather y su función de distribución marginal bivariada es:

$$P(Z(x_i) \leq z_i, Z(x_j) \leq z_j) = \exp \left\{ \frac{-1}{2} \left( \frac{1}{z_i} + \frac{1}{z_j} \right) \left\{ 1 + \sqrt{1 - 2 \frac{\{\rho(h)+1\} z_i z_j}{(z_i+z_j)^2}} \right\} \right\}$$

donde  $h = \|x_i - x_j\|$  es la distancia euclidiana y  $\rho(h)$  es la función de correlación del proceso Gaussiano fundamental,  $\theta(h) = 1 + \sqrt{\{1 - \rho(h)\}/2}$  es la función de dependencia extremal, que tiene un límite superior de  $\theta(h) < 1.838$  cuando  $h \in \mathbb{R}^2$  y  $\theta(h) < 1.781$  cuando  $h \in \mathbb{R}^3$ . Con este modelo es posible usar una gran variedad de funciones de correlación, sin embargo  $\rho(h)$  es una función definida positiva, lo cual impone la restricción sobre la dependencia extrema, por lo tanto el modelo no puede tomar en cuenta extremos que sean independientes a distancias  $h$  que crecen indefinidamente.

La densidad por pares de  $Z(x_i), Z(x_j)$  es calculada por diferenciación particularmente con respecto a  $z_i, z_j$ .

$$f(z_i, z_j) = \left\{ \frac{\partial V(z_i, z_j)}{\partial z_i} \frac{\partial V(z_i, z_j)}{\partial z_j} - \frac{\partial^2 V(z_i, z_j)}{\partial z_i \partial z_j} \right\} \exp\{-V(z_i, z_j)\}, z_i, z_j > 0$$

donde la medida exponente es:

$$V(z_i, z_j) = -\frac{1}{2} \left( \frac{1}{z_i} + \frac{1}{z_j} \right) \left\{ 1 + \sqrt{1 - 2 \frac{\{\rho(h)+1\} z_i z_j}{(z_i+z_j)^2}} \right\}$$

### 4.2.1. Modelo Gaussiano Geométrico

Este modelo es una derivación de la caracterización de Schlather, es tratado en [Davison et al. \(2012\)](#), si  $Y_s(x) = \exp\{\sigma \varepsilon(x) - \sigma^2/2\}$ ,  $\sigma > 0$ , donde  $\varepsilon(x)$  es un proceso Gaussiano estándar estacionario con función de correlación  $\rho(h)$ , entonces la distribución marginal bivariada es:

$$Pr\{Z(x_i) \leq z(x_i), Z(x_j) \leq z(x_j)\} = \exp \left\{ -\frac{1}{z_i} \Phi \left( \frac{a}{2} + \frac{1}{a} \log \frac{z_j}{z_i} \right) - \frac{1}{z_j} \Phi \left( \frac{a}{2} + \frac{1}{a} \log \frac{z_i}{z_j} \right) \right\} \quad (4.3)$$

con  $a^2(h) = 2\sigma^2\{1 - \rho(h)\}$  y coeficiente extremal  $\theta(h) = 2\Phi[\sigma\{1 - \rho(h)\}^2 / \sqrt{2}]$ , cuyos límites son cuando  $\sigma \rightarrow 0$  o  $\rho \rightarrow 1$ , entonces  $\theta \rightarrow 1$ , mientras que cuando  $\sigma \rightarrow \infty$ ,  $\theta \rightarrow 2$ , para cualquier  $\rho$ , lo que significa que este modelo puede tener como límites procesos max-stable independientes y dependientes totalmente.

## 4.3. Modelo de Brown Resnik

El proceso aleatorio de Brown Resnik (Brown and Resnik (1977); Kabluchko et al. (2009)) se deriva al generalizar al proceso Gaussiano Geométrico, es decir si se toma a

$$Y_s(x) = \exp\{W(x) - \gamma(x)\}, x \in \mathbb{R}^d$$

donde  $W(\cdot)$  es un proceso Gaussiano centrado con patrones muestrales continuos, incrementos estacionarios, semivariograma  $\gamma(h)$ , tal que  $W(0) = 0$  casi seguramente. El proceso Brown Resnik aparece cuando  $W$  es un proceso Browniano fraccional, es decir,  $\gamma(h) \propto h^\alpha$ ,  $0 < \alpha \leq 2, h > 0$ . Para el proceso de Brown Resnik, la distribución en forma cerrada bivariada tiene la forma:

$$P(Z(x_i) \leq z_i, Z(x_j) \leq z_j) = \exp\left[-\left(\frac{1}{z_i}\Phi\left[\frac{a}{2} + \frac{1}{a}\right]\log\frac{z_j}{z_i} + \left(\frac{1}{z_j}\Phi\left[\frac{a}{2} + \frac{1}{a}\right]\log\frac{z_i}{z_j}\right)\right]$$

donde

$$a = \{2\gamma(x_i - x_j)\}^{1/2}, \Phi \text{ es la función de distribución acumulada normal estándar.}$$

El coeficiente extremal correspondiente a este modelo es

$$\theta(h) = 2\Phi[\{\gamma(h)/2\}^{1/2}] \text{ para todo } h \in \mathbb{R}^d$$

## Familias de funciones de correlación

Las familias de correlación para el coeficiente extremal en los modelos max-stable incluyen:

- Whittle Matern:  $\rho(h) = \frac{2^{k-1}}{\Gamma(k)} \left(\frac{\|h\|}{r}\right)^k K_k\left(\frac{\|h\|}{r}\right), k > 0$
- Exponencial:  $\rho(h) = \sigma \exp\left[-\left(\frac{\|h\|}{r}\right)^k\right], 0 < k \leq 2$
- Cauchy:  $\rho(h) = \sigma \left[1 + \left(\frac{\|h\|}{r}\right)^2\right]^{-k}, k > 0$
- Bessel:  $\rho(h) = \sigma \left(\frac{2r}{\|h\|}\right)^k \Gamma(k+1) J_k\left(\frac{\|h\|}{r}\right)$

En las cuatro familias  $r > 0$  representa un parámetro de escala con la dimensión de distancia,  $k$  es un parámetro de forma o suavizamiento que controla las propiedades del campo aleatorio,  $h$  es la distancia de separación entre dos posiciones,  $\Gamma$  denota la función gamma,  $J_k$  y  $K_k$  son las funciones Bessel y Bessel modificadas de orden  $k$ .

## 4.4. Medidas de dependencia espacial para campos aleatorios Max-stable

En estadística espacial, cuando se trata con datos no extremos, la dependencia espacial es estudiada por medio del semivariograma:

$$\gamma(h) = \frac{1}{2} \text{Var}[Z(x+h) - Z(x)]$$

En datos extremos ésta herramienta puede no existir, así que la medida propuesta del grado de dependencia espacial entre valores extremos en campos aleatorios max-stable son el coeficiente extremal y algunos estimadores de éste están basados en madogramas.

### 4.4.1. Coeficiente extremal

El coeficiente extremal es una medida de dependencia extrema entre  $N$  localizaciones en  $R^d$  y está definido para un proceso estacionario max-stable  $Z(\cdot)$  con marginales unitarias Fréchet de la siguiente forma:

$$\Pr[Z(x_1) \leq z, \dots, Z(x_N) \leq z] = \exp(-\theta_N/z)$$

donde  $1 \leq \theta_N \leq N$  son límites inferior y superior correspondientes a dependencia completa e independencia respectivamente,  $\theta_N$  se interpreta como el número de localizaciones independientes.

En procesos max-stable con marginales unitarias Fréchet, las marginales de orden superior  $k \geq 2$  son distribuciones de valor extremo multivariadas, éstas funciones de distribución tienen la forma:

$$\Pr[Z(x_1) \leq z_1, \dots, Z(x_N) \leq z_N] = \exp\{-V(z_1, \dots, z_N)\}$$

$V$  es la función llamada medida exponente para extremos multivariados tratada por [Coles \(2001\)](#), la relación que guarda con el coeficiente extremal es:

$$\theta_N = V(1, \dots, 1).$$

Como caso particular, para el caso bivariado donde la función de distribución es  $F(z, z') = \exp(-V(z, z'))$  se tiene el coeficiente extremal por pares  $\theta_{zz'}$ . Éste es una medida de dependencia para máximos espaciales en dos localizaciones  $x_i, x_j \in X$  separadas por una distancia  $h \in R^d$ . El coeficiente extremal se define como:

#### 4.4. Medidas de dependencia espacial para campos aleatorios Max-stable

---

$\theta = (-V(1, 1))$ , es decir,

$$Pr(Z(h) < z, Z'(0) < z) = \exp(-V(z, z)/z) = \exp(-1/z)^{V(1,1)} = \exp(-\theta(h)/z), z > 0$$

donde  $\theta(h) = \int_s \max\{g(s, 0), g(s, h)\} v ds$

Si  $Z$  es un proceso max-stable el coeficiente extremal es una medida de dependencia exacta. [Schlather \(2002\)](#) da las siguientes propiedades para el coeficiente extremal  $\theta(h)$ :

- Es invariante para distribuciones marginales.
- Simétrico sobre el origen.
- $\theta(0) = 1$ .
- $1 \leq \theta(h) \leq 2$ , con límites inferior y superior correspondientes a dependencia e independencia completa en la separación  $h$  respectivamente.

### Estimadores del Coeficiente Extremal

Algunos estimadores para el coeficiente extremal han sido propuestos por Schlather y Tawn, Smith. Algunos otros estimadores son el Madograma, F-Madograma y lambda madograma.

### Estimador de Smith

El estimador por pares propuesto por [Smith \(1990\)](#), se obtiene de suponer que las marginales de  $Y(x_i)$  y  $Y(x_j)$  son unitarias Fréchet para todo  $x \in R^d$ , entonces  $1/Y(x_i)$  y  $1/Y(x_j)$  tienen distribuciones exponenciales unitarias y  $1/\max(Y(x_i), Y(x_j))$  tiene una distribución exponencial con media  $1/\theta$ . Como los datos han sido transformados  $\{Z_{ni}, 1 \leq n \leq N, 1 \leq i \leq p\}$  y el ajuste por máxima verosimilitud de las distribuciones marginales asegura que  $\sum_{i=1}^n 1/Z_{ni}(x_i) = \sum_{i=1}^n 1/Z_{nj}(x_j) = N$ , entonces el estimador de Smith denominado raw es

$$\widehat{\theta}(x_i - x_j) = N / \sum_{n=1}^N \min\{Z_{ni}(x_i)^{-1}, Z_{nj}(x_j)^{-1}\}$$

Estas cantidades son llamadas estimaciones raw del coeficiente extremal.



## 4.4. Medidas de dependencia espacial para campos aleatorios Max-stable

### Estimador de Schlather y Tawn

Un estimador de la dependencia extrema multivariada que cumple la propiedad de consistencia fue propuesto por [Schlather and Tawn \(2003\)](#), el estimador se obtiene al maximizar la log verosimilitud:

$$l_A(\theta_A) = \text{card} \left\{ i : \max_{j \in A} (Z_i(x_j) \overline{Z(x_j)}) > z \right\} \log \theta_A - \theta_A \sum_{i=1}^n [\max\{z, \max_{j \in A} (Z_i(x_j) \overline{Z(x_j)})\}]^{-1}$$

donde  $\overline{Z(x_j)} = n^{-1} \sum_{i=1}^n \frac{1}{Z_i(x_j)}$  y el correspondiente estimador es denotado por  $\widehat{\theta}_{A,t}$ . Este es llamado estimador naive. Aquí las variables aleatorias Fréchet son escaladas por  $\overline{Z(x_j)}$  para asegurar que  $\widehat{\theta}_j = 1$  donde  $t = 0$ ,  $t$  denota la probabilidad marginal de que no se exceda un umbral  $z$ , esto es  $t = \exp(-1/z)$ .

#### 4.4.2. Enfoque basado en Madograma

El variograma es una herramienta usual para medir el grado de dependencia espacial en procesos gaussianos, sin embargo, en procesos max-stable unitarios Fréchet, éste no existe teóricamente dado que la media y la varianza del proceso no existen. Así, se han derivado para medir el grado de dependencia por pares para extremos, otro tipo de herramientas basadas en madogramas. Estos son el Madograma, el F-madograma y el  $\lambda$ -madograma. El primero esta ligado a un coeficiente extremal cuando las marginales tienen parámetro de forma  $\xi \leq 0.5$ , el segundo es tratado cuando las marginales del proceso son de cola pesada (Fréchet), ya que si el parámetro de forma es  $\xi > 0.5$  los momentos de esta distribución no existen y al considerar a  $F(Z(x))$  que tiene distribución uniforme entonces los momentos existen por lo que la esperanza del valor absoluto de la diferencia existe. Para ambos fue encontrada por [Cooley et al. \(2006\)](#) la relación con el coeficiente extremal. Para estos se considera el caso de valores iguales para  $z$ ,  $z_i = z_j = z$ , es decir  $V(z, z)$ . [Naveau et al. \(2009\)](#) introduce el  $\lambda$ -madograma, el cual a diferencia de los anteriores caracteriza la dependencia espacial del campo aleatorio totalmente, es decir, se considera el caso donde  $z_i \neq z_j$ .

- Madograma

Supóngase a  $\{Z(x)\}$  como cualquier campo aleatorio espacial estacionario max-stable con marginales GEV y coeficiente extremal  $\theta(h)$ . Si el parámetro de forma  $\xi$  satisface  $\xi < 1$ , entonces el madograma  $v(h)$ , y el coeficiente extremal  $\theta(h)$  verifican

$$\theta(h) = \begin{cases} \mu_\beta \left( \mu + \frac{v(h)}{\Gamma(1-\xi)} \right), & \text{si } \xi < 1, \text{ y } \xi \neq 0, \\ \exp\left(\frac{v(h)}{\lambda}\right), & \text{si } \xi = 0, \end{cases}$$

donde  $u_\beta(u)$  pertenece a la expresión de la distribución GEV

#### 4.4. Medidas de dependencia espacial para campos aleatorios Max-stable

$$P(Z(x) \leq u) = \exp\left(\frac{-1}{u_{\beta}(u)}\right), \text{ con } u_{\beta}(u) = \left(1 + \xi \frac{u-\mu}{\lambda}\right)_+^{1/\xi}$$

donde  $a_+ = a$  si  $a > 0$  y 0 d.o.m.. El vector de parámetros  $B$  es definido como  $\beta = (\mu, \lambda, \xi)$ , los cuales son parámetros de localización, escala y forma respectivamente.  $\Gamma(\cdot)$  es la función gamma clásica.

La aplicación directa de la relación madograma coeficiente extremal da  $v(h) = 1 - \frac{1}{\theta(h)}$ , si  $\{Z(x)\}$  tiene marginales Weibull con parámetros ( $u = 0, \lambda = 1, \xi = -1$ ) y  $v(h) = \log\theta(h)$ , si  $\{Z(x)\}$  tiene marginales unitarias Gumbel con parámetros ( $u = 0, \lambda = 1, \xi = 0$ ).  $v(h) = \sqrt[1]{\Pi}(\theta(h)\frac{1}{2} - 1)$ , si  $\{Z(x)\}$  tiene marginales Fréchet con parámetros ( $u = 1, \lambda = 1, \xi = 1/2$ ). Un estimador del coeficiente extremal  $\theta(h)$  es

$$\widehat{\theta}(h) = \mu_{\beta} \left( \mu + \frac{\widehat{v}(h)}{\Gamma(1-\xi)} \right)$$

donde

$$\widehat{v}(h) = \frac{1}{2N_h} \sum_{(x_i, x_j) \in N_h} |Z(x_j) - Z(x_i)|,$$

y  $N_h$  es el conjunto de pares de localizaciones cuyas distancias están contenidas en el conjunto  $[h - \xi, h + \xi]$  para  $\xi > 0$ .

Las propiedades asintóticas del estimador  $\widehat{\theta}(h)$  dependen del comportamiento asintótico del estimador de madograma  $\widehat{v}(h)$ . Este madograma es simplemente una suma de incrementos absolutos estacionarios.

Trabajar con estos variogramas de primer orden tiene algunas ventajas como que es simple de definir y fácil de calcular, además tiene una liga clara con la teoría de valores extremos a través del coeficiente extremal.

- F-Madograma

Un segundo estimador de la dependencia extrema es el F-madograma propuesto por [Cooley et al. \(2006\)](#), éste es definido como:

$$v^F(h) = \frac{1}{2} E|F(Z(x+h)) - F(Z(x))|,$$

la marginal de  $\{Z(x)\}$  es  $F(u) = \exp(-1/u_{\beta}(u))$

Para derivar la equivalencia entre el F-madograma y el coeficiente extremal, tenemos que

$$v^F(h) = \theta E[F(Z(x))F(Z(x))^{\theta(h)-1}] - E[F(Z(x))]$$

también

$$E[F^r(Z(x))] = \frac{1}{1+r} \text{ debido a que } F(Z(x)) \text{ sigue una distribución uniforme.}$$

Resolviendo la igualdad anterior se obtiene la relación entre el coeficiente extremal y el F-madograma.

$$2v^F(h) = \frac{\theta(h)-1}{\theta(h)+1} \text{ o de manera inversa}$$

$$\theta(h) = \frac{1+2v^F(h)}{1-2v^F(h)}.$$

un estimador para el F-madograma es:

#### 4.4. Medidas de dependencia espacial para campos aleatorios Max-stable

$\widehat{v}^F(h) = \frac{1}{2n} \sum_{i=1}^n |\widehat{F}(Z_i(x+h)) - \widehat{F}(Z_i(x))|$ , donde  $Z_i$  son las observaciones  $i$ -ésimas del campo aleatorio en las localizaciones  $x+h$  y  $x$ , y  $n$  es el número de observaciones.

Por lo tanto, [Cooley et al. \(2006\)](#) demuestra que existe una relación uno a uno entre el F-madograma y el coeficiente extremal.

De aquí un estimador para el coeficiente extremal es:

$$\widehat{\theta}(h) = \frac{1+2\widehat{v}^F(h)}{1-2\widehat{v}^F(h)}.$$

- $\lambda$ -Madograma

Un tercer estimador de la dependencia extrema por pares el cual procede del concepto de variograma, fue introducido por [Naveau et al. \(2009\)](#). Éste se define como:

$$V(h, \lambda) = \frac{1}{2} E[|F^\lambda\{Z(x+h)\} - F^{1-\lambda}\{Z(x)\}|]$$

para cualquier  $\lambda \in (0, 1)$ .

