

PROGRAMA DE DOCTORADO EN
ESTADÍSTICA MATEMÁTICA Y APLICADA



**UNIVERSIDAD
DE GRANADA**

Tesis Doctoral

**Estructuras estocásticas notables en procesos puntuales
espacio-temporales y medidas de riesgo bivariantes**

Amalia Isabel Escudero Villa

Directores de tesis

Prof. Jorge Mateu Mahiques

Prof. José Miguel Angulo Ibáñez

Enero, 2023

Editor: Universidad de Granada. Tesis Doctorales
Autor: Amalia Isabel Escudero Villa
ISBN: 978-84-1117-686-6
URI: <https://hdl.handle.net/10481/80004>

Confía en Jehová con todo tu corazón y no te apoyes en tu propio entendimiento. Tómallo en cuenta en todos tus caminos, y él hará rectas tus sendas.

Proverbios 3:5,6

La motivación es lo más importante. Por ello es necesario procurar que nuestra motivación nazca del amor al prójimo.

Isabel

Dedicado a:

Mi familia:

Jehová Dios, mi madre, mi hija, mi esposo y mis hermanos.

Agradecimientos

Mi sincero agradecimiento a:

La Escuela Superior Politécnica de Chimborazo, por la Beca (Resolución No: 376.CP.2019) y Licencia (Resolución Administrativa No. 129-ESPOCH.2019), que hizo posible fortalecer mi formación académica.

La Universidad de Granada y la Universitat Jaume I, por las facilidades de acceso a las áreas y laboratorios de investigación.

Mis directores Jorge Mateu y José Miguel Angulo, que aparte de ser excelentes profesionales con amplia experticia en modelos de procesos puntuales y medidas de riesgos, son personas con múltiples cualidades humanas, de las cuales me es admirable su personalidad y carisma; cualidades ejemplares para las nuevas generaciones de investigadores.

Los directivos de las entidades gubernamentales ecuatorianas que facilitaron los datos e información complementaria de delitos en los periodos estudiados. Especialmente al Ing. Luis Cabezas analista del ECU911-Ecuador, por su predisposición y colaboración. Y al GEAA-ESPOCH por el libre acceso a los datos meteorológicos.

Ing. David Ramos e Ing. Cristian Escudero, quienes formaron un gran equipo de trabajo, y aportaron con las especificaciones técnicas de los servidores necesarios que facilitan la corrida de los modelos espacio-temporales, e incorporaron dispositivos de acceso remoto que facilitaron el control (encendido) y acceso permanente a los servidores desde el teléfono celular.

Mis amigos doctorandos: Somnath Chaudhuri, David Payares, Adriana Medialdea y Patricia Paredes, quienes me apoyaron con su amistad y consejos sobre herramientas de R. Así como también al Profesor Miguel Burgos de la UGR, por su apertura y continuo asesoramiento en herramientas de \LaTeX .

Y a todos quienes me apoyaron con su amistad, compañía, consejos, ánimo y buenos deseos para completar esta meta.

Resumen

La modelización estadística constituye un conjunto de formulaciones matemáticas que proporcionan un mejor entendimiento de los fenómenos, especialmente en los que subyacen relaciones entre variables que a su vez poseen influencia aleatoria. En este trabajo de tesis se realiza la modelización estadística del fenómeno delictivo, mediante el desarrollo de estructuras estocásticas en procesos puntuales espacio-temporales, con el propósito de reproducir eventos de la forma más exacta posible a la realidad en función de su evolución en el espacio, tiempo y espacio-tiempo.

En primer lugar, se estructura la dependencia espacio-temporal de los procesos auto-excitados espacialmente correlacionados, considerando una ecuación en diferencias estocástica para la intensidad del proceso espacio-temporal, que captura tanto la dependencia debida a la auto-excitación como la dependencia en un proceso espacial subyacente. Se sigue el razonamiento de [Clark y Dixon \(2021\)](#) y [Reinhart \(2018\)](#), pero capturando la no linealidad de las covariables espacio-temporales mediante una estructura aditiva generalizada (GAM) con B-splines. En segundo lugar, se estructura un modelo de proceso espacio-temporal latente, es decir, un procesos de Cox log-gaussiano (LGCP), con una función de intensidad separable de primer orden. La componente determinista temporal es el resultado de un modelo lineal generalizado (GLM) con efectos de covariables meteorológicas, días y meses del año, parámetros de regresión armónica, periodicidad anual de las tasas de incidencia y tendencia global. Para la componente determinista espacial se considera un GAM siguiendo el razonamiento de [Diggle et al. \(2005\)](#), [Taylor et al. \(2013\)](#) y [González y Mateu \(2021\)](#), pero con estimación de suavidad integrada con B-splines univariantes. Por su parte, para la componente estocástica, se considera un campo gaussiano que modeliza la dependencia y la variación espacio-temporal de los eventos. En tercer lugar, se identifican los mecanismos de auto-excitación entre las series de delitos en un tiempo continuo mediante los procesos puntuales de Hawkes. Se estima la tasa de background de cada componente con una reconstrucción estocástica no paramétrica; ésta incluye una periodicidad temporal, una componente espacial separable y una tendencia a largo plazo. Para la estimación semi-paramétrica de los coeficientes de relajación se utiliza máxima verosimilitud para estabilizar y asegurar el proceso de estimación. Se sigue el algoritmo iterativo de [Zhuang y Mateu \(2019\)](#) y [Zhuang \(2006\)](#). Los modelos de procesos puntuales espacio-temporales son buenas herramientas matemáticas para analizar datos, por lo que los modelos descritos proporcionan predicciones confiables, como se muestra a lo largo de esta memoria con datos de delitos.

Con la intención de ir más allá de la modelización y las predicciones de patrones puntuales, se formula un procedimiento para identificar las excedencias espaciales en LGCP bivariantes y sus regiones asociadas, así como también la cuantificación de la peligrosidad en términos de probabilidades.

Para esto, se modeliza la descomposición espacial de los hurtos (delito sin uso de violencia) y robos a personas (delitos con uso de violencia o amedrentamiento) para representarlos en una variación asociada a un tipo de delito en particular, lo que es posible mediante una estructura estocástica compuesta por una componente dentro del flujo y otra entre flujos. Y se incorpora un enfoque de medidas de riesgo como el Valor en Riesgo y el Déficit Esperado ([Malevergne y Sornette, 2006](#)). Estas medidas están basadas en percentiles de la distribución de los delitos que exceden un umbral alto. Se considera una estructura de dependencia mediante la cópula extrema de Gumbel-Hougaard ([Salvadori et al., 2007](#)) y distribuciones marginales de Pareto generalizadas ([Castellanos y Cabras, 2007](#); [Abad et al., 2014](#)).

En general, se adopta un marco bayesiano con MCMC-MALA y máxima verosimilitud para la inferencia sobre los parámetros de los diferentes modelos. Se hace estudio de casos con datos reales de delitos registrados en la ciudad de Riobamba-Ecuador y algunas covariables temporales, espaciales y espacio-temporales. Los resultados obtenidos proporcionan información relevante sobre el modelizado, predicción y comportamiento extremo localizado de eventos delictivos; sin embargo, pueden ser de mucha utilidad para aplicarse a conjuntos de datos de diversas áreas. Finalmente, se exponen las conclusiones y se plantean algunas ideas abiertas.