Se define el  $\lambda$ -madograma con la finalidad de estimar la función de dependencia extremal  $V_h(u, v)$ . Esto es, se busca establecer la convergencia del estimador del  $\lambda$ -madograma. Se cuenta para ello con mediciones en cada posición correspondientes al par de variables aleatoria  $\{Z(x+h), Z(x)\}$ , entonces es posible obtener los máximos por bloques en cada posición, considerando que la estacionalidad y tendencia han sido removidas. Por la teoría de valores extremos bajo condiciones débiles los máximos convergen en distribución cuando  $n$  es grande, por lo que es posible obtener las funciones empíricas de las marginales, las cuales son utilizadas para encontrar el estimador del  $\lambda$ -madograma.

Ahora bien la relación que existe entre el  $\lambda$ -madograma y la función de dependencia extremal, se da bajo el supuesto de que los máximos  $\{Z(x+h), Z(x)\}$  tienen marginales unitarias Fréchet y es la siguiente:

$$V_\lambda(h) = \frac{V_{x+h,x}(\lambda, 1-\lambda)}{1+V_{x+h,x}(\lambda, 1-\lambda)} - c(\lambda)$$

donde  $c(\lambda) = 3/2(1+\lambda)(2-\lambda)$ .

O de forma inversa

$$V_{x+h,x}(\lambda, 1-\lambda) = \frac{c(\lambda)+V_\lambda(h)}{1-c(\lambda)-V_\lambda(h)}$$

Bajo las condiciones dadas con anterioridad el estimador del lambda mado-grama dado a continuación no usa el supuesto de que las marginales de las muestras bivariadas sean unitarias Fréchet:

$$\widehat{V}(h, \lambda) = \frac{1}{2T} \sum_{t=1}^T T |\widehat{F}_T^\lambda\{Z_t(x+h)\} - \widehat{G}_T^{1-\lambda}\{Z_t(x)\}|]$$

donde  $T$  denota los años ,esto es  $t = 1, \dots, T$ , y  $\widehat{F}_T$  y  $\widehat{G}_T$  denotan las funciones empíricas definidas por:

$$\widehat{F}_T(u) = \frac{1}{T} \sum_{t=1}^T 1_{\{Z_t(x+h) \leq u\}}$$

## 4.5. Ajuste de un proceso Max-stable a datos

---

$$\widehat{G}_T(u) = \frac{1}{T} \sum_{t=1}^T 1_{\{Z_t(x) \leq u\}}$$

Por la relación encontrada entre el  $\lambda$ -madograma y la función extremal, el estimador de dependencia extrema es:

$$\widehat{V}_{x+h,x}(\lambda, 1 - \lambda) = \frac{c(\lambda) + \widehat{V}(h,\lambda)}{1 - c(\lambda) - \widehat{V}(h,\lambda)}$$

## 4.5. Ajuste de un proceso Max-stable a datos

Para los procesos max-stable únicamente se tiene la distribución bivariada, así que el ajuste a datos no es sencillo. Para esto se cuenta con dos métodos: el de mínimos cuadrados tratado por [Smith \(1990\)](#) en el que el supuesto es marginales Fréchet unitarias y el de máxima verosimilitud compuesta por pares propuesto por [Padoan et al. \(2010\)](#). Para el ajuste y la estimación de parámetros de dependencia espacial junto con los parámetros marginales si los datos en el espacio presentan tendencia, se estima la superficie de respuesta adecuada por medio de regresión o regresión semiparamétrica.

### 4.5.1. Mínimos cuadrados

Suponiendo marginales unitarias Fréchet, el procedimiento de ajuste por mínimos cuadrados consiste en minimizar la función

$$C(\psi) = \sum_{i < j} \left( \frac{\widetilde{\theta}_{ij} - \hat{\theta}_{ij}}{s_{ij}} \right)^2$$

donde  $\psi$  es el vector de parámetros de procesos max-stable,  $\hat{\theta}_{ij}$  es el valor del coeficiente extremal ajustado del modelo max-stable para las posiciones  $i, j$ ,  $\widetilde{\theta}_{ij}$  es el estimador semiparamétrico de [Smith \(1990\)](#),  $s_{ij}$  es la desviación estándar relacionada a la estimación  $\widetilde{\theta}_{ij}$ , obtenida por el estimador de jackknife.

Los errores estándar para este modelo son calculados de la forma

$$s_{ij} = \left\{ \frac{N-1}{N} \sum_{n=1}^N \widetilde{\theta}_{ij}^n - \widetilde{\theta}_{ij}^2 \right\}^{1/2}.$$

Una ventaja de este estimador es que da un buen ajuste en lugares donde la dependencia es fuerte. Las marginales del proceso deben ser unitarias Fréchet por lo que se deben transformar los datos, el estimador carece de la propiedad de eficiencia en la forma del límite de Cramer Rao dado que las observaciones provienen de una GEV y no de una Normal.

## 4.5. Ajuste de un proceso Max-stable a datos

---

Una prueba de ajuste es examinar los residuales contra los valores ajustados o distancia entre puntos.

### 4.5.2. Verosimilitud

La inferencia en el caso de modelos max-stable se realiza por máxima verosimilitud compuesta por pares desarrollada por [Padoan et al. \(2010\)](#) sobre la distribución bivariada de Smith. [Padoan et al. \(2010\)](#) implementó este método a partir de verosimilitud compuesta la cual es tratada por [Varin and Vidoni \(2005\)](#).

El mérito de la verosimilitud compuesta es que reduce la complejidad computacional, por lo que es posible tratar con conjuntos de datos grandes correlacionados y modelos muy complejos, y cuando la verosimilitud estándar y métodos bayesianos no son adecuados.

[Varin and Vidoni \(2005\)](#) definen a la verosimilitud compuesta como:

**Definición:** Sea  $\{f(y \in \theta); y \in Y, \theta \in \Theta\}$  un modelo estadístico paramétrico con  $y \subseteq R^n$ ,  $\theta \subseteq R^d$ ,  $n \geq 1$  y  $d \geq 1$ . Considere un conjunto de eventos  $\{A_i; A_i \subseteq F, i \in I\}$  donde  $I \subseteq N$  y  $F$  es una sigma álgebra sobre  $Y$ . La verosimilitud compuesta es:

$$L_c(\theta; y) = \prod_{i \in I} f(y \in A_i; \theta)^{w_i}$$

donde  $f(y \in A_i; \theta) = f(y_j \in Y_j; y_j \in A_i)$  con  $y = (y_1, \dots, y_n)$ , mientras que  $w_i, i \in I$  es un conjunto de pesos adecuado. La log verosimilitud compuesta asociada es  $l_c(\theta; y) = \log L_c(\theta; y)$ .

Vista a la verosimilitud total  $L(\theta; y)$  como un caso especial, la verosimilitud compuesta puede agruparse en dos clases generales. La primera incluye a la verosimilitud por pares la cual esta basada en eventos marginales relacionados a pares de observaciones, en ésta esta basada la inferencia de procesos max-stable.

La ecuación de estimación asociada a la verosimilitud compuesta  $L_c(\theta; y)$  es la primera derivada de la log verosimilitud compuesta igualada a cero  $\nabla l_c(\widehat{\theta}_{MCL}; y) = 0$ , donde  $\theta$  es el vector de parámetros a estimar.

Dado que cada componente de la verosimilitud compuesta es elemento de la verosimilitud  $L(\theta; y)$ , entonces bajo condiciones de regularidad el estimador de máxima verosimilitud compuesta  $\widehat{\theta}_{MCL} = \widehat{\theta}_{MCL}(Y)$  es insesgado y consistente, este es un estimador distribuido asintóticamente normal de media  $\theta$  y matriz de varianzas covarianzas  $H(\theta)^{-1}J(\theta)H(\theta)^{-T}$ . Aquí  $J(\theta) = V_f\{\nabla l_c(\theta; Y)\}$  y  $H(\theta) = E_f\{\nabla^2 l_c(\theta; Y)\}$ , donde la esperanza es con respecto a  $f(y, \theta)$ .

## 4.5. Ajuste de un proceso Max-stable a datos

### 4.5.3. Selección del modelo

La selección de un modelo en procesos max-stable se realiza con el criterio de información de verosimilitud compuesta (CLIC), el cual ha sido derivado por [Varin and Vidoni \(2005\)](#) para modelos mal especificados, éste tiene la propiedad de considerar que el modelo candidato no incluye al verdadero.

El objetivo de la selección de un modelo es encontrarlo de tal manera que sea aquel que ofrezca la predicción mas satisfactoria de una v.a. futura dada una realización del vector  $Y = (y_1, \dots, y_n)$ . Se cuenta con una gama de modelos para selección relacionados a la familia  $\{f(y\epsilon; \theta), y\epsilon Y, \theta\epsilon\Theta\}$ , los cuales pueden o no pueden contener al modelo verdadero  $g(y)$ .

El procedimiento de selección consiste en elegir el modelo que minimiza a la esperanza de la divergencia compuesta de Kullback-Leibler  $E_{g(y)}I_c(g; f)$  entre la densidad verdadera  $g(y)$  y la densidad estimada  $\widehat{f}(z) = f(z; \widehat{\theta})$  del modelo estadístico elegido. O el que maximiza  $\varphi(g, f) = \sum_{i\epsilon I} E_g(y)[E_{g(z)}\{\log f(z\epsilon A_i; \widehat{\psi}_{MCL})\}]w_i$  para el cual se requiere se conozca la densidad verdadera  $g(z)$ . Por lo que en la práctica se maximiza a la estadística  $\widehat{\varphi}(g, f)$  definida como un estimador adecuado de  $\varphi(g, f)$  basado en  $Y$ .

La divergencia compuesta de Kullback-Leibler de una densidad  $f(z)$  relativa a  $g(z)$ , donde  $Z = (Z_1, \dots, Z_n)$  v.a., está definida como:

$$I_c(g, h) = E_g(z)[\log\{\frac{L_c(g; Z)}{L_c(h; Z)}\}] = \sum_{i\epsilon I} E_g(z)[\log g(Z\epsilon A_i) - \log f(Z\epsilon A_i)]w_i \text{ donde } L_c(g; Z) = \prod_{i\epsilon I} g(Z\epsilon A_i)^{w_i} \text{ y } L_c(f; Z) = \prod_{i\epsilon I} f(Z\epsilon A_i)^{w_i}$$

Por otra parte un estimador adecuado de  $\varphi(g, f)$  es:

$$l_c(\widehat{\theta}; Y) = \log L_c(\widehat{\theta}; Y) = \sum_{i\epsilon I} \log f(y\epsilon A; \widehat{\theta})w_i \text{ la cual es la log verosimilitud compuesta maximizada.}$$

Bajo condiciones de regularidad, el criterio de información de verosimilitud compuesta es un estimador insesgado y el mejor modelo es aquel que lo maximiza. Para una muestra aleatoria  $Y$  éste esta definido como:

$$l_c(\theta^*; Y) + tr \left\{ \widehat{J}(y) \widehat{H}(y)^{-1} \right\}$$

donde  $\widehat{J}(y)$  y  $\widehat{H}(y)$  son estimadores insesgados y consistentes de  $J(\theta^*)$  y  $H(\theta^*)$

La estadística para la selección del modelo es una generalización del criterio de Takeuchy, el cual tiene la misma forma, pero con los estimadores calculados de la verosimilitud ordinaria.

### Criterio de Información de Takeuchy

Los modelos mal especificados presentan dos características: los modelos candidatos no contienen al verdadero y la inferencia se realiza con máxima verosimilitud compuesta.

Takeuchy en 1976 creo un criterio, el cual debido a sus propiedades es una herramienta poderosa para seleccionar modelos mal especificados. Este ha sido derivado de máxima verosimilitud total por lo que los estimadores para los parámetros contenidos en el mismo son consistentes, insesgados y eficientes.

El TIC es definido por:

$$TIC = -2\{l_p(\theta; z)(\widehat{\theta}_{MCPL E}; z) - tr[\widehat{J}(\widehat{\theta}_{MCPL E})\widehat{H}(\widehat{\theta}_{MCPL E})^{-1}]\}$$

donde H y J son la matriz de información observada y la estadística score correspondiente al logaritmo de la verosimilitud por pares  $l_p(\theta; y)$ ,  $\widehat{\theta}_{MCPL E}$  es el estimador de máxima verosimilitud compuesta por pares.

La selección del modelo consiste en elegir aquel que minimiza al criterio de Takeuchy, entre más bajo sea el valor, de mejor calidad será el modelo.

El criterio de Takeuchy es utilizado para la selección del modelo en procesos max-stable, en esta representación los modelos con los que se cuenta para el modelado de la dependencia espacial tienen forma cerrada bivariada  $f(z)$ , y las observaciones muestrales provienen de una distribución  $g(z)$  multivariada pero desconocida, por lo que los modelos candidatos a ser seleccionados no contendrán al modelo verdadero. Además para la inferencia se utiliza máxima verosimilitud compuesta por pares. De aquí que los modelos max-stable son mal especificados.

# Capítulo 5

## Aplicación

En este capítulo se da una reseña de la contaminación por ozono en la Ciudad de México. Se aplica metodología de procesos max-stable para el análisis espacio temporal del mismo.

### 5.1. Contaminación por ozono

En la estratosfera, segunda capa de la atmósfera que comprende entre 12 y 50 Km a partir del suelo, se encuentra la capa de ozono, la cual protege de los rayos ultravioleta lo que hace posible la vida en la tierra. El ozono ( $O_3$ ) es un gas reactivo de color azul que se produce mediante el efecto de la luz solar sobre el oxígeno. El ozono ( $O_3$ ) en la primera capa troposfera denominado ozono urbano es muy dañino para la salud y el medio ambiente, su presencia en el aire es resultado de la combinación de óxidos de Nitrógeno ( $NO_x$ ) e Hidrocarburos volátiles o compuestos orgánicos volátiles ( $COV_s$ )(derivados de gases por el uso de combustibles, escapes vehiculares, industrias, servicios públicos (hoteles, restaurantes, baños públicos, etc.)) y la radiación ultravioleta los que consecuentemente actúan como precursores.

La contaminación por ozono en la Ciudad de México empezó a estudiarse desde 1950, debido a los daños que esta causa a las plantas, personas e inmuebles.

Las concentraciones de ozono se empiezan a analizar en forma continua a partir de 1980, y con el objetivo de llevar un registro de la contaminación atmosférica en 1985 empieza a operar un sistema de monitoreo denominado Red Automática de Monitoreo Atmosférico (RAMA) que contaba con 25 estaciones en toda la ciudad. La regulación sobre los límites de este contaminante es debido a la Norma Oficial



## 5.1. Contaminación por ozono

---

Mexicana NOM-020-SSA1/1993, la cual establece que los niveles no deben exceder el límite de 0.11 ppm, lo que es equivalente a  $216\mu\text{g}/\text{m}^3$ , en una hora, una vez al año en un periodo de tres años para protección a la salud de la población susceptible.

A partir del 2000 se empiezan a realizar estudios estadísticos sobre las concentraciones de ( $\text{O}_3$ ) y el grado de asociación con sus precursores ( $\text{NO}_x$ ,  $\text{COV}_s$ ) y con algunos parámetros del estado del tiempo; temperatura, humedad relativa y velocidad del viento. Estos estudios muestran tendencias cíclicas en 24 horas presentándose los picos más altos ente 1 y 4 de la tarde, lo cual se asocia al incremento de la temperatura y a la actividad vehicular. Tendencias decrecientes a partir de 1992, lo cual muestra que las medidas preventivas empiezan a funcionar (Correa et al. (2000), Huerta et al. (2004)).

### 5.1.1. Descripción de los datos

La base de datos para este trabajo fue proporcionada por el Sistema Meteorológico de Medio Ambiente (SIMAT) de la Ciudad de México. Se utilizaron mediciones diarias de los niveles de ozono registradas cada hora por 14 estaciones meteorológicas ubicadas en la zona metropolitana, en la Tabla 5.1 se muestran las estaciones meteorológicas y su localización geográfica. La información corresponde al periodo de 1998 a 2010. La Fig. 5.1 presenta la Red Automática de Monitoreo Atmosférico la cual muestra la ubicación de cada una de las estaciones meteorológicas de la Ciudad de México:

## 5.1. Contaminación por ozono

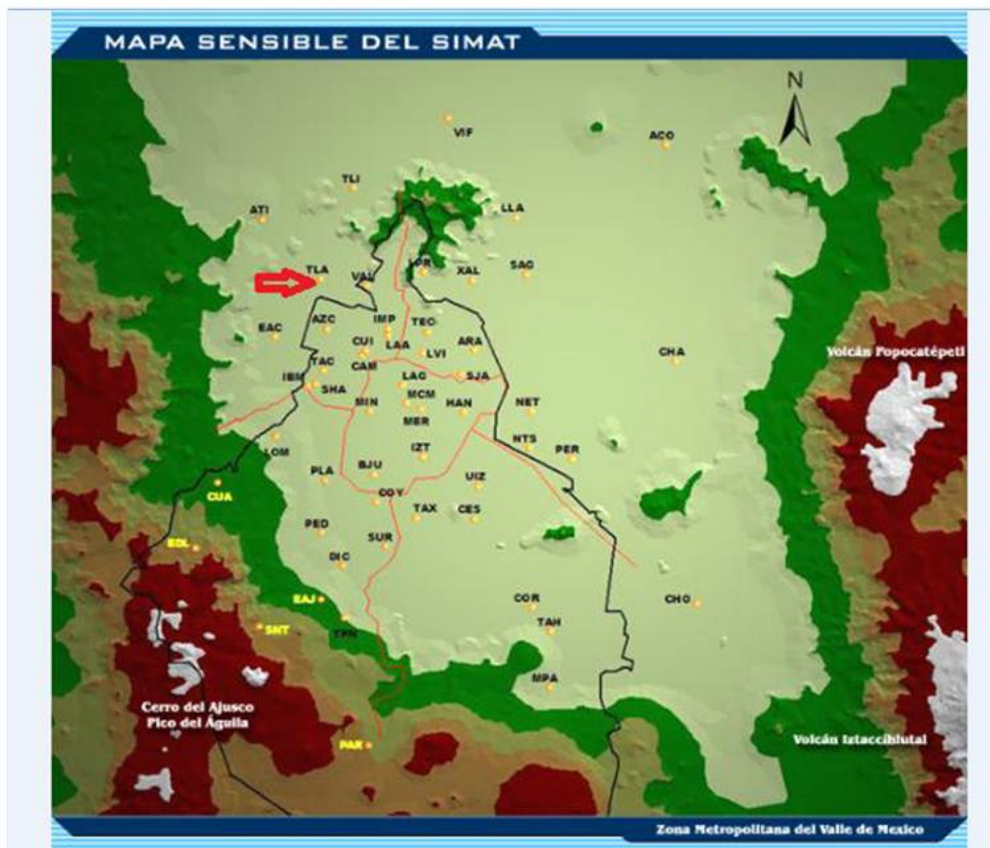
---

**Tabla 5.1:** Ubicación de las estaciones meteorológicas

Estación	abreviación	Longitud	Latitud
Lagunilla	LAG	485810	2149915
Tacuba	TAC	478750	2151193
San Agustín	SAG	496859	2159721
Azcapotzalco	AZC	479142	2154936
Tlalnepantla	TLA	478573	2159310
Xalostoc	XAL	491980	2159225
Merced	MER	487478	2147790
Pedregal	PED	478601	2136772
Cerro de la Estrella	CES	492156	2137995
Plateros	PLA	478986	2141451
UAM Iztapalapa	UIZ	492526	2140927
Taxqueña	TAX	487059	2138103
Cuajimalpa	CUA	4686635	2141112
Tlahuac	TAH	498926	2128017

## 5.1. Contaminación por ozono

Figura 5.1: Mapa de las estaciones meteorológicas en la Ciudad de México



## 5.2. Metodología

1. Para lograr la independencia en series temporales se pueden considerar los valores máximos de observaciones de dos o más días (Coles (1993)). Por lo tanto se formaron bloques de datos de 72 observaciones consecutivas (tres días). Debido a que la base de datos carecía de algunos registros, los bloques correspondientes no se tomaron en cuenta para el análisis. De esta forma, el tamaño de muestra de los máximos de bloques en cada posición quedaron de los siguientes tamaños: 119,112,94,104,82,109,119,115,91,93,68,56,74; para los años consecutivos de 1998 a 2010, respectivamente. Por el teorema de Fisher y Tippett tales máximos se distribuyen como una GEV. Con el fin de aproximar las series temporales en cada posición al supuesto de estacionariedad de los procesos max-stable, los análisis se hicieron por periodos anuales.

Con la finalidad de hacer el análisis espacial de los niveles máximos de ozono con procesos max-stable, los datos en cada una de las posiciones se ajustaron a una DGVE y mediante la prueba de Kolmogorov-Smirnov se verificó el ajuste. Posteriormente los datos se transformaron a una distribución Fréchet unitaria en cada posición mediante  $Y = \{1 - \frac{\xi(x-\mu)}{\sigma}\}^{-1/\xi}$  (Davison et al. (2012)).

2. Para analizar la dependencia extrema espacial se obtuvieron los estimadores no paramétricos de Smith, Schlather y Tawn y los estimadores de madograma, f-madograma y lambda madograma de esta dependencia.

3. Ajustes por mínimos cuadrados y máxima verosimilitud

Para realizar un análisis de la dependencia extremal espacial con los datos transformados, se estimaron los parámetros del modelo de Smith y de Schlather por mínimos cuadrados de todos los años. Se consideraron a las correlaciones exponencial y cauchy para el modelo de Schlather.

Se ajustaron a los datos ya transformados a distribución Fréchet unitaria en cada posición los modelos espaciales max-stable de Smith, Schlather, Gaussiano Geométrico y de Brown Resnik. Lo dos siguientes al de Smith con las funciones de correlación whitmat, Bessel, exponencial y cauchy. Los parámetros de los modelos ajustados se estimaron por máxima verosimilitud por pares. El modelo que mejor ajustó se eligió con el criterio de información de máxima verosimilitud o criterio de Takeuchi.

4. Se determinó la tendencia por medio del rango del modelo de Schlather, fijando el parámetro de suavizamiento. Davison et al. (2012) consideran que los parámetros de dependencia rango y suavizamiento de los modelos Schlather y Gaussiano Geométrico no son fáciles de encontrar, por lo que ellos fijan alguno de los dos y obtienen el resto de parámetros. Debido a que esta dificultad se presentó en esta

### 5.3. Resultados

---

investigación el parámetro de suavizamiento fue fijado en el modelo Schlather y se obtuvieron el resto de parámetros. En cambio el modelo Gaussiano Geométrico fue más complicado de ajustar, pues el parámetro con dificultad de encontrar fue el de rango. Debido a tal inestabilidad en estos modelos se decidió realizar el análisis de la tendencia con el modelo Schlather y fijar el parámetro de suavizamiento.

5. Con los Modelos Lineales Vectoriales Generalizados(VGLM) se confirma la tendencia espacio temporal lineal en los parámetros. La significancia estadística la da el valor absoluto de  $t$  si  $|t| > 2$ , entonces existe una relación lineal de la covariable respecto al parámetro, para esto se usó el paquete VGAM en R ([R DEVELOPMENT CORE TEAM \(2012\)](#)).

Para lo realizado en los pasos del 1 al 4 se utilizó el paquete SpatialExtremes en R ([R DEVELOPMENT CORE TEAM \(2012\)](#)).

### 5.3. Resultados

1. Los ajustes de la DGVE fueron estadísticamente significativos ( $\alpha = 0.05$ ) al realizar la prueba Kolmogorov Smirnov. Se aceptó la hipótesis nula

$H_0$ : Los datos analizados siguen una distribución DGEV.

Con esto se asegura que los datos transformados tengan distribuciones unitarias Fréchet.

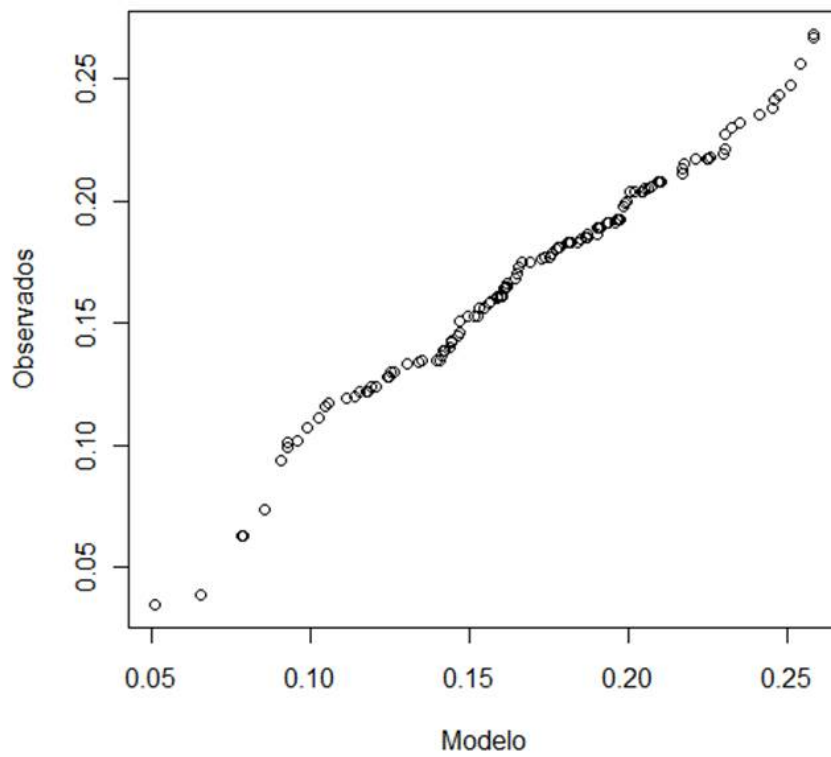
En el apéndice C se incluyen las tablas para los resultados en todas las posiciones correspondientes al estadístico de Kolmogorov y el p-value.

La Fig. 5.2 muestra los valores observados contra los generados por el modelo para la estación meteorológica lagunilla, esta gráfica es una muestra del buen ajuste para una de las posiciones.

### 5.3. Resultados

---

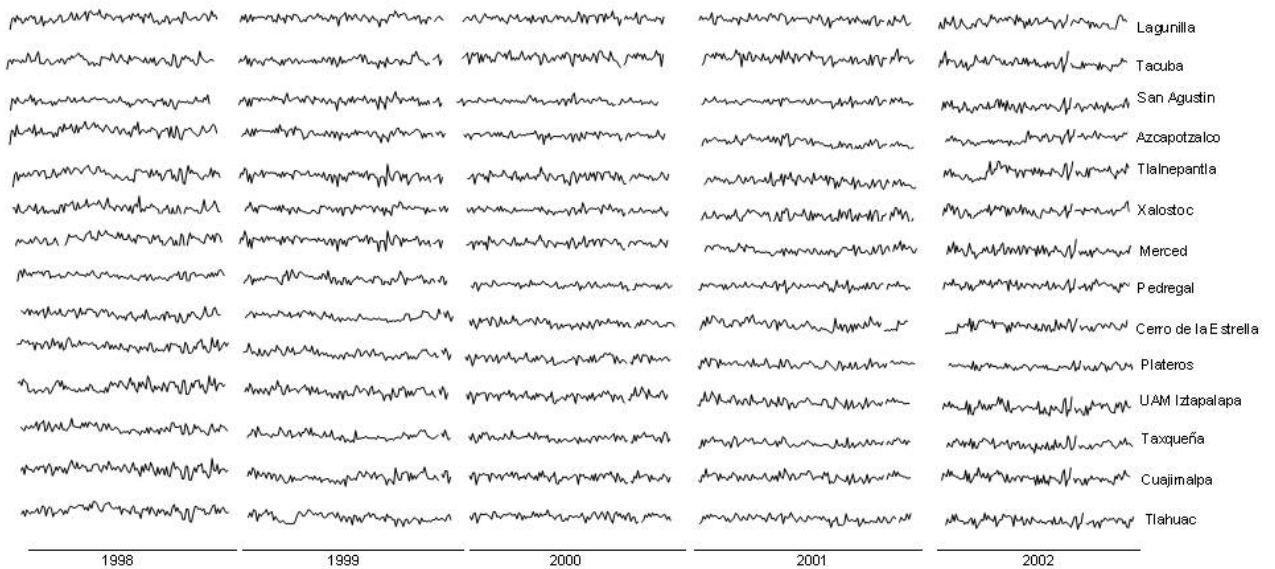
Figura 5.2: Valores observados & los del Modelo ajustado en la estación Lagunilla



### 5.3. Resultados

Para aproximar a la estacionariedad del proceso max-stable se decidió realizar este análisis por periodos de tiempo anuales. La Fig. 5.3 muestra las series de tiempo de 5 años. En cada uno de los periodos anuales se obtuvieron los máximos por bloques bajo el enfoque de "blok máxima" en grupos de observaciones diarias de cada hora de tres días, esto asegura independencia entre las observaciones y es posible ajustar una distribución DGVE.

Figura 5.3: Series de tiempo de 5 años para todas las estaciones meteorológicas



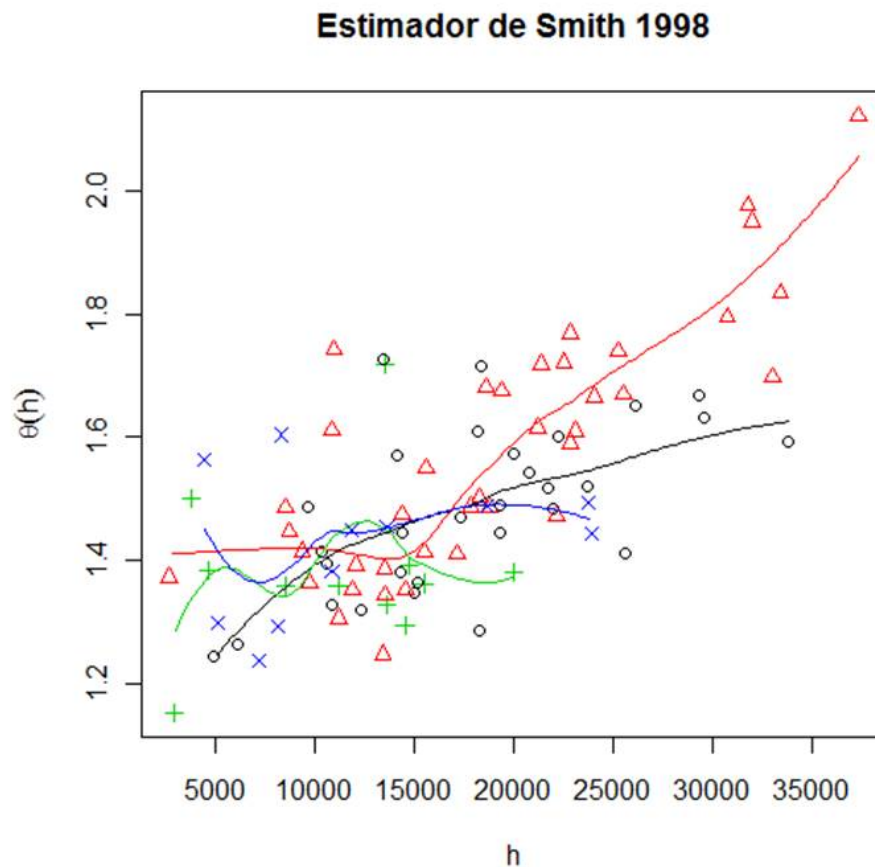
### 5.3. Resultados

#### 2. Estimadores del coeficiente extremal

##### Estimadores no paramétricos

Las Figuras 5.4, 5.5, 5.6, 5.7 muestran los coeficientes extremales no paramétricos Smith y Schlather y Tawn para los años 1998 y 2010. El estimador de Smith muestra variabilidad espacial en colores para cuatro direcciones diferentes, el comportamiento se visualiza anisotrópico. En 1998 la variabilidad espacial es grande y parecida hacia dos direcciones y menor en las dos restantes, mientras que para el 2010 la variabilidad es grande en las dos direcciones correspondientes y muy parecidas a las de 1998 y en las otras dos, la variabilidad sigue siendo pequeña. Otra característica que se observa en ambos estimadores es una mayor variabilidad de la dependencia en el año 1998 y menor en el 2010:

**Figura 5.4:** Estimador no paramétrico de Smith de la función extremal para el año 1998

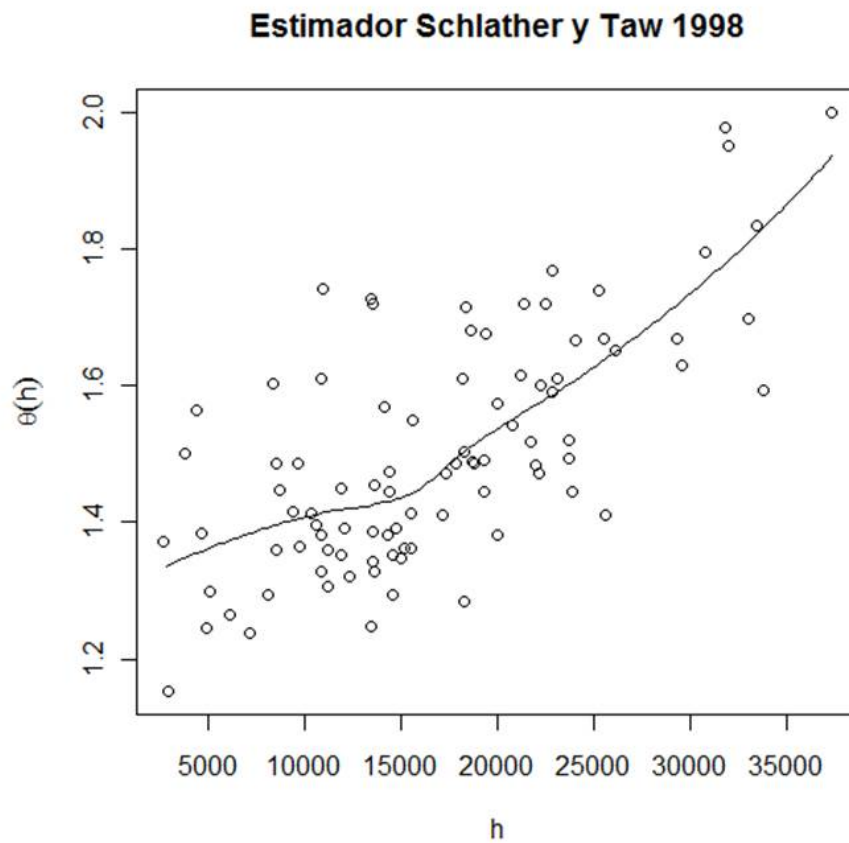




### 5.3. Resultados

---

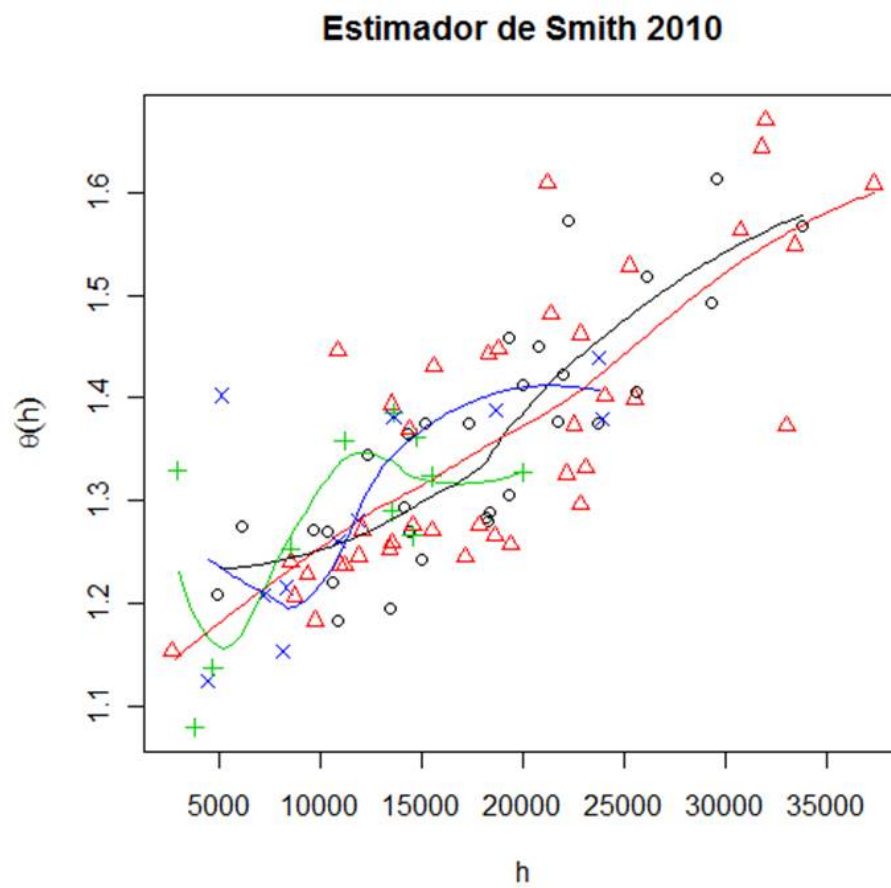
**Figura 5.5:** Estimador no paramétrico de Schlather y Tawn de la función extremal para el año 1998