Summary

Statistical modeling is a set of mathematical formulations which provide a better understanding of phenomena, especially those that underlie relationships between variables that in turn have random influence. In this thesis work, the statistical modeling of criminal phenomena is carried out through the development of stochastic structures in spatio-temporal point processes, in order to reproduce events as accurately as possible to reality in terms of their evolution in space, time and space-time.

First, the spatio-temporal dependence of spatially correlated self-excited processes is structured, considering a stochastic difference equation for the intensity of the spatio-temporal process, which captures both the dependence due to self-excitation and the dependence on an underlying spatial process. The reasoning of [Clark and Dixon \(2021\)](#) and [Reinhart \(2018\)](#) is followed, but capturing the nonlinearity of the spatio-temporal covariates by means of a generalized additive structure (GAM) with B-splines. Second, a latent spatio-temporal process model, i.e. a LGCP, is structured with a separable first order intensity function. The deterministic time component is the result of a generalized linear model (GLM) with meteorological covariate effects, days and months of the year, harmonic regression parameters, annual periodicity of incidence rates and global trend. For the spatial deterministic component a generalized additive model (GAM) is considered, following the reasoning, of [Diggle et al. \(2005\)](#), [Taylor et al. \(2013\)](#) and [González and Mateu \(2021\)](#), but with integrated smoothness estimation with univariate B-splines. And for the stochastic component, we consider a Gaussian field that models the spatio-temporal dependence and variation of the events. Thirdly, self-excitation mechanisms are identified among the crime serie in a continuous time by means of Hawkes point processes. The background rate of each component is estimated with a nonparametric stochastic reconstruction; it includes a temporal periodicity, a separable spatial component and a long-term trend. The semi-parametric maximum likelihood estimation of the relaxation coefficients to stabilize and secure the estimation process, and the iterative algorithm of [Zhuang and Mateu \(2019\)](#) and [Zhuang \(2006\)](#) are followed. Spatio-temporal point process models are good mathematical tools for data analysis, and therefore, the described models provide reliable predictions, as shown throughout this thesis in crime data.

In order to go beyond modeling and predictions of point patterns, a procedure is formulated to identify the spatial exceedances in bivariate LGCP and their associated regions, as well as the quantification of the hazard in terms of probabilities. For this, we model the spatial decomposition of thefts (crime without the use of violence) and robberies (crime with the use of violence or intimidation) to represent them in a variation associated with a particular type of crime, which is possible by means of a stochastic structure composed of a within-flow component and another one between-flows. For

this modeling, we use log-Gaussian Cox processes along the same line of [Taylor et al. \(2015\)](#). And a focus on risk measures such as Value-at-Risk and Expeted Shorfall ([Malevergne and Sornette, 2006](#)) is incorporated. These measures are based on percentiles of the distribution of offenses exceeding a high threshold. A dependency structure using the extreme Gumbel-Hougaard copula is considered ([Salvadori et al., 2007](#)), and generalized Pareto marginal distributions ([Castellanos and Cabras, 2007](#); [Abad et al., 2014](#)).

In general, a Bayesian framework with MCMC-MALA and maximum likelihood is adopted for the inference on the parameters of the different models. A case study is developed with real data of crimes registered in the city of Riobamba-Ecuador and some temporal, spatial and spatio-temporal covariates. The results obtained provide relevant information on the modeling, prediction and extreme localized behavior of criminal events; however, they can be very useful to be applied to data sets of different areas. Finally, conclusions are drawn and some open ideas are put forward.

Índice general

Introducción	4
1 Procesos estocásticos de Poisson y medidas de riesgo	5
1.1 Introducción	5
1.2 Proceso de Poisson	6
1.3 Procesos de Poisson no homogéneos	7
1.4 Procesos espaciales	7
1.4.1 Clasificación	8
1.4.2 Procesos puntuales	8
1.4.3 Procesos de Cox	10
1.4.4 Procesos de Cox log-gaussianos (LGCP)	13
1.5 Procesos espacio-temporales	15
1.5.1 Tipo de datos	15
1.5.2 Procesos puntuales espacio-temporales	15
1.5.3 Características	16
1.5.4 Función de intensidad espacio-temporal	17
1.5.5 Clasificación	22
1.6 Medidas de riesgo	22
1.6.1 Teoría de valores extremos	23
1.6.2 Valores extremos multivariantes	28
1.6.3 Cópulas	29
1.6.4 Medidas de riesgo	33
2 Modelos espacio-temporales auto-regresivos para datos de conteo (SPINGARCH)	35
2.1 Introducción	35
2.2 Descripción de los datos	37
2.3 Metodología	42
2.3.1 Modelo	42
2.3.2 Inferencia bayesiana de parámetros	44
2.3.3 Estimación de la función de factores exógenos	45
2.4 Resultados	46
2.5 Discusión y conclusiones	52

3	Modelos espacio-temporales de Cox log-gaussianos (LGCP)	55
3.1	Introducción	55
3.2	Descripción de los datos	57
3.3	Metodología	62
3.3.1	Componente determinista temporal	62
3.3.2	Componente determinista espacial	63
3.3.3	Componente estocástica espacio-temporal	63
3.3.4	Modelos de predicción	65
3.3.5	Modelo para la identificación de excedencias	65
3.4	Resultados	66
3.5	Discusión y conclusiones	73
4	Modelos espacio-temporales semiparamétricos auto-excitados de Hawkes	75
4.1	Introducción	75
4.2	Descripción de los datos	77
4.3	Metodología	78
4.3.1	Componente de background	80
4.3.2	Componente de triggering	81
4.3.3	Estimación de coeficientes	81
4.3.4	Suavizado y corrección del efecto de borde	82
4.3.5	Bondad de Ajuste	83
4.4	Resultados	84
4.5	Discusión y conclusiones	90
5	Medidas de riesgo en LGCP bivalentes	93
5.1	Introducción	93
5.2	Descripción de los datos	95
5.3	Metodología	97
5.3.1	Modelo LGCP bivalente	97
5.3.2	Estimación de parámetros para el LGCP bivalente	98
5.3.3	Estimación de las funciones de distribución acumulativas marginales	100
5.3.4	Estimación del parámetro de dependencia	101
5.4	Resultados	103
5.5	Discusión y conclusiones	121
6	Conclusiones e ideas abiertas	123
6.1	Conclusiones	123
6.2	Ideas abiertas	124
	Bibliografía	127

Índice de figuras

Figura 1.1	Realización de un proceso de Poisson homogéneo	9
Figura 1.2	Generación de un proceso de Cox	12
Figura 1.3	Fase de generación de un proceso de Cox log-gaussiano	13
Figura 1.4	Enfoques para identificar valores extremos en datos reales	23
Figura 1.5	Vida residual media	28
Figura 2.1	Ciudad de Riobamba-Ecuador	38
Figura 2.2	Delitos registrados por tres entidades gubernamentales en la ciudad de Riobamba	39
Figura 2.4	Densidad poblacional (azul) y ubicaciones de los puntos de referencia	40
Figura 2.3	Estructura jerárquica para el registro de delitos en la ciudad de Riobamba	40
Figura 2.5	Promedios mensuales de temperatura y precipitación en la ciudad de Riobamba	41
Figura 2.6	Descripción media de eventos por meses y días, para los delitos registrados por MI (2010-2014), CJCH (2015-2019) y UPC (2015-2017)	41
Figura 2.7	Distribución de las covariables	42