### 5.3. Resultados

---

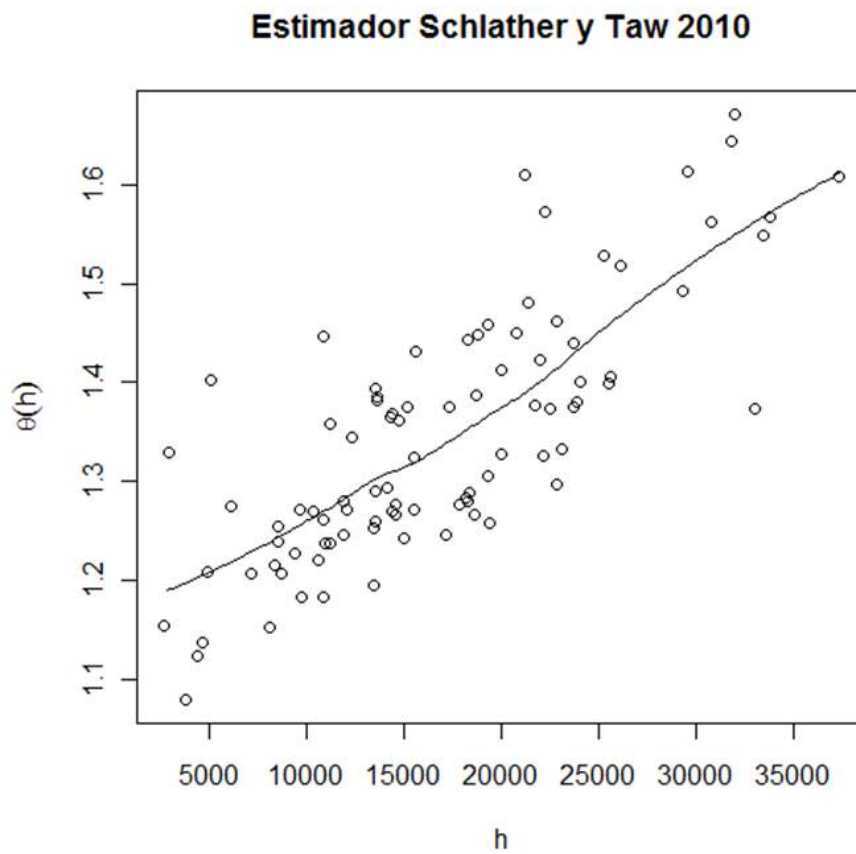
**Figura 5.6:** Estimador no paramétrico de Smith de la función extremal para el año 2010



### 5.3. Resultados

---

**Figura 5.7:** Estimador no paramétrico Schlather y Tawn de la función extremal para el año 2010



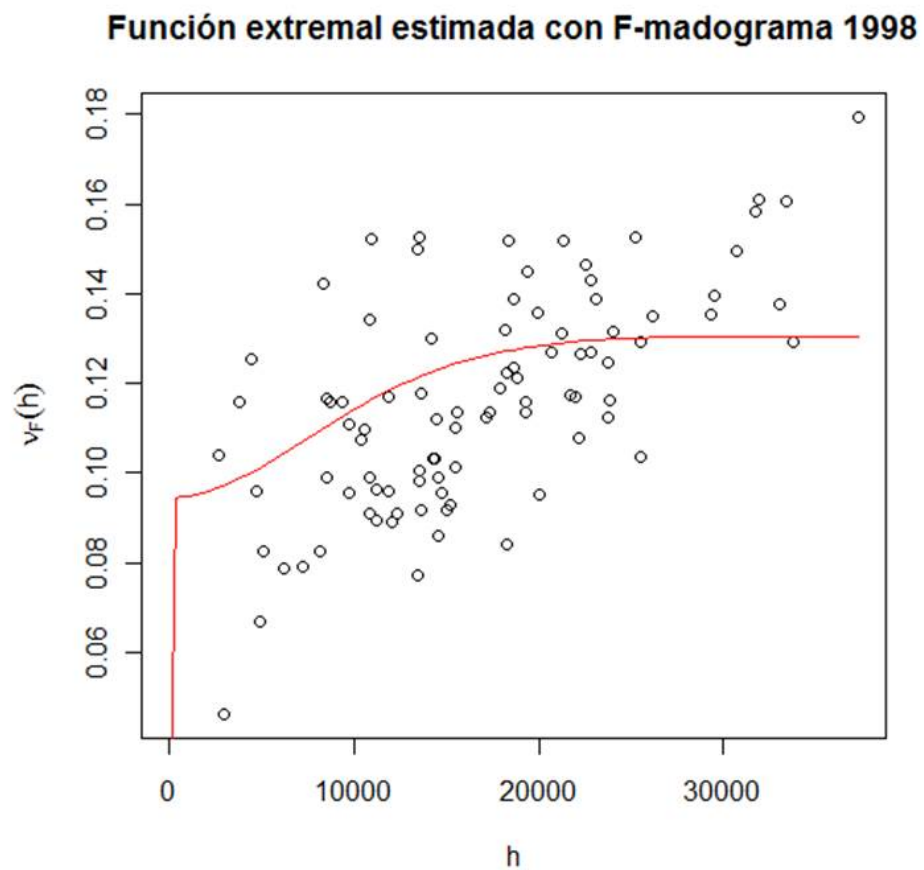
### 5.3. Resultados

---

#### Estimadores de Madograma

Las Figuras 5.8 y 5.9 presentan la función extremal para los años 1998 y 2010 estimada en función del f-madograma, puede notarse que a distancias cortas la dependencia es más fuerte, además de la misma manera que los estimadores no paramétricos, éste muestra una mayor variabilidad de la dependencia para el año 1998 que para 2010:

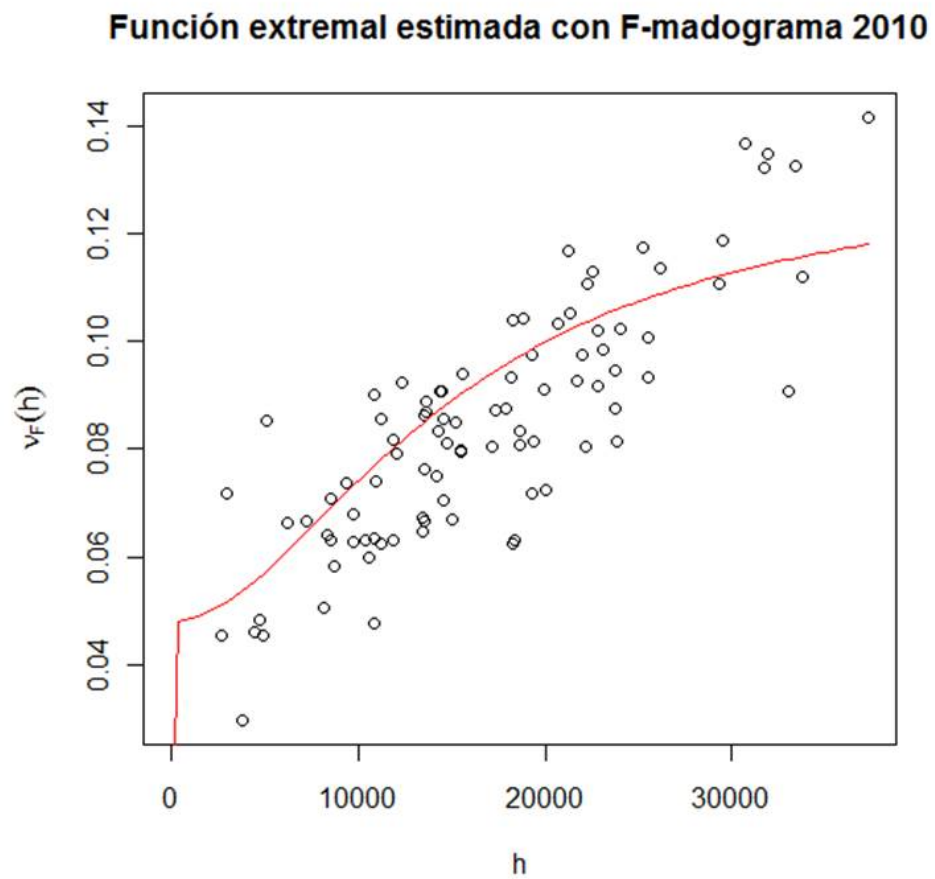
**Figura 5.8:** Función extremal estimada con el f-madograma, para el año 1998



### 5.3. Resultados

---

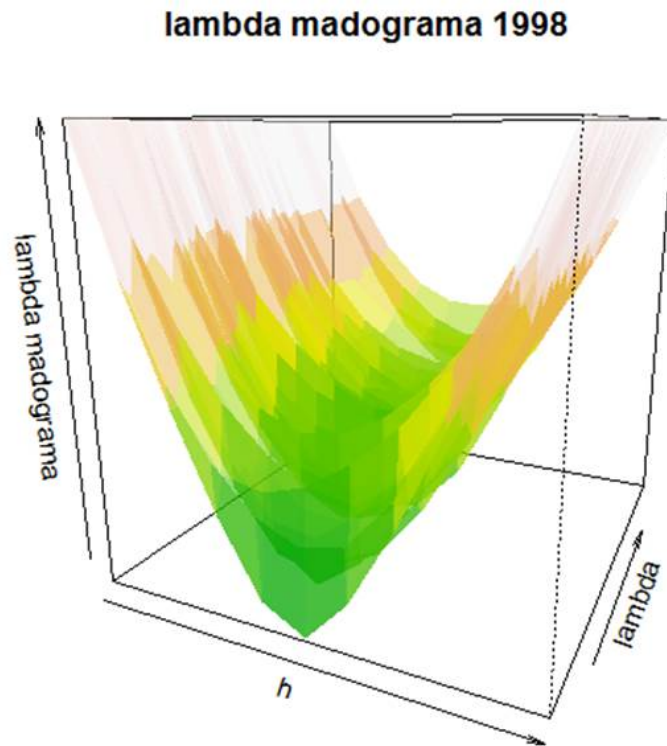
**Figura 5.9:** Función extremal estimada con el f-madograma, para el año 2010



### 5.3. Resultados

Con la dependencia estimada mediante el  $\lambda$ -madograma, con valores de  $\lambda \in (0, 1)$  se confirma que a distancias más cortas la dependencia es más fuerte y decrece al incrementar  $h$ , como puede observarse en las Figuras 5.10 y 5.11.

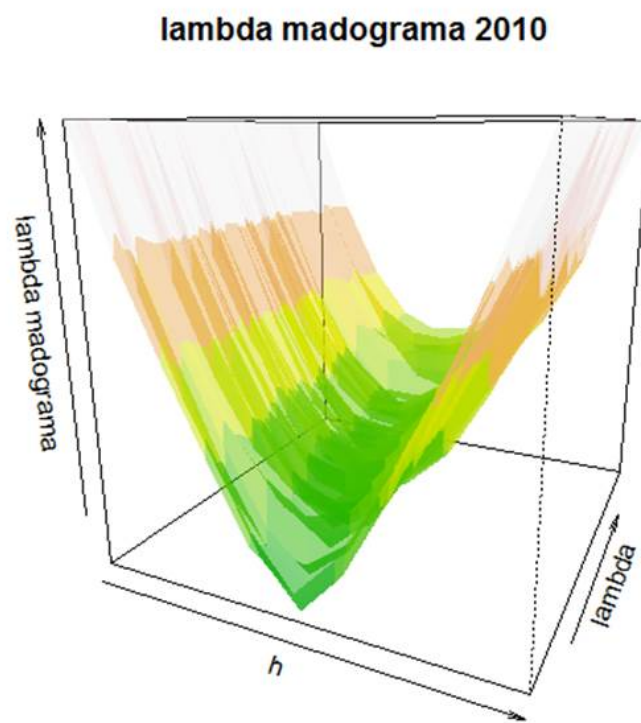
**Figura 5.10:** Función extremal estimada con el lambda madograma, para el año 1998.



### 5.3. Resultados

---

**Figura 5.11:** Función extremal estimada con el lambda madograma, para el año 2010.



### 5.3. Resultados

---

#### 3. Ajuste y selección del modelo

Con los datos transformados a distribuciones unitarias Fréchet en cada posición se estiman los parámetros de los modelos por mínimos cuadrados de Smith y de Schlather para todos los años. Para el modelo de Schlather se consideran las funciones de correlación exponencial y cauchy.

La Tabla 5.2 muestra los parámetros estimados para todos los años del modelo de Smith:

**Tabla 5.2:** Parámetros del modelo de Smith estimados por mínimos cuadrados para 1998 a 2010

Año	COV11	COV12	COV22
1998	247893172	60523011	146057156
1999	197788635	32722675	165454453
2000	209202997	50290125	181351581
2001	232918073	56907203	158182915
2002	179451149	1995284	107772100
2003	200898713	112075863	249452785
2004	314625751	90579822	280532219
2005	513926575	94531204	379788074
2006	312553441	14262492	277206941
2007	406618983	160790072	328905798
2008	560802518	75190952	424152513
2009	467501842	90305792	472556385
2010	429877946	27666732	350559991

Las Tablas 5.3 y 5.4 muestran los parámetros estimados por mínimos cuadrados para el modelo de Schlather:



### 5.3. Resultados

**Tabla 5.3:** Parámetros del modelo de Schlather con correlación exponencial estimados por mínimos cuadrados para 1998 a 2010

Modelo	Año	nugget	rango	smooth
Exp.	1998	1.566e-01	2.467e+04	2
Exp.	1999	1.499e-01	2.584e+04	1.854
Exp.	2000	6.796e-02	2.343e+04	2
Exp.	2001	1.419e-01	2.474e+04	2
Exp.	2002	1.395e-01	2.049e+04	1.822
Exp.	2003	5.371e-08	4.101e+03	2.147e-01
Exp.	2004	1.132e-01	3.589e+04	1.783
Exp.	2005	5.372e-02	3.932e+04	2
Exp.	2006	8.523e-02	3.307e+04	2
Exp.	2007	8.157e-02	3.476e+04	2
Exp.	2008	6.004e-02	4.397e+04	2
Exp.	2009	9.927e-02	4.954e+04	2
Exp.	2010	5.767e-02	3.778e+04	2

**Tabla 5.4:** Parámetros del modelo de Schlather con correlación Cauchy estimados por mínimos cuadrados para 1998 a 2010.

Modelo	Año	nugget	rango	smooth
Cauchy	1998	1.565e-01	2.462e+05	100
Cauchy	1999	1.623e-01	2.561e+05	100
Cauchy	2000	6.750e-02	2.337e+05	100
Cauchy	2001	9.053e-08	4.100e+03	2.280e-01
Cauchy	2002	8.057e-08	4.101e+03	2.965e-01
Cauchy	2003	5.371e-04	4.101e+03	2.147e-01
Cauchy	2004	1.315e-02	4.098e+03	1.382e-01
Cauchy	2005	5.222e-08	4.108e+03	9.083e-02
Cauchy	2006	4.344e-08	4.103e+03	1.329e-02
Cauchy	2007	1.707e-08	4.097e+03	1.195e-01
Cauchy	2008	5.297e-03	4.096e+03	7.958e-02
Cauchy	2009	9.919e-02	4.947e+05	9.999e+01
Cauchy	2010	3.652e-07	4.099e+03	9.734e-02

### 5.3. Resultados

Como puede observarse en las Tablas 5.3 y 5.4 para el ajuste del modelo de Schlather con correlaciones exponencial y Cauchy, el rango es creciente en la correlación exponencial y en la Cauchy no se observa tendencia alguna.

Estimación por máxima verosimilitud

En la Tabla 5.5 se muestran los mejores modelos ajustados para los años 1998-2010, la selección se hizo en base al TIC más bajo dado que los modelos son mal especificados:

**Tabla 5.5:** Mejores modelos seleccionados de acuerdo al TIC

Año	Modelos	correlación	sigma	nugget	rango	suavizamiento
1998	G.G.	Bessel	32.82825(5.783)	0.03437(6.713e-03)	4000)	39.96637(14.115e-06)
1999	Schlather	Bessel		0.3278(0.048)	5922.8726(827.38099)	1*
2000	Schlather	Whitmat		6.4467e-02(0.02672)	5.230e+03(860.64462)	2*
2001	G. G.	Whitmat	11.27892(2.02196)	0.06786(0.02688)	3000*	99.96453(1.51077)
2002	G. G.	Bessel	4.1792(0.7643)	0.1532(0.0442)	3000*)	25.3315(9.0872)
2003	G.G.	Bessel	4.1792(0.7643)	0.1532(0.0442)	3000*)	25.3315(9.0872)
2004	Schlather	Whitmat		6.874e-02(1.568e-02)	1.452e+04(2.132e+03)	1*
2005	Brown Resnik			1.924e+04(2.046e+03)	7.769e-01(4.438e-02)	
2006	Brown-Resnik				3.018e+04(3.635+03)	8.287e-01(4.686e-02)
2007	G.G.	Whitmat	6.91440(1.06009)	0.04688(0.01974)	3000*	99.98517(1.06347)
2008	G.G.	Bessel	2.88549(0.28432)	0.08784(0.02172)	4000*	22.95349(2.85915)
2009	Schlather	Whitmat		9.062e-02(1.893e-02)	2225e+04(4.045e+03)	1*
2010	G.G.	Whitmat	5.71293(0.98834)	0.04627(0.01815)	3000*)	97.60453(0.95487)

4. Determinación de tendencia por medio de los parámetros rango del modelo de Schlather, fijando el parámetro de suavizamiento.

Se ajustó el Modelo de Schlather con correlaciones , exponencial y cauchy, para esto se fijo el parámetro de suavizamiento. Los resultados se presentan en las Tablas 5.6 y 5.7:

### 5.3. Resultados

---

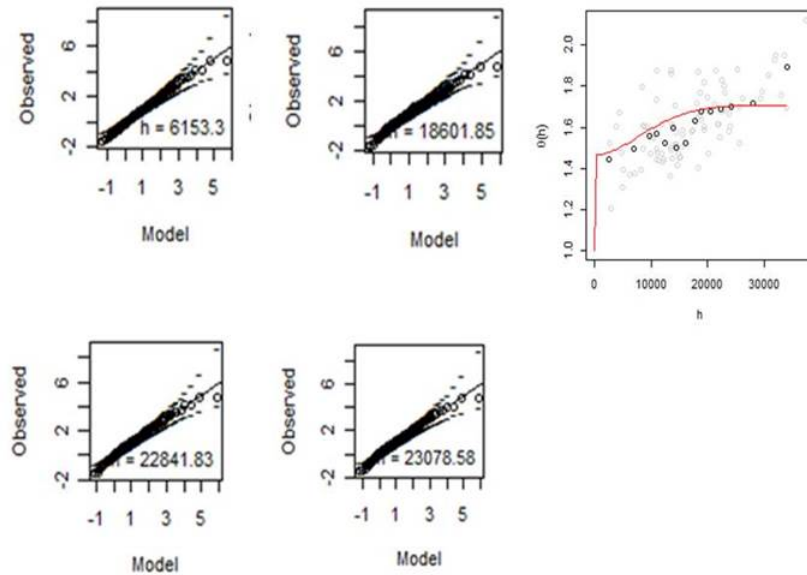
**Tabla 5.6:** Mejores modelos de Schlather ajustados para todos los años con correlación exponencial.

Año	nugget	suavizamiento	rango	TIC
1998	4351e-01(3789e-02)	2	1.257e+04(2.657e+03)	91456.27
1999	2.911e-01(3.246e-02)	2	1.428e+04(2.505e+03)	84305.77
2000	1.056e-01(2.019e-02)	2	1.370e+04(1.836e+03)	70809.62
2001	2.915e-01(3.132e-02)	2	1.479e+04(2.357e+03)	79248.86
2002	3.014e-01(3.853e-02)	2	1.460e+04(2.751e+03)	62290.34
2003	1.744e-01(2.236e-02)	2	1.601e+04(2.069e+03)	80972.26
2004	1.941e-01(0.022)	2	1.956e+04(2540.986)	88200.45
2005	1.359e-01(1.648e-02)	2	2.317e+04(2.5448e+03)	84468.04
2006	1.441e-01(2.228e-02)	2	2.257e+04(2.761e+03)	66343.38
2007	1.594e-01(2.246e-02)	2	22670(2558e+03)	68291.31
2008	1.263e-01(0.0213)	2	2.840e+04(3813.6511)	48001.89
2009	1.341e-01(2.197e-02)	2	3.227e+04(4.590e+03)	40005.15
2010	1.068e-01(1.433e-02)	2	2.599e+04(3.325e+03)	52906.58

### 5.3. Resultados

En la Fig. 5.12 se muestra el ajuste del año 1998 con correlación exponencial.

**Figura 5.12:** Coeficiente extremal con parámetros estimados por máxima verosimilitud para el año 1998.



### 5.3. Resultados

---

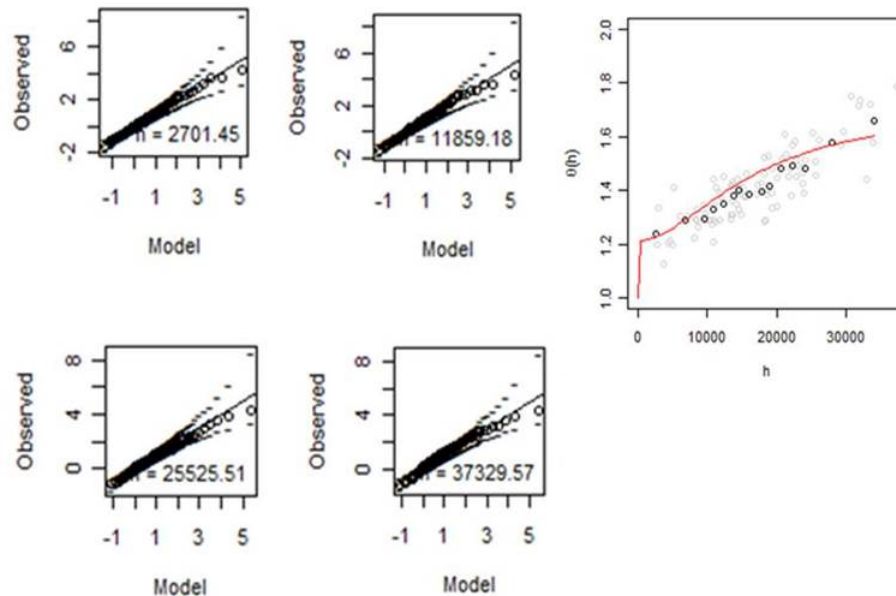
**Tabla 5.7:** Mejores modelos de Schlather ajustados para todos los años con correlación Cauchy.

Año	nugget	suavizamiento	rango	TIC
1998	0.3256(7.889e-02)	1	7378.4259(2.798e+03)	91476.15
1999	2.437e-01(4.437e-02)	1	1.027e+04(2.710e+03)	84328.08
2000	4.762e-02(3.388e-02)	1	9.933e+03(1.926e+03)	70823.86
2001	2.25e-01(4.482e-02)	1	1.030e+04(2.455e+03)	79264.16
2002	0.2228(5.578e-02)	1	9929.8594(2.821e+03)	62303.58
2003	1.777e-01(2.887e-02)	1	1.178e+04(2.102e+03)	80984.84
2004	1.454e-01(2.369e-02)	1	1.494e+04(2.525e+03)	88205.13
2005	1.024e-01(1.426e-02)	1	1.867e+04(2.502e+03)	84467.66
2006	1.022e-01(2.124e-02)	1	1.774e+04(2.663e+03)	66334.9
2007	0.128(2.152e-02)	1	18277.833(2.524e+03)	68306.14
2008	1.109e-01(1.947e-02)	1	2.451e+04(3.893e+03)	48007.38
2009	1.193e-01(1.988e-02)	1	2.816e+04(4.580e+03)	40006.43
2010	1.994e-02(1.291e-02)	1	2.226+04(3.414e+03)	52916.77

### 5.3. Resultados

La Fig 5.13 muestra el modelo ajustado para el año 2010 con correlación Cauchy.

**Figura 5.13:** Coeficiente extremal con parámetros estimados por máxima verosimilitud para el año 2010.



Como puede observarse el ajuste a los datos es bueno y la dependencia es más fuerte a distancias cortas. Puede además decirse que la variabilidad de la dependencia extremal es mayor en 1998 que en 2010.

Con los valores de los parámetros de las correlaciones Exponencial y Cauchy se ajusta una regresión lineal simple al rango, con un nivel de significancia de 0.05 se concluye que existe tendencia creciente en el rango. La tendencia se presenta a continuación en las Figuras 5.14 y 5.15 Se anexa además su ecuación de tendencia:

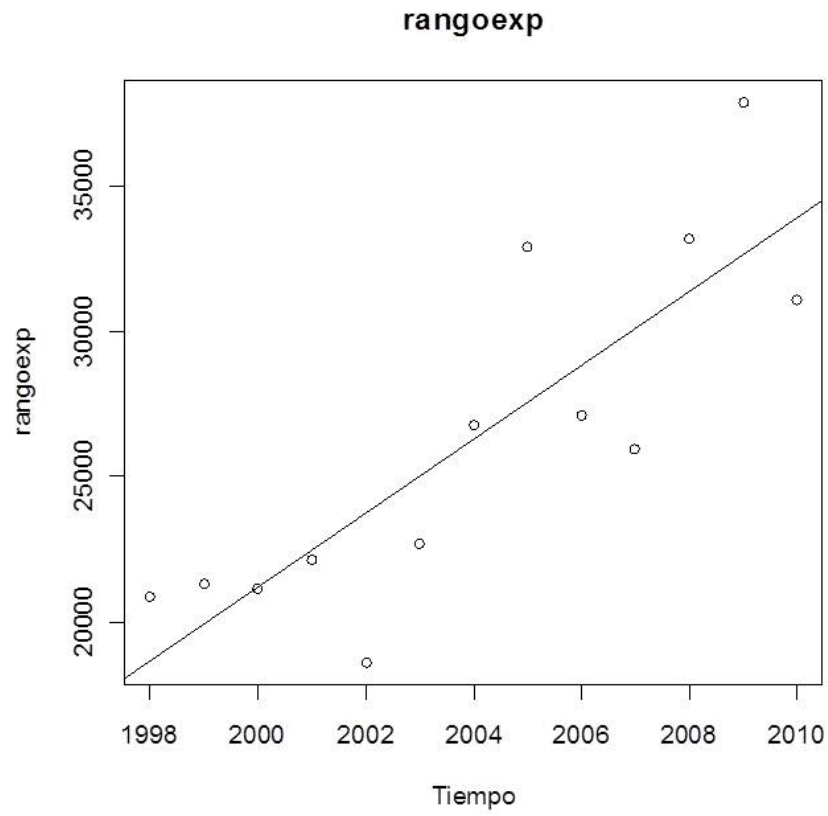
El p-valor para el intercepto y la pendiente del rango ( $3.46e-06$ ,  $3.25e-06$ ) de la correlación exponencial y la ecuación de tendencia es:  $Rango = -3019098.5 + 1516.5t$ .

El p-valor para el intercepto y la pendiente del rango ( $3.10e-06$ ,  $2.95e-06$ ) de la correlación Cauchy y la ecuación es:

### 5.3. Resultados

---

Figura 5.14: Tendencia en el rango de la correlación exponencial

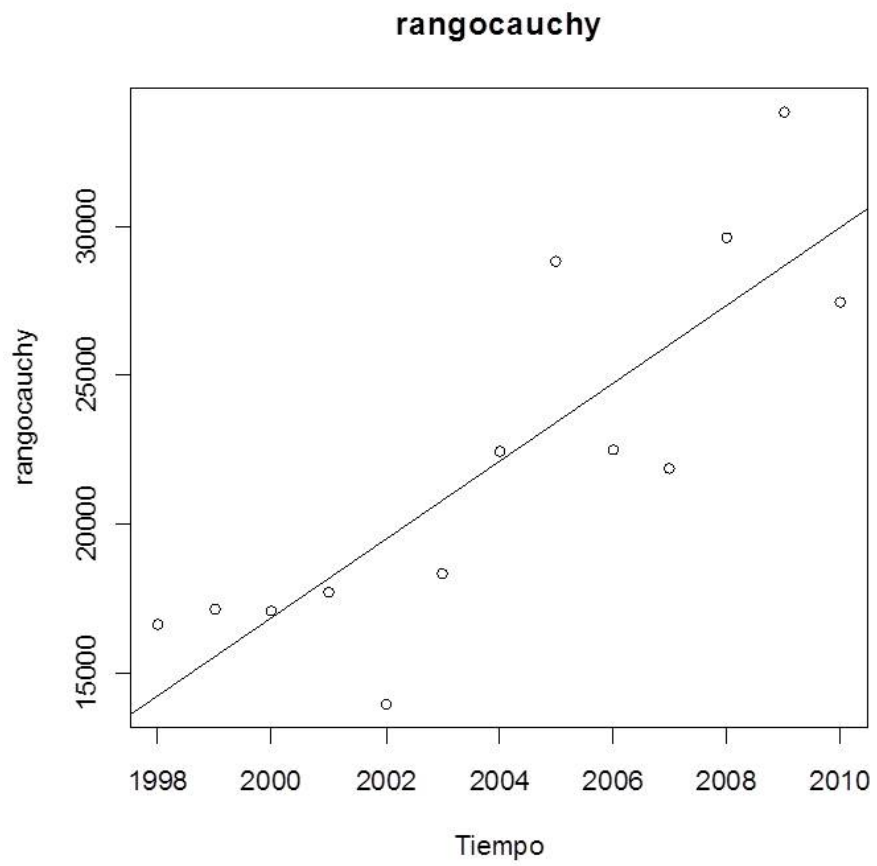


$$Rango = -3105843.2 + 1557.7t.$$

### 5.3. Resultados

---

Figura 5.15: Tendencia en el rango de la correlación Cauchy

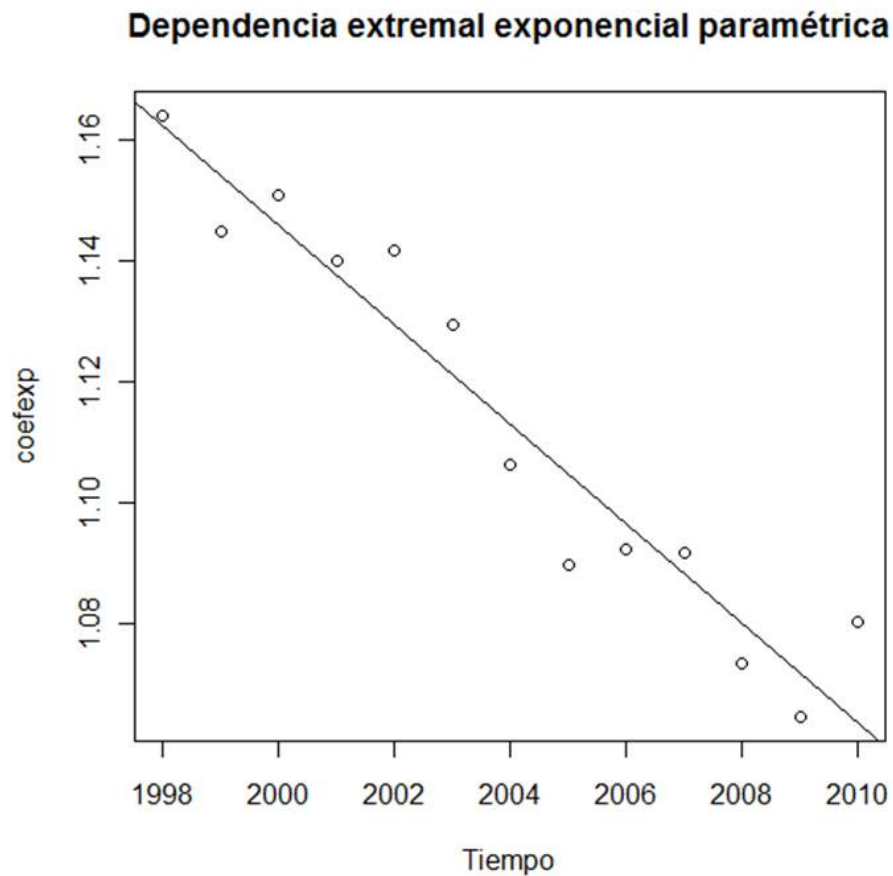




### 5.3. Resultados

Sustituyendo el rango en la función de dependencia extremal del modelo de Schlather con correlaciones exponencial y cauchy, y la distancia  $h = 2955$  de separación mínima entre dos de las estaciones meteorológicas se obtiene una tendencia decreciente en las funciones de dependencia respectivas como puede observarse en las Figuras 5.16 y 5.17:

**Figura 5.16:** Tendencia en la dependencia extremal de Schlather con corr. exponencial



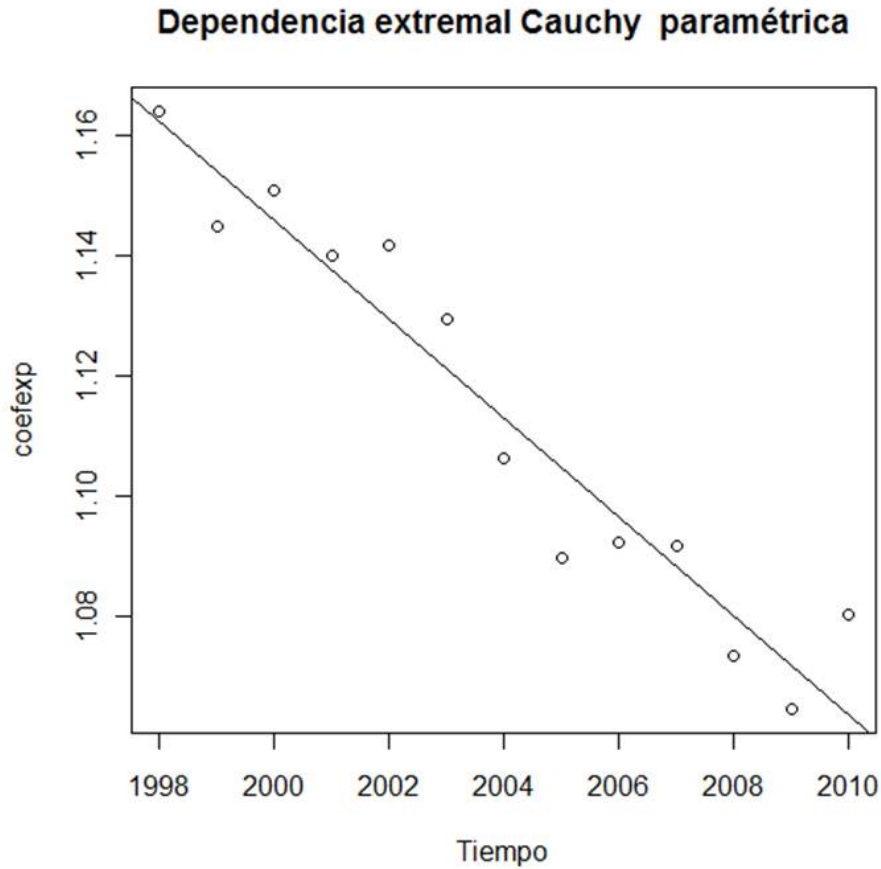
El p-valor para el intercepto y la pendiente de la dependencia extremal ( $8.41e-08$ ,  $1.65e-07$ ) del modelo Schlather con correlación exponencial y su ecuación de tendencia es:

$$\text{dependencia} = 17.5865590 - 0.0082204t.$$

Para la función de dependencia con correlación cauchy se observa también una dependencia decreciente en la Figura 5.17:

### 5.3. Resultados

Figura 5.17: Tendencia en la dependencia extremal de Schlather con corr. cauchy.



El p-valor para el intercepto y la pendiente de la dependencia extremal ( $8.41e-08$ ,  $1.65e-07$ ) del modelo de Schlather con correlación cauchy y su ecuación de tendencia es:

$$\text{Rango} = 17.5865590 - 0.0082204t.$$

Se concluye de esto que al fijar el parámetro de suavizamiento en el modelo de Schlather, el rango tiene una tendencia creciente, por lo que la función de dependencia extremal presenta una tendencia decreciente, lo cual indica una tendencia decreciente espacio-tiempo de los niveles máximos de ozono.

### 5.3. Resultados

---

Determinación de tendencia temporal por medio del ajuste del estimador de Smith de una regresión no paramétrica Loess.