Figura 2.8	B-splines cúbicos de las covariables exógenas para los delitos de los tres grupos de datos	47
Figura 2.9	Inferencia bayesiana para los registros delictivos de MI (2010–2014), donde $\eta = \eta$, $\kappa = \kappa$, $\sigma = 1/\tau$ y $\sigma_\epsilon = \sigma_\epsilon$	48
Figura 2.10	Inferencia bayesiana para los registros delictivos de CJCH (2015–2019), donde $\eta = \eta$, $\kappa = \kappa$, $\sigma = 1/\tau$, $\sigma_\epsilon = \sigma_\epsilon$	49
Figura 2.11	Inferencia bayesiana para los registros delictivos de UPC (2015–2017), donde $\eta = \eta$, $\kappa = \kappa$, $\sigma = 1/\tau$, $\sigma_\epsilon = \sigma_\epsilon$	49
Figura 2.12	Diferencias espacio-temporales entre delitos reales y predicciones	50
Figura 2.13	Delitos reales y predicciones	51
Figura 2.14	Predicciones espaciales de delitos para los tres grupos de datos	51
Figura 3.1	Ubicación geográfica de la ciudad de Riobamba-Ecuador	58
Figura 3.2	Delitos registrados durante el período 2010-2014 en la ciudad de Riobamba	58
Figura 3.3	Eventos delictivos mensuales en la ciudad de Riobamba	59
Figura 3.4	Covariables espaciales de la ciudad de Riobamba	60
Figura 3.5	Promedio semanal de las covariables temporales	61
Figura 3.6	GLM para la componente determinista temporal $\mu(t)$	67
Figura 3.7	GAM para la componente determinista espacial $\lambda(\mathbf{s})$	68
Figura 3.8	Estimación de los parámetros de escala espacial y temporal mediante la función de correlación por pares	69
Figura 3.9	Series temporales de los delitos reales y predicciones	70
Figura 3.10	Delitos originales y predicciones para las semanas [256, 261]	71
Figura 3.11	Diferencias temporales entre los delitos originales y predicciones con LGCP	72
Figura 3.12	Probabilidad de excedencias a posteriori	72
Figura 4.1	Distribución de los eventos delictivos en la ciudad de Riobamba-Ecuador y su función acumulativa	78
Figura 4.2	Historial de eventos delictivos en la ciudad de Riobamba-Ecuador	78
Figura 4.3	Periodicidad semanal y tendencia a largo plazo	84
Figura 4.4	Funciones de respuesta temporal y espacial	85
Figura 4.5	Tasa de ocurrencia en el dominio del tiempo transformado	86
Figura 4.6	Tasa de background $\mu_s(x, y)$	87

Figura 4.7	Periodicidad semanal y tendencia a largo plazo, clasificadas según los puntos cardinales y los años	88
Figura 4.8	Eventos delictivos reales (negro) y predicciones (verde) para las seis semanas, sobre el background $\mu_s(x, y)$	89
Figura 5.1	Hurtos y robos a personas registrados durante 2014 en la ciudad de Riobamba	96
Figura 5.2	Malla computacional	96
Figura 5.3	Covariables espaciales ajustadas a la malla computacional determinada	97
Figura 5.4	Diagnóstico de la convergencia de la distribución a posteriori de los parámetros (log-target)	104
Figura 5.5	A priori y a posteriori de cada parámetro	104
Figura 5.6	Covarianzas a posteriori para Y_1, Y_2 e Y_3	105
Figura 5.7	Intensidad estimada de hurtos (Λ_1), robos a personas (Λ_2) y conjunta (Λ_3)	106
Figura 5.8	Probabilidades condicionadas	107
Figura 5.9	Discretización de delitos mediante una ventana deslizante	107
Figura 5.10	Vida residual media de exceso de delitos	108
Figura 5.11	Parámetro de forma ξ de la GPD para $c = 15, c = 30$ y $c = 45$, respectivamente	109
Figura 5.12	Histogramas de los parámetros de escala $\sigma(\mathbf{s})$ y de forma $\xi(\mathbf{s})$ de la GPD, para tres de las localizaciones \mathbf{s}	110
Figura 5.13	Densidades predictivas de las distribuciones a posteriori de los parámetros de la GPD para algunas de las localizaciones \mathbf{s}	110
Figura 5.14	Niveles de retorno de la GPD de hurtos y robos a personas para tres de las localizaciones \mathbf{s}	111
Figura 5.15	Resumen de los valores de los parámetros de forma $\xi(\mathbf{s})$ resultantes	112
Figura 5.16	Curvas de nivel de los valores del parámetro de dependencia (θ) de las cópulas de Gumbel-Hougaard	112
Figura 5.17	Valores del parámetro de dependencia (θ) de la cópula de Gumbel-Hougaard en la región de estudio	113
Figura 5.18	Cópulas extremas de Gumbel-Hougaard para tres de las localizaciones \mathbf{s}	114
Figura 5.19	Distribución bivalente (\mathcal{C}_s^{GH}) de delitos sobre los cuantiles 50%, 75% y 90%, respectivamente	114
Figura 5.20	Distribuciones de u y v normalizadas	115
Figura 5.21	Simulación de hurtos y robos a personas para algunas de las localizaciones \mathbf{s} mediante \mathcal{C}_s^{CH} y sus contornos	116
Figura 5.22	Error cuadrático medio (RMSE) para hurtos y robos a personas, respectivamente	117
Figura 5.23	$VaR_{0.95}(\mathbf{s})$ bivalente	117

Figura 5.24	$VaR_{0.95}(\mathbf{s})$ univariante	118
Figura 5.25	Resumen de los $VaR_{0.95}(\mathbf{s})$ univariante y bivalente . .	118
Figura 5.26	Ratio de excepciones para $VaR_{0.95}(\mathbf{s})$ univariante y bivalente	119
Figura 5.27	Resumen del ratio de excepciones del $VaR_{0.95}(\mathbf{s})$ univa- riante y bivalente	119
Figura 5.28	$ES_{0.95}(\mathbf{s})$ univariante y bivalente	120
Figura 5.29	Resumen estadístico de los valores del $ES_{0.95}(\mathbf{s})$, univa- riante y bivalente	121

Índice de tablas

Tabla 2.1	Distribuciones a posteriori de los parámetros para los registros delictivos de MI (2010-2014)	48
Tabla 2.2	Distribución de parámetros para los delitos de CJCH (2015-2019)	49
Tabla 2.3	Distribución a posteriori de los parámetros para delitos de UPC (2015-2017)	50
Tabla 2.4	Errores cuadráticos medios temporales (MSPE)	51
Tabla 3.1	Coefficientes estimados mediante GLM para la componente temporal	66
Tabla 3.2	Coefficientes estimados mediante GAM con B-splines para la componente espacial	68
Tabla 5.1	Valores de los parámetros estimados	105

Siglas

acf: autocorrelation function

AIC: Akaike Information Criterion

ARMA: Auto-Regressive Moving Average

ATM: Automatic Teller Machine

BIC: Bayesian Information Criterion

BM: Block Maxima

BYM: Besag York Mollié

CAR: Conditional Auto-Regressive

CJCH: Consejo de la Judicatura de Chimborazo

CNN: Convolutional Neural Network

CSR: Completely Spatially Random

DP: Distribuciones dobles de Poisson

ECU911: Servicio Integrado de Seguridad 911

ES: Expected Shortfall

ESPOCH: Escuela Superior Politécnica de Chimborazo

EV: Extreme Value

EVC: Extreme Value Copula

EVT: Extreme Value Theory

GAM: Generalized Additive Models

GARCH: Generalized Auto-Regressive Conditional Heteroskedastic

GCV: Generalized Cross Validation

GEAA: Grupo de Energías Alternativas y Ambiente

GEV: Generalized Extreme Value

GED: Generalized Error Distribution

GEVD: Generalized Extreme Value Distribution

GH: Gumbel-Hougaard

GHC: Gumbel-Hougaard Copula
GLM: Generalized Linear Models
GP: Generalized Poisson
GPD: Generalized Pareto Distribution
i.i.d.: independientes e idénticamente distribuidas
INGARCH: Integer Generalized Auto-Regressive Conditionally Heteroskedastic
INLA: Integrated Nested Laplace Approximation
INEC: Instituto Nacional de Estadísticas y Censos
LGCP: Log-Gaussian Cox Process
MALA: Metropolis-Adjusted Langevin Algorithm
MCMC: Markov Chain Monte Carlo
MEV: Multivariate Extreme Value
MEVT: Multivariate Extreme Value Theory
MI: Ministerio del Interior
MLE: Maximum Likelihood Estimator
MSE: Mean Squared Error
NB: Negative Binomial
NHP: Next Hit Predictor
POT: Peaks Over Threshold
RMSE: Root Mean Squared Error
SAV: Symmetric Absolute Value
SEHP: Self-Exciting Hawkes Process
SPINGARCH: Spatial Integer Generalized Auto-Regressive Conditionally Heteroskedastic
UPC: Unidades Policiales Comunitarias
VaR: Value-at-Risk

Introducción

Los métodos de análisis de datos para los procesos puntuales han recibido cierta atención en el campo de la Estadística durante las décadas de 1950 y 1960. Concretamente, [Cox \(1955\)](#) se centró en los eventos que ocurren al azar en el espacio o el tiempo, las componentes de la varianza y la correlación entre los eventos de diferentes tipos, e introdujo la noción de un proceso de Poisson doblemente estocástico, que en la actualidad se le conoce como proceso de Cox. Posteriormente, [Bartlett \(1963\)](#) introdujo métodos estadísticos para procesos puntuales basados en sus densidades espectrales de potencia o también llamadas funciones de densidad de covarianza correspondientes. [Lewis y Govier \(1964\)](#) demostraron que la densidad de autocovarianza para un proceso es proporcional a la densidad de probabilidad de convolución de la diferencia de dos variables aleatorias independientes, con un resultado estrechamente relacionado para la función de autocorrelación de los recuentos. La metodología sobre el estudio de las observaciones en forma de eventos puntuales que ocurren en un continuo, espacio o tiempo, fue descrita por [Cox y Lewis \(1966\)](#) junto con ejemplos de aplicaciones de procesos puntuales univariantes y multivariantes. En este contexto, Hawkes se propuso llevar el enfoque del análisis espectral de Bartlett a un nuevo tipo de proceso, un proceso puntual auto-excitado como un resultado particular, y con las mismas propiedades de segundo orden que un cierto proceso doblemente estocástico ([Hawkes, 1971b,a](#); [Hawkes y Oakes, 1974](#)).

Los modelos de procesos puntuales espacio-temporales surgen de la necesidad de caracterizar la evolución temporal del comportamiento espacial de cantidades aleatorias. Estos modelos enfatizan el avance de los procedimientos de inferencia sobre la función de intensidad y las características de segundo orden del proceso. En palabras sencillas, estos modelos estiman una función de intensidad que predice la tasa de eventos en cualquier localización espacial (x, y) y tiempo t . En las últimas décadas, se han realizado avances en la estimación, inferencia, simulación y herramientas de diagnóstico para los modelos de procesos puntuales en general. Estos avances han tomado mayor importancia en aplicaciones como medio ambiente ([Siino et al., 2018](#); [Nasirzadeh et al., 2021](#); [Lieshout et al., 2012](#); [Pratiwi et al., 2017](#); [Reinhart, 2018](#)), presencia de plantas o animales ([Møller y Díaz-Avalos, 2010](#); [Myllymäki y Penttinen, 2009](#); [Illian et al., 2012](#); [Serra et al., 2014](#); [Sørbye et al., 2019](#); [Flagg y Hoegh, 2022](#)), salud ([Beneš et al., 2002](#); [Rostami et al., 2017](#); [Johnson et al., 2019](#); [Chiang et al., 2022](#); [Heinen, 2003](#)), finanzas ([Aït-Sahalia et al., 2015](#); [Hawkes, 2018](#); [Moreno Trujillo, 2019](#); [Heinen, 2003](#); [Fokianos y Tjøstheim, 2011](#)), análisis de medios sociales ([Ilhan y Kozat, 2020](#)), accidentes de tráfico ([Tang et al., 2022](#)) y otras múltiples áreas.

Al ser la delincuencia un fenómeno que afecta a todos los países, y que a través del tiempo trae consecuencias para la sociedad en general, este trabajo

de tesis concentra el desarrollo de estructuras estocásticas de procesos puntuales espacio-temporales como los procesos de Cox log-gaussianos (LGCP) y auto-excitados como el de Hawkes (SEHP) y el Spatial Integer Generalized Auto-Regressive Conditionally Heteroscedastic Correlation (SPINGARCH) para eventos delictivos, siguiendo los razonamientos principalmente de [Baddeley et al. \(2000, 2014\)](#), [Taylor et al. \(2013, 2015\)](#), [Clark y Dixon \(2021\)](#), [Zhuang y Mateu \(2019\)](#), además de formular un procedimiento que permite identificar las excedencias en LGCP bivariantes, sus regiones asociadas y una cuantificación de la peligrosidad en términos de probabilidad o riesgo, por lo que se incluye la estructura estocástica LGCP espacial bivalente.

Una forma de cuantificar la peligrosidad es mediante medidas de riesgo extremo, puesto que están definidas en términos de la distribución de pérdidas, que en el campo delictivo se refiere a la probabilidad de identificar anomalías en términos de un número de eventos delictivos igual o que exceda algún valor predeterminado. Las medidas de riesgo más comunes son Value at Risk (VaR) y Expected Shortfall (ES). El VaR es el α -ésimo percentil de la distribución y el ES es el tamaño de las anomalías esperadas, dado que se supera el VaR ([Malevergne y Sornette, 2006](#)). Una de las principales dificultades para estimar el VaR es modelizar la estructura de dependencia, especialmente debido a que el VaR está relacionado con la cola de la distribución, por lo que su punto de partida es la teoría de valores extremos ([Abad et al., 2014](#)).