Con el estimador de Smith al cual se ajustó una regresión no paramétrica Loess, fue posible determinar tendencia temporal.

La tendencia a distancias fijas ajustando una regresión no paramétrica Loess fue significativa en un 63%. Un ejemplo de tal tendencia a tres distancias diferentes se muestra en la Tabla 5.8:

**Tabla 5.8:** Tendencia con regresión Loess

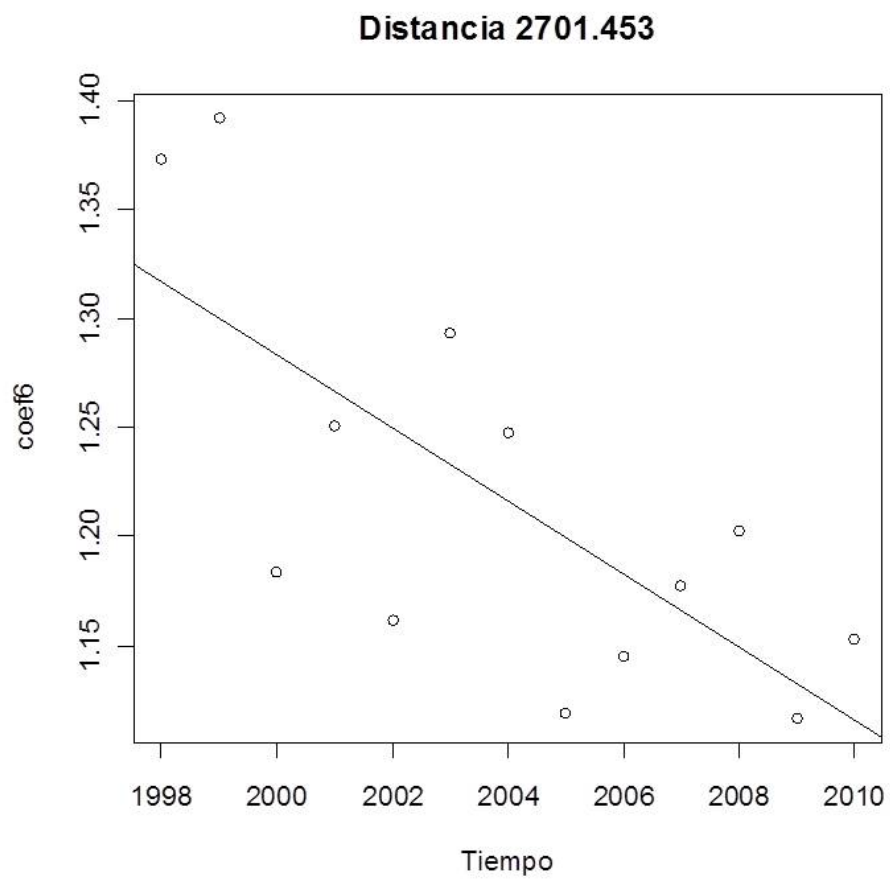
Distancia	Intercepto	Pendiente	Significancia Intercepto	Significancia Pendiente
2701.453	34.767704	-0.016742	0.00421	0.00525
11859.182	42.929388	-0.020738	0.000131	0.000173
37329.565	79.641901	-0.038814	0.00157	0.00185

Las Figuras 5.18, 5.19 y 5.20 muestra la tendencia a tales distancias.

### 5.3. Resultados

---

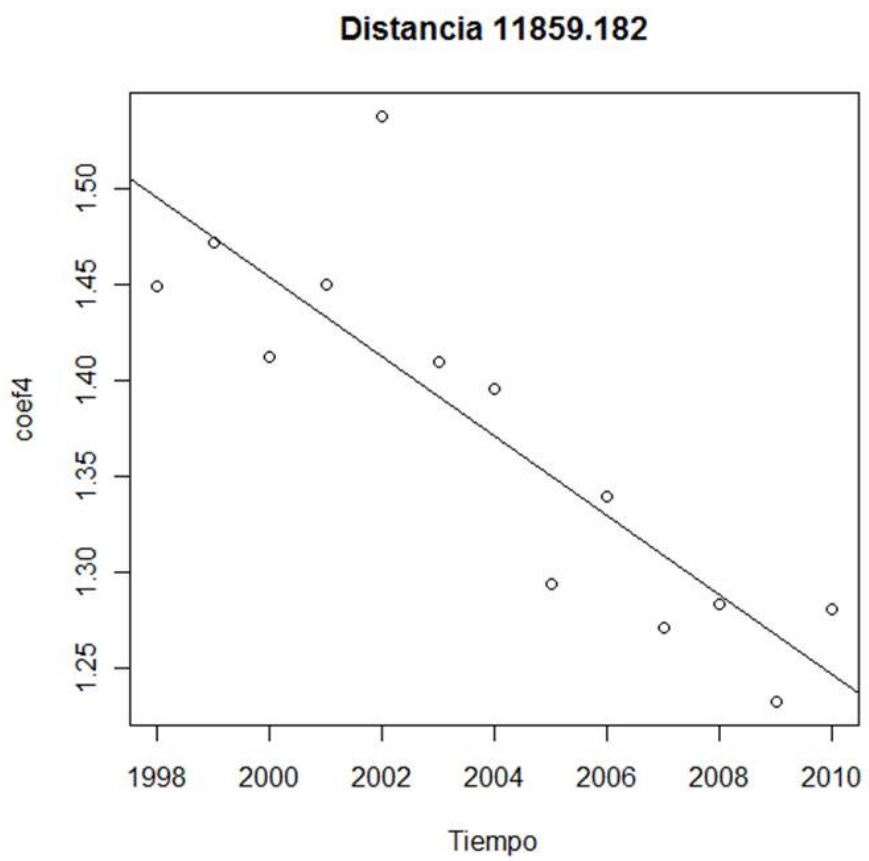
**Figura 5.18:** Tendencia del coeficiente extremal a diferentes distancias en el tiempo



### 5.3. Resultados

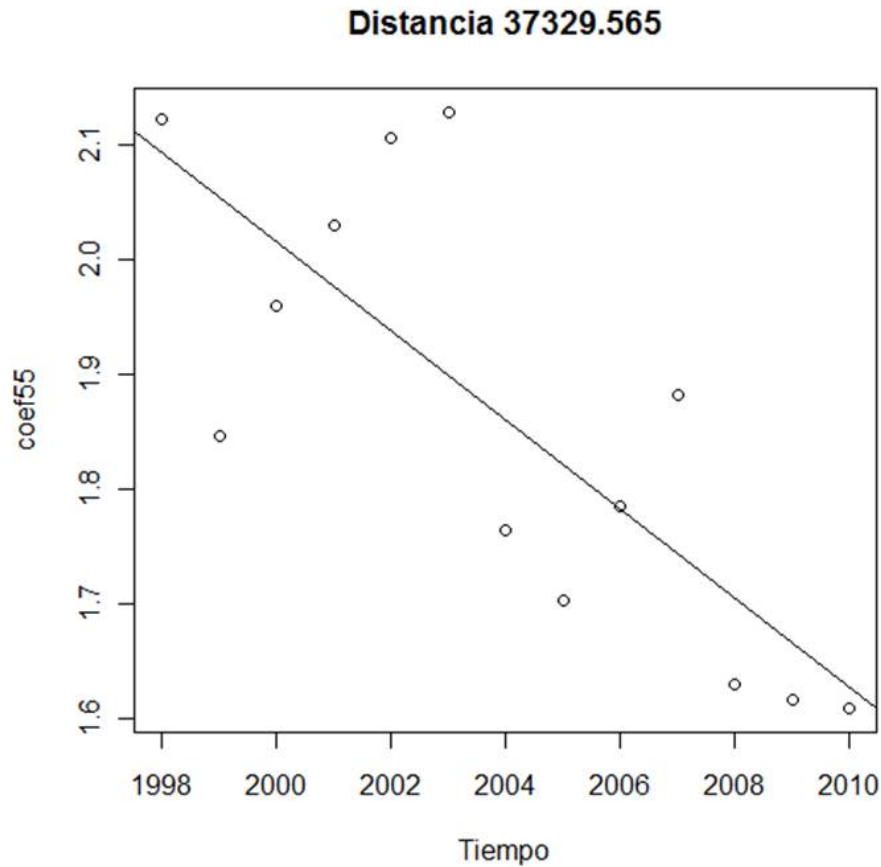
---

**Figura 5.19:** Tendencia del coeficiente extremal a diferentes distancias en el tiempo



### 5.3. Resultados

Figura 5.20: Tendencia del coeficiente extremal a diferentes distancias en el tiempo



Se concluye de esto una tendencia de los niveles máximos de ozono decreciente en espacio-tiempo.

#### 5. Confirmación de tendencia con el Modelo Lineal Vectorial Generalizado

Con el Modelo Lineal Vectorial Generalizado(VGLM) se confirma la tendencia espacio temporal existente que se muestra con los procesos Max-stable, como puede verse en la Tabla 5.9, en la cual se observa que en el parámetro  $\xi$  la longitud resulta no significativa estadísticamente en la relación lineal. Para el resto de las covariables el valor absoluto de  $t$  es mayor que 2, lo que indica una relación lineal de los parámetros con las covariables y por tanto una tendencia espacial y temporal de los parámetros de la GEV.

### 5.3. Resultados

---

**Tabla 5.9:** Parámetros de la GEV en función de covariables, longitud, latitud y tiempo

Covariable	Valor	Std. Error	t-valor
(Intercept)( $\mu$ )	9.8558e+00	1.6667e-01	59.13491
(Intercept)( $\lambda$ )	1.1154e+02	3.6271e+00	30.75086
(Intercept)( $\xi$ )	-4.7774e+01	8.9869e+00	-5.31591
lon( $\mu$ )	-7.5919e-07	3.4183e-08	-22.20938
lon( $\lambda$ )	-1.0864e-05	7.2804e-07	-14.92264
lon( $\xi$ )	2.9102e-07	1.8265e-06	0.15934
lat( $\mu$ )	-6.1783e-07	2.7760e-08	-22.25574
lat( $\lambda$ )	-3.3301e-06	6.3219e-07	-5.26763
lat( $\xi$ )	7.7049e-06	1.5731e-06	4.89795
tiempo( $\mu$ )	-4.0211e-03	7.6897e-05	-52.29246
tiempo( $\lambda$ )	-5.1138e-02	1.6633e-03	-30.74507
tiempo( $\xi$ )	1.4900e-02	4.1221e-03	3.61463

### 5.3. Resultados

En las Tablas 5.10 y 5.11 puede observarse que resultado significativa (valor absoluto de  $t$  es mayor que 2) la interacción longitud-tiempo y latitud-tiempo, lo cual confirma la tendencia espacio temporal obtenida con los modelos max-stable.

**Tabla 5.10:** Parámetros de la GEV en función de covariables, longitud, latitud, tiempo e interacción longitud-tiempo

Covariable	Valor	Std. Error	t-valor
(Intercept):( $\mu$ )	4.7540e+01	9.2801e+00	5.12276
(Intercept):( $\lambda$ )	5.5653e+02	1.9508e+02	2.85286
(Intercept):( $\xi$ )	4.3723e+02	4.9562e+02	0.88219
lon( $\mu$ )	-7.8048e-05	1.9063e-05	-4.09416
lon( $\lambda$ )	-9.2503e-04	4.0175e-04	-2.30254
lon( $\xi$ )	-1.0106e-03	1.0211e-03	-0.98977
lat( $\mu$ )	-6.2044e-07	2.7844e-08	-22.28326
lat( $\lambda$ )	-3.3857e-06	6.3118e-07	-5.36406
lat( $\xi$ )	8.0333e-06	1.5891e-06	5.05512
tiempo( $\mu$ )	-2.2816e-02	4.6287e-03	-4.92920
tiempo( $\lambda$ )	-2.7323e-01	9.7363e-02	-2.80631
tiempo( $\xi$ )	-2.2749e-01	2.4735e-01	-0.91970
lon * tiempo( $\mu$ )	3.8553e-08	9.5090e-09	4.05440
lon * tiempo( $\lambda$ )	4.5638e-07	2.0052e-07	2.27597
lon * tiempo( $\xi$ )	5.0447e-07	5.0957e-07	0.98999



### 5.3. Resultados

**Tabla 5.11:** Parámetros de la GEV en función de covariables, longitud, latitud, tiempo e interacción latitud-tiempo

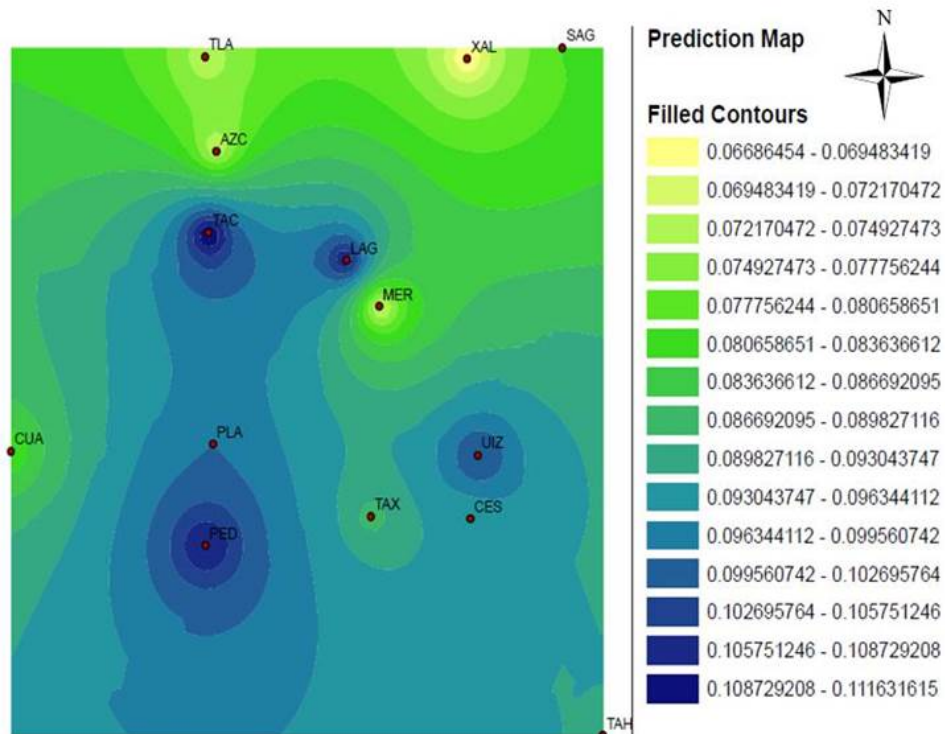
Covariable	Valor	Std. Error	t-valor
(Intercept):( $\mu$ )	1.6737e+02	3.3248e+01	5.0341
(Intercept):( $\lambda$ )	-1.6038e+03	7.4899e+02	-2.1413
(Intercept):( $\xi$ )	4.9077e+03	1.8735e+03	2.6195
lon( $\mu$ )	-7.6279e-07	3.4154e-08	-22.3342
lon( $\lambda$ )	-1.0752e-05	7.2767e-07	-14.7754
lon( $\mu$ )	1.1724e-06	1.8362e-06	0.6385
lat( $\mu$ )	-7.3975e-05	1.5488e-05	-4.7762
lat( $\lambda$ )	7.9522e-04	3.4899e-04	2.2787
lat( $\xi$ )	-2.3022e-03	8.7258e-04	-2.6384
tiempo( $\mu$ )	-8.2589e-02	1.6585e-02	-4.9799
tiempo( $\lambda$ )	8.0501e-01	3.7384e-01	2.1534
tiempo( $\xi$ )	-2.4577e+00	9.3500e-01	-2.6285
lat * tiempo( $\mu$ )	3.6590e-08	7.7259e-09	4.7361
lat * tiempo( $\lambda$ )	-3.9857e-07	1.7419e-07	-2.2882
lat * tiempo( $\xi$ )	1.1525e-06	4.3548e-07	2.6465

### 5.3. Resultados

#### Niveles de retorno

En la Fig. 5.21 se muestra un mapa de 1998 construido con Kriging para los niveles de retorno. En él puede observarse que las zonas con mayor contaminación ambiental en ese año son: Pedregal, Lagunilla, Tacuba y UAM Iztapalapa.

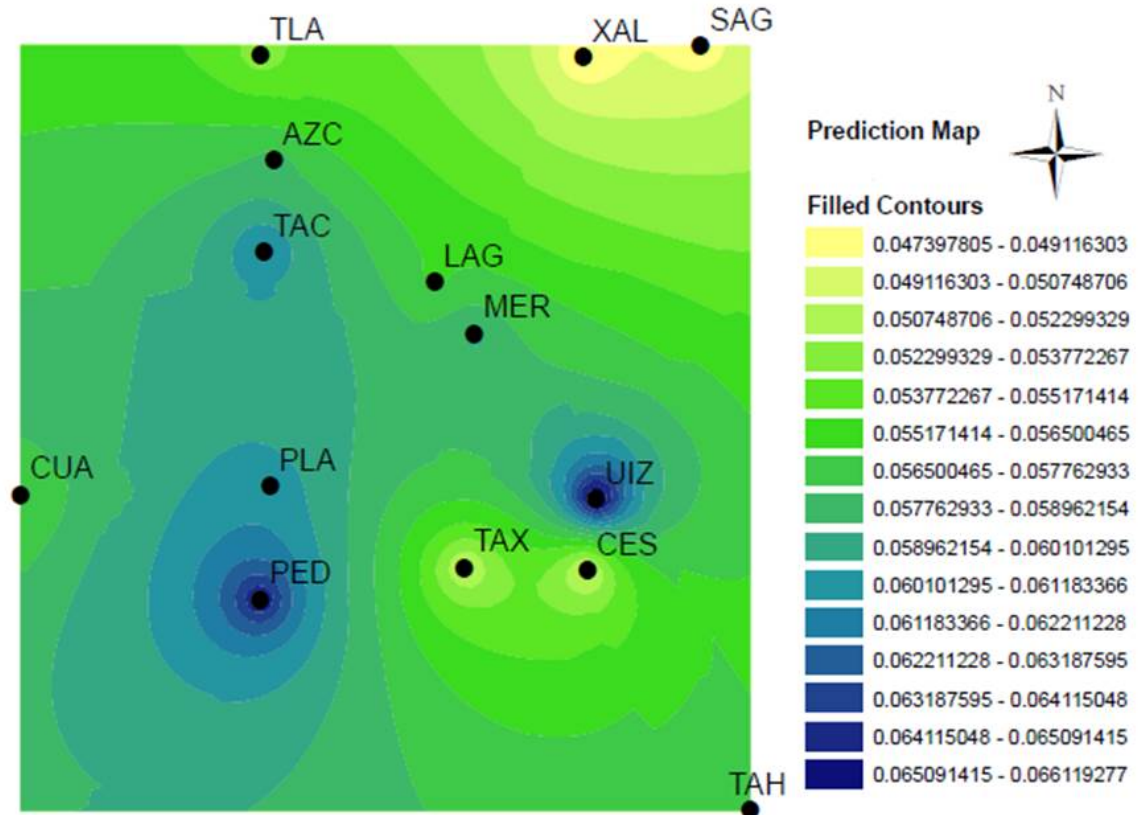
**Figura 5.21:** Kriging de los niveles de retorno para el año 1998



Para el año 2010 puede observarse en la Fig. 5.22 que las zonas con mayor contaminación ambiental son: Pedregal, Tacuba y UAM Iztapalapa.