La Extreme Value Theory (EVT) surge con [Fréchet \(1927\)](#), [Fisher y Tippett \(1928\)](#), [Mises \(1936\)](#), [Gnedenko \(1943\)](#) y [Gumbel \(1958\)](#), quienes demostraron que el mayor o menor valor de un conjunto de variables aleatorias distribuidas independientemente tiende a una distribución asintótica que sólo depende de la variable básica, y con el uso de constantes de normalización y centrado la distribución límite pertenece a un dominio de atracción de Gumbel, Fréchet o Weibull ([Gnedenko, 1943](#)). El método más frecuente para identificar valores extremos es el Peak Over Threshold (POT). Para el caso bivalente se incluye el uso de cópulas ([Malevergne y Sornette, 2006](#); [Salvadori et al., 2007](#)), puesto que permiten derivar distribuciones conjuntas independientemente de su distribución marginal, especialmente para la medición del riesgo.

En este contexto, la memoria de tesis se compone de seis capítulos, que se describen a continuación.

El capítulo 1 introduce en forma general los principios básicos y características de los procesos puntuales espaciales y espacio-temporales de Poisson, y tópicos necesarios para las medidas de riesgo. Los primeros apartados se concentran en los procesos de Poisson y el último expone definiciones sobre valores extremos, cópulas, cópulas extremas y medidas de riesgo como el VaR y ES.

El capítulo 2 muestra una extensión de los modelos mecanísticos auto-excitados de Hawkes, denominados por sus siglas en inglés SPINGARCH

(Spatial Integer Generalized Auto-Regressive Conditionally Heteroskedastic). Este es un estudio publicado con el título *A Spatially Correlated Model with Generalized Auto-regressive Conditionally Heteroskedastic Structure for Counts of Crimes* en *Entropy Journal*, que se puede encontrar en <https://www.mdpi.com/1099-4300/24/7/892> (Escudero et al., 2022). La introducción resume algunos avances y aplicaciones. La metodología describe la estructura del modelo, el enfoque para abordar adecuadamente la relación espacio-temporal y la estimación de parámetros mediante inferencia bayesiana con MCMC. Los resultados detallan las estrategias utilizadas para la incorporación de variables exógenas temporales, espaciales y espacio-temporales mediante Generalized Additive Model (GAM) con B-splines para capturar la no linealidad. El estudio de caso con datos reales se realiza con eventos delictivos registrados en tres instituciones gubernamentales de la ciudad de Riobamba-Ecuador.

El capítulo 3 muestra los modelos LGCP espacio-temporales. La introducción resume algunos avances y aplicaciones en diversas áreas. La metodología describe la estructura de las componentes deterministas temporales y espaciales, la componente estocástica que modeliza la dependencia y la variación espacio-temporal de los eventos, así como la estimación de los parámetros mediante un enfoque bayesiano con MCMC y MALA. Los resultados detallan las estrategias para la inclusión de efectos de variables meteorológicas para la componente temporal y B-splines para las covariables espaciales basadas en distancias. El estudio de caso con datos reales se realiza sobre delitos registrados en la ciudad de Riobamba-Ecuador durante el periodo 2010–2014.

El capítulo 4 muestra el modelo mecanístico auto-excitado semi-paramétrico de Hawkes. La introducción resume algunos avances y diversas áreas de aplicación. La metodología describe la técnica de reconstrucción estocástica para la estructura de las componentes de background (basal) y triggering (desencadenante), y la estimación de los parámetros por máxima verosimilitud. Los resultados detallan el historial de las observaciones, la componente de periodicidad y tendencia a largo plazo, funciones de respuesta espacial, temporal y bondad de ajuste. Se incluyen las predicciones obtenidas con LGCP sobre el background como un enfoque de modelización complementaria entre los dos modelos. El estudio de caso con datos reales se realiza con los eventos delictivos utilizados en el LGCP, para identificar el grado de influencia de eventos pasados sobre nuevos eventos.

El capítulo 5 muestra la formulación de un procedimiento para cuantificar riesgos extremos en LGCP espaciales bivariantes. La introducción resume algunos aspectos sobre la medición del riesgo considerando un enfoque de cópulas extremas. La metodología describe las estructuras de la modelización basada en LGCP espaciales bivariantes, distribuciones de Pareto Generalizadas, cópulas extremas de Gumbel-Hougaard, medidas de riesgo (VaR y ES) y la estimación de los parámetros mediante un enfoque bayesiano con MCMC

y máxima verosimilitud. Los resultados detallan las estrategias de simulación, discretización, uso de las cópulas y cuantificación de la peligrosidad delictiva mediante el cálculo del VaR y ES univariante y bivariante. El estudio de caso con datos reales se realiza con delitos de hurto y robo a personas registradas en la ciudad de Riobamba durante el año 2014.

El capítulo 6 resume las conclusiones y las aportaciones del presente trabajo, y señala alguna de sus posibles evoluciones como futuras líneas de investigación.

La memoria finaliza con una exhaustiva y actualizada bibliografía que ha sido revisada para el desarrollo de la investigación.

1. Procesos estocásticos de Poisson y medidas de riesgo

1.1. Introducción

En esta sección se hace una descripción rápida de los fundamentos teóricos necesarios, especialmente para quienes inician el estudio del impresionante campo de los modelos de procesos puntuales espacio-temporales. Se muestra una definición general de procesos estocásticos de Poisson, y se incluye la descripción de algunos procesos puntuales espaciales y espacio-temporales. Como complemento a los modelos de procesos puntuales, se incorporan los principios básicos sobre teoría de valores extremos, cópulas y medidas de riesgo.

Se toman como referencia principal los trabajos de [Davis et al. \(2021\)](#), [Zhuang y Mateu \(2019\)](#), [Clark y Dixon \(2018\)](#), [González et al. \(2016\)](#), [Baddeley et al. \(2015\)](#), [Taylor et al. \(2015\)](#), [Taylor et al. \(2013\)](#), [Cressie y Moores \(2021\)](#), [Salvadori et al. \(2007\)](#), [Malevergne y Sornette \(2006\)](#), [Coles \(2001\)](#), [Kriele y Wolf \(2014\)](#), entre otros autores.

Proceso estocástico

Un proceso estocástico es una colección de variables aleatorias, definidas en un dominio como el tiempo, el espacio, el espacio-tiempo u otro parámetro o argumento. Se suele utilizar:

- $Y(t)$ para denotar un proceso estocástico que varía durante un período de tiempo continuo $0 \leq t \leq T$. Cuando se considera que el tiempo es discreto, por ejemplo, la hora, el día, el mes, el año, etc., es frecuente usar la notación Y_t ; en este caso, Y_t se denomina serie temporal.
- $Y(\mathbf{s})$ para denotar un proceso estocástico espacial definido sobre un dominio $S \in R^d$, donde \mathbf{s} denota una ubicación espacial descrita por un número finito de dimensiones, como latitud, longitud, altitud, entre otras.

- $Y(\mathbf{s}, t)$ para denotar un proceso estocástico espacio-temporal, es decir, que varía en el espacio y en el tiempo.