### 5.3. Resultados

Figura 5.22: Kriging de los niveles de retorno para el año 2010



La contaminación en niveles máximos de ozono se observa decreciente.

### 5.4. Conclusiones

La metodología empleada para investigar tendencias espacio tiempo usando el modelo max-stable resultó factible de aplicar a los niveles de ozono urbano en el valle de México. Se modeló la dependencia extrema de los niveles máximos del ozono urbano y se concluye que existe una tendencia espacio temporal decreciente en el valle de Mexico. El modelo que ajusto mejor a los datos en los diferentes años según el criterio de Takeuchy(TIC) al fijar el parámetro de suavizamiento fue el de Schlather con correlaciones Exponencial y Cauchy, con este modelo se presento la tendencia por medio del rango.

El modelo VGAM permitio corroborar los resultados obtenidos usando el proceso max-stable.

Según lo que mostró el estimador de la dependencia extrema de Smith la variabilidad espacial tiene un comportamiento anisotrópico. Además los estimadores de máxima verosimilitud y f-madograma mostraron que la mayor variabilidad de la dependencia extrema se presentó en 1998.

# Referencias

- Brown, B. M. and Resnik, S. I. (1977). Extreme values of independent stochastic processes. *J. Appl. Probab.*, 14:732–739.
- Coles, S. G. (1993). Regional modelling of extreme storms via max  $\alpha$ -stable process. *J. Roy. Statist. Soc. B*, 58:329–347.
- Coles, S. G. (2001). *An Introduction to Statistical Modelling of Extreme Values*. Springer, London.
- Constanzo, J. (2009). Revista tierraamérica. *Cambio Climático*.
- Cooley, D., Naveau, P., and Poncet, P. (2006). Variograms for spatial max-stable random fields. In Bertail, P., Soulier, P., and Doukhan, P., editors, *Dependence in Probability and Statistics*, volume 187 of *Lecture Notes in Statistics*, pages 373–390. Springer New York.
- Correa, G. A., Muñoz, C. R., and García, C. G. (2000). Análisis de ozono y óxidos de nitrógeno mediante series de tiempo y regresión lineal. Technical report, XII Congreso Nacional 2000. Ciencia y Conciencia, 21 a 24 de marzo del 2000. Morelia, Mich. México.
- Davison, A. C., Padoan, S., and M., R. (2012). Statistical modelling of spatial extremes. *Statistical Science*, 27(2):161–186.
- de Haan, L. (1984). A spectral representation for max-stable processes. *The Annals of Probability*, 12(4):1194–1204.
- de Haan, L. and Pereyra, T. T. (2006). Spatial extremes: Models for the stationary case. *Annals of Statistics*, 34:146–168.
- de Haan, L. and Pickands, J., I. (1986). Stationary min-stable stochastic processes. *Probability Theory and Related Fields*, 72:477–492.
- de Haan, L. and Resnick, S. (1977). Limit theory for multivariate sample extremes. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 40:317–337.
- Genton, M. G., Ma, J., and Sang, H. (2011). On the likelihood function of gaussian max stable processes. *Biométrica*, 98:481–488.

## Referencias

---

- Huerta, G., Sanso, B., and Stroud, R. J. (2004). A spatiotemporal model for Mexico city ozone levels. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 53(2):231–248.
- IPCC (2007). Managing the risks of extreme events and disasters to advance climate change adaptation. Technical report.
- Jenkinson, A. F. (1955). The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society*, 87:158–171.
- Kabluchko, Z., Schlather, M., and de Haan, L. (2009). Stationary max-stable fields associated to negative definite functions. *The Annals of Probability*, 37(5):2042–2065.
- Katz, R. and Brown, B. (1992). extreme events in a changing climate: Variability is more important than averages. *Climatic Change*, 21:289–302.
- Kotz, S. and Nadarajah, S. (2000). *Extreme Value Distributions: Theory and Applications*.
- Min, S. K., Zhang, X., Zwiers, F. W., and Hegerl, G. (2011). Human contribution to more-intense precipitation extremes. *Nature*, 470:378–381.
- Naveau, P., Guillou, A., Cooley, D., and Diebolt, J. (2009). Modelling pairwise dependence of maxima in space. *Biometrika*, 96(1):1–17.
- Padoan, S. A., Ribatet, M., and Sisson, S. A. (2010). Likelihood-based inference for max-stable processes. *Journal of the American Statistical Association*, 105(489):263–277.
- R DEVELOPMENT CORE TEAM (2012). R: A language and environment for statistical computing. *Vienna, Austria: R Foundation for Statistical Computing, ISBN 3-900051-07-0*.
- Reyes, H. J., Vaquera, H. H., and Villaseñor, J. A. (2010). Estimation of trends in high urban ozone levels using the quantiles of (gev). *Environmetrics*, 21(5):470–481.
- Schlather, M. (2002). Models for stationary max-stable random fields. *Extremes*, 5:33–44. 10.1023/A:1020977924878.
- Schlather, M. and Tawn, J. A. (2003). A dependence measure for multivariate and spatial extreme values: Properties and inference. *Biometrika*, 90(1):139–156.
- Smith, R. L. (1990). Max-stable processes and spatial extremes. *Unpublished*.
- Varin, C. and Vidoni, P. (2005). A note on composite likelihood inference and model selection. *Biometrika*, 92:519–528.
- Von Mises, R. (1954). La distribution de la plus grande de  $n$  valeurs. *In: American Mathematical Society, Providence, Rhode Island, USA, pages 271–294*.

## Referencias

---

Yee, T. and Stephenson, A. (2007). Vector generalized linear and additive extreme value models. *Extremes*, 10:1–19.

# Apéndice A

## Instrucciones de R

INSTRUCCIONES DEL PAQUETE SPATIAL EXTREMES DE R Con esta instrucción se carga la librería Spatial Extremes de R la cual se utiliza para el análisis espacial de extremos con procesos max-stable:

```
library(SpatialExtremes)
```

Instrucciones para la lectura de datos y coordenadas de las 14 posiciones, así como la asignación de etiquetas a las columnas de cada una de las localizaciones:

```
dato4 <- matrix(scan("199814o.txt", 0), nrow = 119, ncol = 14, byrow = "T")
```

```
dat <- matrix(scan("coordozono.txt", 0), nrow = 14, ncol = 2, byrow = "T")
```

```
colnames(dat) = c("lon", "lat")
```

*dato4* archivo que contiene los datos de todas las posiciones en forma de una matriz para que R los pueda leer, las columnas corresponden a cada una de las posiciones y las filas contienen las observaciones correspondientes en cada una de las localizaciones.

*dat* contiene las coordenadas de la posiciones, las columnas corresponden al número de localizaciones y las filas a las coordenadas en cada posición.

Instrucciones para leer el archivo de texto y llevarlo a un formato matricial, éstos archivos contienen 14 parámetros de las distribuciones GEV (localización, escala y forma) estimados en cada posición.

```
l <- matrix(scan("l1998.txt", 0), nrow = 14, ncol = 1, byrow = "T")
```



## A. Instrucciones de R

---

```
s <- matrix(scan("s1998.txt", 0), nrow = 14, ncol = 1, byrow = "T")
```

```
f <- matrix(scan("f1998.txt", 0), nrow = 14, ncol = 1, byrow = "T")
```

Con la siguiente instrucción se transforman los datos a que tengan una distribución unitaria Fréchet:

```
n.site <- 14
```

```
for(i in 1 : n.site)
```

```
dato4[, i] <- gev2frech(dato4[, i], l[i], s[i], f[i])
```

$l[i]$ ,  $s[i]$ ,  $f[i]$  son los parámetros estimados de la GEV en cada una de las posiciones.

`fitmaxstab` es la instrucción para ajustar un proceso max stable a los datos con correlación `cov.mod`, la cual para el modelo de Smith es "gauss"; para el modelo Schlather "Whitmat", "Exponencial", "Cauchy", o "Bessel", de las familias de correlación whittle Matern, Powered Exponential, Cauchy y Bessel; Para el modelo Gaussiano Geométrico se antecede una `g`, ejemplo "gwhitmat". Los datos han sido anteriormente transformados a una distribución unitaria Fréchet. Como resultado se obtienen los parámetros de dependencia y sus errores estándar respectivos (range, nugget y suavizamiento), así como la matriz de varianzas covarianzas.

```
fitted1 <- fitmaxstab(dato4, dat, cov.mod = "powexp", smooth = 2)
```

`cov.mod = "powexp"` es el modelo de correlación utilizado.

`smooth = 2` es el parámetro de suavizamiento.

Gráfica el ajuste del proceso max stable, usa la salida de la función `fitmaxstab`.

```
plot(fitted1)
```

Estimadores del madograma, `fmadograma` y `lambda madograma`.

Estima el `fmadograma` del modelo que ha sido ajustado (usa la salida de la función `fitmaxstab`) con marginales unitarias fréchet, y gráfica el coeficiente extremal estimado en función del `f-madograma`, junto con el `f-madograma` (`which = c("ext", "mado")`):

```
fmadogram(dato4, dat, fitted1, which = c("ext", "mado"), main = "Coeficiente extremal  
estimado con F - madograma 2010")
```

## A. Instrucciones de R

---

Estima y gráfica el madograma y coeficiente extremal (*which* = ("*mado*", "*ext*")) del modelo max stable ajustado (usa la salida de la función *fitmaxstab*), esta instrucción es utilizada cuando se ajustan modelos con datos transformados a que cumplan que su parámetro de forma es menor o igual a 0.5.

```
madogram(dato4, dat, fitted = fitted1, which = ("mado", "ext"),  
  
main = "Madograma 1998 Cov. cauchy")
```

Estima y gráfica el lambda madograma de datos transformados a unitarias Fréchet. Si los datos no han sido transformados, la instrucción *marge* indica la estimación de parámetros de la GEV por máxima verosimilitud y transformación de los datos a que tengan distribución unitaria Fréchet:

```
lmadogram(dato4, dat, n.lambda = 11, xlab = "h", ylab = "lambda", zlab = "lambda madograma",  
marge = "mle", col = terrain.colors(50, alpha = 0.5), theta = 45, phi = 15, main =  
"lambda madograma 1998")
```

Estima los parámetros de dependencia espacial de un proceso max stable por mínimos cuadrados, la instrucción *marge* = "*frech*" indica que los datos ya han sido transformados a unitarias Fréchet, de lo contrario puede utilizarse *marge* = "*mle*" para la estimación de los parámetros de la GEV en cada posición por máxima verosimilitud y transforma posteriormente los datos a unitarias Fréchet:

```
lsmxstab(dato4, dat, cov.mod = "gauss", marge = "frech", iso = FALSE, weighted =  
TRUE)
```

Gráfica el comparativo de los estimadores extremales paramétricos (salida de la función *fitmaxstab*) y no paramétricos Smith y Schlather y Tawn:

```
qqextcoeff(fitted1, estim = "Smith", main = "Comparativo de estimadores MLE y Smith 2010")
```

```
qqextcoeff(fitted1, estim = "ST", main = "Comparativo de estimadores MLE y Schlather y Tawn")
```

Esta instrucción extrae el logaritmo de la máxima verosimilitud de un proceso max-stable ajustado:

```
logLik(fitted1)
```

La siguiente instrucción calcula el criterio de información de Takeuchi:

```
TIC(fitted1)
```

Las funciones *fitcovariance* y *fitcovmat* estiman la función covarianza para el modelo Schlather y la matriz covarianza para el modelo Smith respectivamente, usando

## A. Instrucciones de R

---

estimadores no paramétricos. Para realizar esto se transforman los datos a que tengan distribución unitaria Fréchet y el ajuste se realiza por mínimos cuadrados. Las instrucciones son las siguientes:

```
fitcovariance(dato4, dat, "powexp", marge = "frech", weighted = TRUE)
```

```
fitcovmat(dato4, dat, marge = "frech", weighted = TRUE)
```

Con la siguiente instrucción se calculan funciones de covarianza de un modelo max-stable ajustado:

```
covariance(fitted1, nugget = 0.09050, range = 31070, smooth = 2, main = "Covarianza  
exponencial 2010")
```

```
fitcovariance(dato4, dat, "whitmat")
```

### INSTRUCCIONES DE LA LIBRERIA VGAM

```
library(VGAM)
```

```
dato1 <- as.data.frame(dato1)
```

```
colnames(dato1) = c("tiempo", "ozono", "lon", "lat")
```

```
cm = list("Intercept" = diag(3), "lat" = rbind(0, 1, 0), "lon" = rbind(0, 1, 0), tiempo =  
rbind = (0, 1, 0))
```

```
(fit1 <- vglm(ozono lon + lat + tiempo, family = gev(llocation = "identity", lscale =  
"logit", lshape = "logof f", zero = NULL), data = dato, constraints = cm))
```

```
(fit1 <- vglm(ozono cbind(lon, lat, tiempo), family = gev(llocation = "identity", lscale =  
"logit", lshape = "logof f", zero = 2 : 3), data = dato1, trace = TRUE))
```

# Apéndice B

## Tablas de ajuste de modelos de todos los años

Tabla B.1: Modelos ajustados para el año 1998

Modelos	correlación	$\sigma_{11}(m)$	$\sigma_{12}(m)$	$\sigma_{22}(m)$		TIC
Smith	Anis.					
Modelos	correlación	sigma	nugget	rango	suavizamiento	TIC
Schlather	Whitmat		0.3722(4.615e-02)	4372.7133(1.191e+03)	2*	91461.64
Schlather	Exp.		7.076e-01(8.114e-02)	2.535e+04(8.251e+03)	2(1.241)	91539.16
Schlather	Bessel		0.48(0.04271)	4352.86(759.93227)	2*	91448.9
Schlather	Cauchy		3.811e-01(4.495e-02)	1.373e+04(3.822e+03)	2*	91465.01
G.G.	Whitmat	13.62527(2.98433)	0.08388(0.04082)	3000*	99.98093(2.21080)	91317.64
G.G.	Exp.					
G.G.	Bessel	32.82825(5.783)	0.03437(6.713e-03)	4000*	139.96637(14.115e-06)	91315.69
Brown-Resnik				4226.3769(406.66708)	0.5847(0.05208)	91333.33

## B. Tablas de ajuste de modelos de todos los años

**Tabla B.2:** Modelos ajustados para el año 1999

Modelos	correlación	$\sigma_{11}(m)$	$\sigma_{12}(m)$	$\sigma_{22}(m)$		TIC
Smith	Anis.	41973070(1840908)	21380708(1231822)	26589373(1161691)		85561.15
Modelos	correlación	sigma	nugget	rango	suavizamiento	TIC
Schlather	Whitmat		0.2037(5.051e-02)	7866.7415(2040)	1*	84323.72
Schlather	Exp.		5.178e-01(6.929-02)	2.506e+04(5.975e+03)	1.935(7.808e-01)	84403.12
Schlather	Bessel		0.3278(0.03481)	5922.8726(827.38099)	1*	84297.02
Schlather	Cauchy		2.37e-01(4.437e-02)	1.027e+04(2.710e+03)	1*	84328.08
G.G.	Whitmat	9.1311(1.26011)	0.1021(0.03104)	3000*	50.9156(2.53889)	84758.9
G.G.	Exp.	3.161(4.460e-01)	2.221e-01(3.207e-02)	1.603e+04(2.017e+03)	2(1.957e-01)	84782.18
G.G.	Bessel	9.0816(1.07973)	0.1031(0.02622)	4000*	27.2010(2.76939)	84758.74
Brown-Resnik				5132.2040(367.40443)	0.6707(0.04352)	84770.9

**Tabla B.3:** Modelos ajustados para el año 2000

Modelos	correlación	$\sigma_{11}(m)$	$\sigma_{12}(m)$	$\sigma_{22}(m)$		TIC
Smith	Anis.	47567963(2173282)	1473705(783108)	28782212(506417)		71988.56
Modelos	correlación	sigma	nugget	rango	suavizamiento	TIC
Schlather	Whitmat		6.467e-02(0.02672)	5.230e+03(860.64462)	2*	70810.97
Schlather	Exp.		6.958e-02(5.065e-02)	1.348e+04(1.793e+03)	1.716(4.283e-01)	79253.9
Schlather	Bessel		0.1494(0.02034)	5734.5208(626.16331)	1*	70812.28
Schlather	Cauchy		7.522e-02(2.498e-02)	1.649e+04(2.669e+03)	2*	70812.48
G.G.	Whitmat	8.59241(1.18651)	0.06802(0.02319)	3000*	45.23153(1.90558)	71262.3
G.G.	Exp.	3.684(0.8438)	1.337e-01(0.0324)	2.099e+04(3467.7860)	2(0.1978)	71280.41
G.G.	Bessel	5.2866(1.31608)	0.1042(0.01203)	4000*	11.9277(3.83508)	71263.08
Brown-Resnik				7652.2825(556.48465)	0.8487(0.04847)	71290.33

## B. Tablas de ajuste de modelos de todos los años

**Tabla B.4:** Modelos ajustados para el año 2001

Modelos	correlación	$\sigma_{11}(m)$	$\sigma_{12}(m)$	$\sigma_{22}(m)$		TIC
Smith	Anis.	57803975(2343191)	22907457(919047)	34081664(854027)		80119.55
Modelos	correlación	sigma	nugget	rango	suavizamiento	TIC
Schlather	Whitmat		0.2911(0.2161)	743.8771(101.9845)	99.2905(45.0101)	79252.26
Schlather	Exp.		2.863e-01(0.0680)	1.469e+04(0.0680)	1.987(0.6358)	79253.9
Schlather	Bessel		0.3315(0.0328)	6134.6653(794.8325)	1*)	79243.42
Schlather	Cauchy		2.235e-01(4.482e-02)	1.030e+04(2.455e+03)	1*	79264.16
G.G.	Whitmat	11.27892(2.02196)	0.06786(0.02688)	3000*	99.96453(1.51077)	78992.97
G.G.	Exp	2.1518(0.18125)	0.1727(0.04944)	13000	1.4671(0.28803)	79039.45
G.G.	Bessel	15.58295(2.90383)	0.04953(0.02001)	4000*	79.67777(1.05958)	78993.22
Brown-Resnik				7681.9804(702.67067)	0.6605(0.05029)	79005.88

**Tabla B.5:** Modelos ajustados para el año 2002

Modelos	correlación	$\sigma_{11}(m)$	$\sigma_{12}(m)$	$\sigma_{22}(m)$		TIC
Smith	Anis.	64472941(5031768)	-13427470(3106186)	28101083(1962188)		63013.89
Modelos	correlación	sigma	nugget	rango	suavizamiento	TIC
Schlather	Whitmat		0.1812(6.021e-02)	(7680.6866(2.087e+03)	1*	62298.17
Schlather	Exp.		3.003e-01(7.444e-02)	1.459e+04(2.317e+03)	1.997(7.355e-01)	62295.43
Schlather	Bessel		0.341(0.0392)	6058.490(915.1075)	1*	62283.9
Schlather	Cauchy		2.606e-01(4.368e-02)	1.709e+04(3.946e+03)	2*	62296.17
G.G.	Whitmat	4.151(0.78732)	0.153(0.04539)	3000*)	26.492(9.29007)	62148.61
G.G.	Exp	2.942(9.569e-01)	1.626e-01(6.926e-02)	2.251e+04(7.960e+03)	1.614(4.132e-01)	62155.33
G.G.	Bessel	4.1792(0.7643)	0.1532(0.0442)	3000*	25.3315(9.0872)	62148.88
Brown-Resnik				8506(895.85347)	0.6962(0.05978)	62155.26

## B. Tablas de ajuste de modelos de todos los años

**Tabla B.6:** Modelos ajustados para el año 2003

Modelos	correlación	$\sigma_{11}(m)$	$\sigma_{12}(m)$	$\sigma_{22}(m)$		TIC
Smith	Anis.	96268979(6425843)	51274658(4654914)	72412323(4180394)		81503.11
Modelos	correlación	sigma	nugget	rango	suavizamiento	TIC
Schlather	Whitmat		0.2102(0.02379)	(6528.3585(727.8551)	1*	80970.71
Schlather	Exp.		1.382e-01(0.0478)	1.571e+04(2055.9113)	1.720(0.4210)	80977.35
Schlather	Bessel		0.2102(0.02379)	6528.3585(727.8551)	1*	80970.71
Schlather	Cauchy		1.177e-01(2.887e-02)	1.178e+04(2.102e+03)	1*	80984.84
G.G.	Whitmat	10.55466(1.75129)	0.04172(0.01682)	3000*	99.78087(0.83410)	80560.67
G.G.	Exp	1.8001(0.17011)	0.1669(0.01622)	15000*	2(0.20298))	80635.4
G.G.	Bessel	4.1792(0.7643)	0.1532(0.0442)	3000*	25.3315(9.0872)	62148.88
Brown-Resnik				1.236e+04(1.097e+03)	8.257e-01(5.267e-02)	80597.13

**Tabla B.7:** Modelos ajustados para el año 2004

Modelos	correlación	$\sigma_{11}(m)$	$\sigma_{12}(m)$	$\sigma_{22}(m)$		TIC
Smith	Anis.	79960026(2979303)	42381724(2089444)	71072786(2139867)		89394.93
Modelos	correlación	sigma	nugget	rango	suavizamiento	TIC
Schlather	Whitmat		1.088e-01(2.753e-02)	1.151e+04(2.047e+03)	1*	88200.9
Schlather	Exp.		1.632e-01(3.576e-02)	1.933+04(2.784e+03)	1.733(3.445e-01)	88202.88
Schlather	Bessel		0.219(0.02387)	7735.676(886.52750)	1*	88201.94
Schlather	Cauchy		1.454-01(2.369e-02)	1.494e+04(2.525e+03)	1*	88205.13
G.G.	Whitmat	3.0351(0.50948)	0.1328(0.01475)	3000*	20.8669(5.24507)	88207.28
G.G.	Exp	2.801(9.543e-01)	1.248e-01(4.851e-02)	2.568e+04(7.980e+03)	1.776(1.935e-01)	88210.29
G.G.	Bessel	2.9042(0.87686)	0.1402(0.03516)	3680.0791(555.23247)	11.7852(4.34631)	88212.28
Brown-Resnik				1.247e+04(958.27925)	7.955e-01(0.04249)	88215.69

## B. Tablas de ajuste de modelos de todos los años

**Tabla B.8:** Modelos ajustados para el año 2005

Modelos	correlación	$\sigma_{11}(m)$	$\sigma_{12}(m)$	$\sigma_{22}(m)$		TIC
Smith	Anis.	87947740(3475064)	261555061(2291245)	62387990(1223601)		86137.55
Modelos	correlación	sigma	nugget	rango	suavizamiento	TIC
Schlather	Whitmat		6.874e-02(1.568e-02)	(1.452e+04(2.132e+03)	1*	84462.3
Schlather	Exp.		1.359e-01(1.648e-02)	2.317e+04(2.548e+03)	2*	84468.04
Schlather	Bessel		0.1521(0.01837)	8925.7238(893.79548)	1*	84470.86
Schlather	Cauchy		1.024e-01(1.426e-02)	1.867e+04(2.502e+03)	1*	84467.66
G.G.	Whitmat	3.9421(0.3707)	0.1022(0.0232)	3000*	38.9229(2.3882)	84844.91
G.G.	Exp	1.55529(0.11028)	0.07855(0.03257)	14000*	1.39874(0.17390)	84895.75
G.G.	Bessel	3.1053(0.85029)	0.1262(0.03326)	4033.7711(597.09196)	14.1453(4.07709)	84846.95
Brown-Resnik				1.471e+04(1.177e+03)	8.031e-01(3.626e-02)	84840.79

**Tabla B.9:** Modelos ajustados para el año 2006

Modelos	correlación	$\sigma_{11}(m)$	$\sigma_{12}(m)$	$\sigma_{22}(m)$		TIC
Smith	Anis.	88063628(3251041)	-1050327(2091620)	77634711(1567717)		67364.51
Modelos	correlación	sigma	nugget	rango	suavizamiento	TIC
Schlather	Whitmat		6.915e-02(2.377e-02)	1.384e+04(2.252e+03)	1*	66329.35
Schlather	Exp.		1.441e-01(2.228e-02)	2.257e+04(2.761e+03)	2*	66343.38
Schlather	Bessel		0.1655(0.02412)	8819.0657(977.70264)	1*	66349.22
Schlather	Cauchy		1.022e-01(2.124e-02)	1.774e+04(2.663e+03)	1*	66334.9
G.G.	Whitmat	1.6640(0.31051)	0.1712(0.02027)	3000*	13.5735(3.99271)	66219.47
G.G.	Exp	1.37436(0.10963)	0.07047(0.04069)	16000*	1.32165(0.22215)	66216.89
G.G.	Bessel	3.5467(0.3761)	0.1031(0.0281)	4000	25.9092(3.3126)	66227.46
Brown-Resnik				1.924e+04(2.046e+03)	7.769e-01(4.438e-02)	66194.38



## B. Tablas de ajuste de modelos de todos los años

**Tabla B.10:** Modelos ajustados para el año 2007

Modelos	correlación	$\sigma_{11}$ (m)	$\sigma_{12}$ (m)	$\sigma_{22}$ (m)		TIC
Smith	Anis.	110355570(2889700)	37651548(2325851)	85760097(1045318)		69138.47
Modelos	correlación	sigma	nugget	rango	suavizamiento	TIC
Schlather	Whitmat		0.1583(0.1005)	1365.9634(131.5332)	69.0806(15.3416)	68294.63
Schlather	Exp.		1.597e-01(2.928e-02)	2.268e+04(2.880e+03)	1.998(3.942e-01)	68296.26
Schlather	Bessel		0.1699(0.02327)	6956.6615(730.92201)	2*	68285.3
Schlather	Cauchy		0.128(2.152e-02)	18277.833(2.524e+03)	1*	68306.14
G.G.	Whitmat	6.91440(1.06009)	0.04688(0.01974)	3000*	99.98517(1.06347)	67840.45
G.G.	Exp	1.220(0.1283)	7.127e-02(0.0520)	1.328e+04(0.2634)	1.330(0.2634)	68004.1
G.G.	Bessel	16.0270(2.723049)	0.0209(0.009518)	4000*	139.3326(0.393348)	67836.8
Brown-Resnik				1.921e+04(1.698e+03)	8.284e-01(3.841e-02)	67895.11

**Tabla B.11:** Modelos ajustados para el año 2008

Modelos	correlación	$\sigma_{11}$ (m)	$\sigma_{12}$ (m)	$\sigma_{22}$ (m)		TIC
Smith	Anis.	135993596(5832014)	36669307(2169775)	104294054(2839065)		48844.03
Modelos	correlación	sigma	nugget	rango	suavizamiento	TIC
Schlather	Whitmat		0.1263(0.09129)	1553.4104(208.44773)	84.2173(11.90134)	48005.62
Schlather	Exp.		1.265e-01(2.718e-02)	2.842e+04(4.827e+03)	2(4.162e-01)	48006.71
Schlather	Bessel		1.335e-01(2.257e-02)	1.063e+04(1.318e+03)	1*	48001
Schlather	Cauchy		1.109e-01(1.947e-02)	2.451e+04(3.893e+03)	1*	48007.38
G.G.	Whitmat	5.77821(0.92979)	0.04601(0.01858)	3000*	99.99266(1.09846)	47963
G.G.	Exp	1.713(4.155e-01)	1.309e-01(3.129e-02)	2.645e+04(4.552e+03)	1.968(1.715e-01)	47967.72
G.G.	Bessel	2.88549(0.28432)	0.08784(0.02172)	4000*	22.95349(2.85915)	47961.88
Brown-Resnik				2.432e+04(2.962e+03)	8.331e-01(4.483e-02)	47971.48

## B. Tablas de ajuste de modelos de todos los años

**Tabla B.12:** Modelos ajustados para el año 2009

Modelos	correlación	$\sigma_{11}(m)$	$\sigma_{12}(m)$	$\sigma_{22}(m)$		TIC
Smith	Anis.	115670813(4496805)	910549(2970464)	96988502(1708155)		41177.95
Modelos	smallcorrelación	sigma	nugget	rango	suavizamiento	TIC
Schlather	Whitmat		9.062e-02(1.893e-02)	2.225e+04(4.045e+03)	1*	40004.96
Schlather	Exp.		1.341e-01(2.197e-02)	3.227e+04(4.590e+03)	2*	40005.15
Schlather	Bessel		0.1343(0.07414)	3207.0340(491.26222)	24.5215(12.31924)	40008.7
Schlather	Cauchy		1.193e-01(1.988e-02)	2.816e+04(2.580e+03)	1*	40006.43
G.G.	whitmat	3.73760(0.79688)	0.09523(0.04302)	3000*	99.93629(2.23952)	40188.45
G.G.	Exp.	0.06092(0.001168)	0.63581(0.005975)	20000	1.91587(0.095508)	53872.94
G.G.	Bessel	8.44084(2.08741)	0.04278(0.02186)	4000*	139.34386(0.81877)	40187.74
Brown-Resnik				3.190e+04(5.492e+03))	6.259e-01(5.248e-02)	40185.75

**Tabla B.13:** Modelos ajustados para el año 2010

Modelos	correlación	$\sigma_{11}(m)$	$\sigma_{12}(m)$	$\sigma_{22}(m)$		TIC
Smith	Anis.	47567963(2173282)	1473705(783108)	28782212(506417)		71988.56
Modelos	correlación	sigma	nugget	rango	suavizamiento	TIC
Schlather	Whitmat		6.171e-02(1.357e-02)	1.750e+04(3.002e+03)	1*	52919.84
Schlather	Exp.		1.068e-01(1.433e-02)	2.599e+04(3.324e+03)	2*	52906.58
Schlather	Bessel		0.1127(0.0152)	7856.7264(944.3623)	2*	52904.64
Schlather	Cauchy		9.755e-02(1.337e-02)	3.392e+04(4.271e+03)	2*	52911
G.G.	whitmat	5.71293(0.98834)	0.04627(0.01815)	3000	97.60453(0.95487)	52852.13
G.G.	Exp.	0.06092(0.001168)	0.63581(0.005975)	20000*	1.91587(0.095508)	53872.94
G.G.	Bessel	5.61570(0.92882)	0.04708(0.01734)	4000*	52.53006(1.04331)	52852.71
Brown-Resnik				2.431e+04(3154.2020)	8.146e-01(0.0473)	52886.27

# Apéndice C

## Resultados de prueba Kolmogorov Smirnov

Pruebas de bondad de ajuste de Kolmogorov Smirnov para la GEV en cada posición

**Tabla C.1:** Estadísticos de Prueba y P-value. Estaciones: LAG, TAC, SAG, AZC.

año	est. D	P-value	est. D	P-value	est. D	P-value	est. D	P-value
1998	0.0481	0.9462	0.0543	0.8736	0.0546	0.8698	0.0688	0.6268
1990	0.0615	0.7899	0.0944	0.2706	0.0753	0.5497	0.0496	0.09453
2000	0.0601	0.8859	0.0536	0.9501)	0.0497	0.9745	0.0541	0.9462
2001	0.0504	0.9546	0.0499	0.9579	0.095	0.9969	0.0382	0.9981
2002	0.055	0.9652	0.0741	0.7588	0.0791	0.6843	0.0769	0.7168
2003	0.0623	0.7917	0.0483	0.9613	0.0452	0.9793	0.1018	0.2084
2004	0.0644	0.7078	0.0497	0.9303	0.0781	0.4626	0.06157	0.684
2005	0.0516	0.9189	0.0575	0.8415	0.066	0.06982	0.671	0.6784
2006	0.0661	0.8217	0.0455	0.9918	0.0708)	0.7521	0.0627	0.8671
2007	0.04	0.9984	0.0548	0.988	0.071	0.7362	0.572	0.9216
2008	0.0826	0.7504	0.05	0.9741	0.1005	0.5078	0.0729	0.869
2009	0.089	0.7668	0.0948	0.6955	0.0848	0.8159	0.0788	0.8772
2010	0.0817	0.7063	0.0737	0.8165	0.0919	0.5598	0.0581	0.9639

## C. Resultados de prueba Kolmogorov Smirnov

**Tabla C.2:** Estadísticos de Prueba y P-value. Estaciones: TLA, XAL, MER, PED.

año	est. D	P-value	est. D	P-value	est. D	P-value	est. D	P-value
1998	0.06	0.7851	0.0451	0.9686	0.0412	0.9877	0.0476	0.9503
1990	0.0489	0.952	0.071	0.6254	0.0787	0.4921	0.116	0.09819
2000	0.0532	0.9531	0.0372	0.9995)	0.0587	0.902	0.0663	0.8033
2001	0.0542	0.92	0.0418	0.9934	0.0509	0.9508	0.062	0.8186
2002	0.0868	0.5677	0.0704	0.8105	0.0646	0.8838	0.0495	0.9879
2003	0.0951	0.2775	0.0534	0.915	0.0561	0.8822	0.0543	0.9046
2004	0.0595	0.7928	0.0798	0.4354	0.0573	0.8302	0.0559	0.8506
2005	0.0564	0.8577	0.045	0.974	0.042	0.9873	0.0891	0.3212
2006	0.0704	0.7586	0.0638	0.8521	0.0695	0.7718	0.0682	0.7906
2007	0.0576	0.9171	0.0506	0.9711	0.0519	0.9635	0.0598	0.8934
2008	0.0665	0.9281	0.0567	0.9823	0.082	0.7589	0.0795	0.7912
2009	0.072	0.9333	0.0751	0.9101	0.077	0.8939	0.1018	0.6076
2010	0.0581	0.9642	0.0777	0.7633	0.0528	0.986	0.0925	0.5518

**Tabla C.3:** Estadísticos de Prueba y P-value. Estaciones: CES, PLA, UIZ, TAX.

año	est. D	P-value	est. D	P-value	est. D	P-value	est. D	P-value
1998	0.0551	0.8624	0.0481	0.9459	0.0543	0.8748	0.0746	0.5225
1990	0.0949	0.2654	0.1043	0.1744	0.0532	0.9095	0.0944	0.2712
2000	0.0634	0.844	0.0758	0.6533)	0.0583	0.9063	0.0556	0.9336
2001	0.0464	0.9785	0.499	0.9582	0.0557	0.9032	0.07	0.6872
2002	0.0728	0.7781	0.0612	0.9188	0.0756	0.7363	0.1008	0.3756
2003	0.0798	0.4911	0.054	0.9087	0.0865	0.3883	0.0728	0.611
2004	0.0472	0.9537	0.079	0.4473	0.0746	0.5225	0.0708	0.5906
2005	0.0497	0.939	0.0689	0.6454	0.0569	0.8512	0.064	0.7334
2006	0.063	0.8634	0.0599	0.9001	0.0441	0.9945	0.0561	0.9368
2007	0.0581	0.9118	0.0644	0.8346	0.0413	0.9973	0.0548	0.9426
2008	0.0684	0.9123	0.0707	0.8914	0.0658	0.9335	0.0845	0.7249
2009	0.0799	0.8668	0.065	0.9721	0.0913	0.07392	0.0787	0.8789
2010	0.0722	0.8351	0.0825	0.6957	0.0864	0.6385	0.0505	0.9917

## C. Resultados de prueba Kolmogorov Smirnov

---

**Tabla C.4:** Estadísticos de Prueba y P-value. Estaciones: CUA, TAH

año	est. D	P-value	est. D	P-value
1998	0.0353	0.9984	0.0397	0.0992
1990	0.078	0.5027	0.0549	0.808
2000	0.0484	0.9803	0.0662	0.8041
2001	0.036	0.9993	0.0546	0.9163
2002	0.0531	0.9747	0.0592	0.9358
2003	0.0632	0.7771	0.0894	0.3478
2004	0.0425	0.9825	0.0606	0.7744
2005	0.0625	0.7606	0.0438	0.9802
2006	0.066	0.8223	0.0587	0.9121
2007	0.067	0.7978	0.0648	0.8302
2008	0.1062	0.4363	0.1269	0.2307
2009	0.0627	0.9805	0.0976	0.6598
2010	0.08	0.7311	0.0648	0.9148