1.2. Proceso de Poisson

El proceso de Poisson es uno de los procesos estocásticos de recuento más utilizados, especialmente en escenarios donde se cuentan las ocurrencias de ciertos eventos que parecen ocurrir en una cierta tasa completamente aleatoria. Por ejemplo, si a partir de datos históricos, se sabe que los terremotos se producen con un cierto ritmo en tiempos y lugares completamente aleatorios, el proceso de Poisson podría ser un buen modelo para este tipo de fenómenos. En la práctica, el proceso de Poisson o sus extensiones se han utilizado para modelizar fenómenos como el número de accidentes de tráfico, delitos en un lugar o zona, localización de usuarios en una red inalámbrica, solicitudes de documentos individuales en un servidor web, el estallido de guerras, entre otros.

Variable aleatoria de Poisson. Se dice que Y es una variable aleatoria de Poisson con parámetro λ , denotada como $Y \sim Poisson(\lambda)$, si su función de masa de probabilidad es de la forma

$$f_Y(k) = \begin{cases} \frac{\exp(-\lambda)\lambda^k}{k!} & \text{si } k \in \mathbb{N} \\ 0 & \text{en otro caso} \end{cases} \quad (1.1)$$

con $E[Y] = \lambda$ y $Var(Y) = \lambda$. Si $Y_i \sim Poisson(\lambda_i)$ e independientes para $i = 1, 2, \dots, n$, entonces $Y_1 + Y_2 + \dots + Y_n \sim Poisson(\lambda_1 + \lambda_2 + \dots + \lambda_n)$.

Definición. Sea $\lambda > 0$ fijo. El proceso de recuento $Z(t), t \in [0, \infty)$, se llama **proceso de Poisson** con intensidad o tasa λ si se cumple:

1. $Z(0) = 0$.
2. $Z(t)$ tiene incrementos independientes.
3. El número de llegadas en cualquier intervalo de longitud $t > 0$ tiene distribución $Poisson(\lambda t)$.

Detalles de este tipo de procesos pueden verse, por ejemplo, en [Pishro-Nik \(2016\)](#). Es importante notar que la tercera condición implica que la distribución del número de llegadas en cualquier intervalo depende sólo de la longitud del intervalo, y no de la ubicación exacta del intervalo en la recta real. Por tanto, el proceso de Poisson tiene incrementos estacionarios.

Si un proceso estocástico representa la llegada de individuos a una cola a lo largo del tiempo, es de esperar que posea incrementos independientes, pero no necesariamente estacionarios (podría haber momentos con mayor flujo de llegadas).

1.3. Procesos de Poisson no homogéneos

En un proceso de Poisson no homogéneo los tiempos entre dos llegadas consecutivas ya no siguen una distribución exponencial, ni son independientes, es decir, no se cumple la condición tres de la definición de proceso de Poisson. Por ejemplo, sea $Z(t)$ el número de clientes que llegan a un restaurante de comida rápida en el tiempo t . Inicialmente se cree que los clientes llegan de forma algo aleatoria, por lo que se podría querer modelizar $Z(t)$ como un proceso de Poisson. Sin embargo, se observa que la tasa de llegada de clientes es mayor durante la hora de la comida (incrementos no estacionarios). En estos casos, se podría modelizar $Z(t)$ como un proceso de Poisson no homogéneo. Este proceso tiene todas las propiedades de un proceso de Poisson, excepto el hecho de que su tasa es una función del tiempo, es decir, $\lambda = \lambda(t)$.

Definición. Sea $\lambda(t) : [0, \infty) \mapsto [0, \infty)$ una función integrable; el proceso de recuento $Z(t), t \in [0, \infty)$ toma el nombre de proceso de Poisson no homogéneo con tasa $\lambda(t)$ si se cumple que:

1. $Z(0) = 0$.
2. $Z(t)$ tiene incrementos independientes para cualquier $t \in [0, \infty)$.

Para un proceso de Poisson no homogéneo con tasa $\lambda(t)$, el número de llegadas en cualquier intervalo es una variable aleatoria de Poisson. Sin embargo, su parámetro puede depender de la ubicación del intervalo; más concretamente, se puede escribir como

$$Z(t+l) - Z(t) \sim \text{Poisson} \left(\int_t^{t+l} \lambda(\alpha) d\alpha \right). \quad (1.2)$$

1.4. Procesos espaciales

Según [Cressie y Moores \(2021\)](#), en estadística espacial, la incertidumbre en la teoría científica se expresa de forma probabilística a través de un proceso estocástico espacial $\{Y(\mathbf{s}) : \mathbf{s} \in \mathcal{S}\}$, donde $Y(\mathbf{s})$ es el valor del atributo aleatorio de la ubicación \mathbf{s} , y \mathcal{S} es un subconjunto de un espacio d -dimensional. El espacio euclídeo \mathbb{R}^d indexa todas las posibles ubicaciones espaciales de interés. Dentro de \mathcal{S} hay un conjunto S (posiblemente aleatorio) que indexa las partes de \mathcal{S} relevantes para el estudio. S puede tener un conjunto de propiedades diferentes, dependiendo de si el proceso espacial es un proceso geoestadístico, un proceso de malla (lattice) o un proceso puntual.

1.4.1. Clasificación

- Para un **proceso geoestadístico**, $S = S^G$, en donde S^G es un conjunto conocido sobre el que las localizaciones varían continuamente. El área o volumen es mayor que cero.
- Para un **proceso de malla**, $S = S^L$, donde S^L es un conjunto conocido cuyas localizaciones varían discretamente, finito o infinito numerable. El área o volumen de S^L es igual a cero.
- Para un **proceso puntual**, $S = S^P$, donde S^P es un conjunto aleatorio formado por puntos aleatorios en el espacio euclídeo R^d .

1.4.2. Procesos puntuales

Un proceso de puntos espacial es una colección numerable de localizaciones aleatorias $S \equiv S^P \subset \mathcal{S}$. El proceso de recuento, denotado como $\{Z(A) : A \subset \mathcal{S}\}$ está estrechamente relacionado con este conjunto aleatorio de puntos, en donde \mathcal{S} indexa todas las posibles localizaciones de interés, a más de asumir que está acotado. Por ejemplo, si A es un subconjunto dado de \mathcal{S} y dos de los puntos aleatorios $\{\mathbf{s}_i\}$ están contenidos en A , entonces $Z(A) = 2$. Ya que $S^P = \{\mathbf{s}_i\}$ es aleatorio y A es fijo, $Z(A)$ es una variable aleatoria definida en los enteros no negativos.

Claramente, las distribuciones conjuntas $[Z(A_1), \dots, Z(A_m)]$, para cualesquiera subconjuntos $\{A_j : j = 1, \dots, m\}$ contenidos en \mathcal{S} (posiblemente superpuestos) y para cualquier $m = 0, 1, 2, \dots$, están bien definidas. La dependencia espacial se puede ver a través de la proximidad espacial entre los $\{A_j\}$. Si se consideran sólo dos subconjuntos fijos A_1 y A_2 (es decir, $m = 2$) y, para evitar la ambigüedad causada por la posibilidad de compartir puntos, se impone que $A_1 \cap A_2$ sea vacío, entonces no se exhibe dependencia espacial si hay independencia estadística, eso es,

$$[Z(A_1), Z(A_2)] = [Z(A_1)][Z(A_2)]. \quad (1.3)$$

El proceso puntual básico conocido como proceso de Poisson tiene la propiedad de independencia, y su proceso de conteo asociado satisface

$$[Z(A)] = \exp\{-\lambda(A)\} \frac{\lambda(A)^{Z(A)}}{Z(A)!}; \quad A \subset \mathcal{S}, \quad (1.4)$$

donde $\lambda(A) \equiv \int_A \lambda(\mathbf{s}) d\mathbf{s}$. En la ecuación (1.4) $\lambda(\cdot)$ es una función de intensidad definida según

$$\lambda(\mathbf{s}) \equiv \lim_{|\delta\mathbf{s}| \rightarrow 0} \frac{E[Y(\delta\mathbf{s})]}{|\delta\mathbf{s}|}, \quad (1.5)$$

1.4. PROCESOS ESPACIALES

donde $\delta\mathbf{s}$ es un conjunto pequeño centrado en $\mathbf{s} \in \mathcal{S}$, y cuyo volumen es $|\delta\mathbf{s}|$.

La expresión (1.4) da como resultado un proceso de Poisson homogéneo en el caso de $\lambda(\mathbf{s}) \equiv \lambda$ para todo $\mathbf{s} \in \mathcal{S}$. Por ejemplo, la Figura 1.1 muestra una simulación probabilística equivalente para $Z(\mathcal{S})$, en $\mathcal{S} = [0, 1] \times [0, 1]$, con parámetros $\lambda = 50$ y $Z(\mathcal{S}) = 53$.

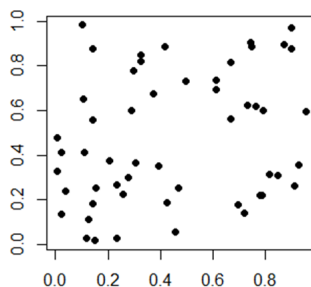


Figura 1.1: Realización de un proceso de Poisson homogéneo

La simulación condicionada a $Z(\mathcal{S})$, $\{\mathbf{s}_1, \dots, \mathbf{s}_{Z(\mathcal{S})}\}$ i.i.d. según la distribución uniforme

$$[\mathbf{u}] = \begin{cases} \frac{1}{\lambda(\mathcal{S})} & \text{if } \mathbf{u} \in \mathcal{S} \\ 0 & \text{en otro caso,} \end{cases} \quad (1.6)$$

explica la razón por la que un proceso de Poisson homogéneo se denomina proceso Completely Spatially Random (CSR) y se usa como línea de base para probar la ausencia de dependencia espacial en un proceso de puntos. Es decir, antes de que un modelo espacial se ajuste a un patrón de puntos, a menudo se lleva a cabo una prueba de la hipótesis nula de que el patrón de puntos se origina en un proceso de CSR. El rechazo de CSR justifica entonces el ajuste de procesos puntuales espacialmente dependientes a los datos (Diggle et al., 2013).

Procesos no homogéneos

Si la densidad media de puntos es una función $\lambda(\mathbf{s})$ de la ubicación espacial \mathbf{s} , dada una región del espacio B , dividida en pequeños píxeles, el número esperado de puntos que caen en el píxel es $\lambda(\mathbf{s})\Delta\mathbf{s}$, donde $\Delta\mathbf{s}$ es el área del píxel. El número total esperado de puntos en B es la suma de estos valores $\lambda(\mathbf{s})\Delta\mathbf{s}$, sobre todos los píxeles con centroides \mathbf{s} dentro de B . En el límite a medida que el tamaño del píxel se hace infinitesimal, esta suma se convierte en

$$\int_B \lambda(\mathbf{s})d\mathbf{s}. \quad (1.7)$$

Un proceso de puntos de Poisson es no homogéneo con función de intensidad $\lambda(\mathbf{s})$ si:

- El número esperado de puntos que caen en una región B es $\mu = \int_B \lambda(\mathbf{s}) d\mathbf{s}$, la integral de la función de intensidad $\lambda(\mathbf{s})$ sobre la región B .
- Si el espacio se divide en regiones no solapadas, los patrones aleatorios dentro de estas regiones son independientes entre sí.
- El número aleatorio de puntos que caen en una región determinada tiene una distribución de probabilidad de Poisson.

Detalles sobre procesos puntuales espaciales no homogéneos pueden consultarse en [Baddeley et al. \(2015\)](#).

1.4.3. Procesos de Cox

Procesos mixtos

El ejemplo más sencillo de un proceso de Cox es el proceso de Poisson mixto en el que se genera un número aleatorio Λ , con el cual se construye un proceso de Poisson uniforme con intensidad Λ . Dada una realización de Λ , el número esperado de puntos que caen en una región B es $E[n(Y(\mathbf{s}) \cap B) | \Lambda] = \Lambda |B|$. Al promediar los diferentes resultados posibles de la intensidad estocástica Λ , se obtiene

$$E[n(Y \cap B)] = E[E[n(Y \cap B) | \Lambda]] = E[\Lambda |B|] = E[\Lambda] |B|. \quad (1.8)$$

Por tanto, la intensidad del proceso de Poisson mixto es $\lambda = E[\Lambda]$, el valor medio de la intensidad de la parte aleatoria. También se puede encontrar la varianza del número de puntos que caen en una región B , utilizando el principio de análisis de la varianza:

$$Var[n(Y \cap B)] = E[Var[n(Y \cap B) | \Lambda]] + Var(E[n(Y \cap B) | \Lambda]), \quad (1.9)$$

donde, en el lado derecho, el primer término es la variabilidad debida al proceso de Poisson, y el segundo término es la variabilidad debida a Λ . Para un valor dado de Λ , el número de puntos tiene una distribución de Poisson, con varianza igual a la media, $Var[n(Y \cap B) | \Lambda] = E[n(Y \cap B) | \Lambda] = \lambda |B|$, por lo que se obtiene

$$Var[n(Y \cap B)] = (E\Lambda) |B| + (Var\Lambda) |B|^2. \quad (1.10)$$

Si la varianza de $n(Y \cap B)$ supera su media, es indicativo de que los recuentos de puntos están sobredispersos.

Modelo log-lineal

El modelo log-lineal general engloba a los modelos de intensidad siguientes:

- Homogénea: $\lambda(\mathbf{s}) \equiv \lambda$.
- Homogénea en diferentes regiones: $\lambda(\mathbf{s})$ toma el valor β_j cuando \mathbf{s} está en la región B_j , donde β_j es el parámetro a estimar.
- Proporcional a la línea de base: $\lambda(\mathbf{s}) = \theta b(\mathbf{s})$, donde $b(\mathbf{s})$ es una función conocida como línea de base y θ es un parámetro a estimar.
- Exponencial con covariable: $\lambda(\mathbf{s}) = \exp(\alpha + \beta X(\mathbf{s}))$, donde $X(\mathbf{s})$ es una covariable espacial, y α y β son parámetros a estimar.
- Incidencia elevada: $\lambda(\mathbf{s}) = b(\mathbf{s}) \exp(\alpha + \beta X(\mathbf{s}))$.

El modelo de intensidad log-lineal o modelo de proceso de Poisson modulado está dado por

$$\begin{aligned} \lambda_\theta(\mathbf{s}) &= \exp(B(\mathbf{s}) + \theta^T X(\mathbf{s})) \\ &= \exp(B(\mathbf{s}) + \theta_1 X_1(\mathbf{s}) + \dots + \theta_p X_p(\mathbf{s})), \end{aligned} \quad (1.11)$$

donde $B(\mathbf{s})$ y $X(\mathbf{s}) = (X_1(\mathbf{s}), \dots, X_p(\mathbf{s}))$ son funciones conocidas, y $\theta = (\theta_1, \dots, \theta_p)$ son parámetros a estimar ([Baddeley et al., 2015](#)).

El logaritmo de la intensidad es una suma de términos que representan diferentes efectos; incluso se puede agregar un término en el modelo que dé cuenta de la probabilidad de observar un punto. Existen supuestos de modelización plausibles que dan lugar a modelos de procesos de Poisson de forma aditiva o multiplicativa. Un modelo de la forma aditiva tiene una función de intensidad $\lambda(\mathbf{s}) = \gamma_1(\mathbf{s}) + \gamma_2(\mathbf{s}) + \dots$, donde los términos $\gamma_k(\mathbf{s})$ son interpretables como las intensidades de los procesos de Poisson asociados con diferentes orígenes o causas que dan lugar al proceso puntual observado. Un modelo multiplicativo tiene la forma

$$\lambda(\mathbf{s}) = \exp(A(\mathbf{s}) + B_1(\mathbf{s}) + B_2(\mathbf{s}) + \dots), \quad (1.12)$$

donde el primer término $A(\mathbf{s})$ es una intensidad de base en escala logarítmica y los siguientes términos $B_1(\mathbf{s}), B_2(\mathbf{s}), \dots$ pueden interpretarse como efectos de *adelgazamiento* o *modulación*. La forma multiplicativa es más conveniente para la inferencia, y corresponde directamente a la elección canónica del enlace logarítmico en un modelo lineal generalizado de Poisson ([Baddeley et al., 2015](#)).

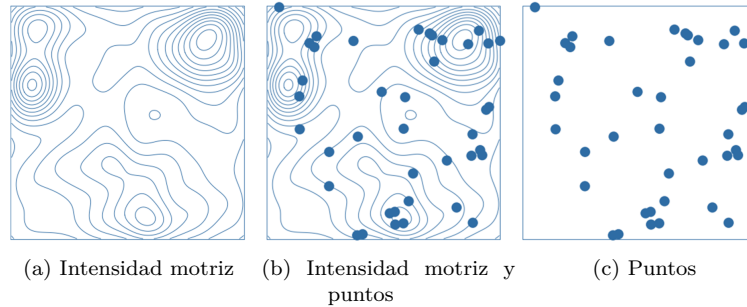


Figura 1.2: Generación de un proceso de Cox

Proceso de Cox

Según [Baddeley et al. \(2015\)](#), el modelo de Cox postula que existe una función de intensidad subyacente espacialmente variable, $\Lambda(\mathbf{s})$, que es aleatoria porque depende de factores externos no observables, así como de covariables observables. Si la superficie de intensidad $\Lambda(\mathbf{s})$ fuera conocida, los puntos constituirían un proceso de Poisson con función de intensidad $\Lambda(\mathbf{s})$. Los procesos de Cox, también conocidos como *procesos de Poisson doblemente estocásticos* o *procesos de Poisson modulados*, fueron introducidos por [Cox \(1955\)](#).

Para generar una realización del proceso de Cox, primero se genera una realización de la función aleatoria subyacente $\Lambda(\mathbf{s})$, también llamada *intensidad motriz* (véase la Figura 1.2a). La intensidad motriz está oculta, sólo se observan los puntos (véase la Figura 1.2b). Dada esta función, se genera una realización de un proceso de Poisson con intensidad $\Lambda(\mathbf{s})$. Para la generación de realizaciones de procesos de Cox, véase el código R en [Baddeley et al. \(2015\)](#).

La función aleatoria $\Lambda(\mathbf{s})$ puede ser generada por cualquier mecanismo aleatorio que se elija, siempre que produzca una función de intensidad válida. Los valores de $\Lambda(\mathbf{s})$ deben ser no negativos, y cada realización de $\Lambda(\mathbf{s})$ debe ser integrable sobre conjuntos acotados.

Según [Møller y Waagepetersen \(2007\)](#), es necesario destacar tres puntos de importancia estadística:

- Aunque se esté modelizando una heterogeneidad ambiental aleatoria, $Y(\mathbf{s})$ es estacionario si $\Lambda(\mathbf{s})$ es estacionaria.
- No se puede distinguir el proceso de Cox $\mathcal{Y}(\mathbf{s})$ de su correspondiente proceso de Poisson $Y(\mathbf{s})|\Lambda(\mathbf{s})$ cuando sólo se dispone de una realización de \mathcal{Y}_S .

1.4. PROCESOS ESPACIALES

- La probabilidad es en general desconocida, mientras que las densidades del producto pueden ser tratables (véase detalles en [Møller y Waagepetersen, 2007](#)).

Campos aleatorios

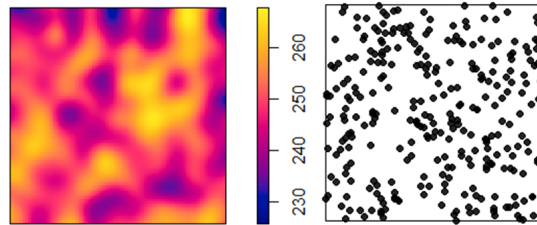
Dado un espacio de probabilidad (Ω, \mathcal{F}, P) , un campo aleatorio es una colección de variables aleatorias definidas sobre dicho espacio, con índice en algún espacio topológico o geométrico, por ejemplo, un espacio euclídeo d -dimensional:

$$Y = \{Y(\mathbf{s}) : \mathbf{s} \in \mathcal{S} \subseteq \mathbb{R}^d\}, \quad (1.13)$$

donde \mathcal{S} define un subconjunto discreto o continuo de \mathbb{R}^d .

1.4.4. Procesos de Cox log-gaussianos (LGCP)

Según [Baddeley et al. \(2015\)](#), el enfoque más natural es construir modelos para $\mathcal{Y}(\mathbf{s}) = \log(\Lambda(\mathbf{s}))$, es decir, el logaritmo de la intensidad estocástica. De este modo, la escala logarítmica también asegura que los valores de intensidad sean positivos. En muchas aplicaciones tiene sentido suponer que $\mathcal{Y}(\mathbf{s})$ tiene una distribución gaussiana para cada localización \mathbf{s} . Además, para cada par de localizaciones (\mathbf{s}, \mathbf{u}) , el par de valores $(\mathcal{Y}(\mathbf{s}), \mathcal{Y}(\mathbf{u}))$ deben tener una distribución gaussiana bivalente. En general, para cualquier conjunto elegido de ubicaciones $\mathbf{s}_1, \dots, \mathbf{s}_n$ los valores correspondientes $\mathcal{Y}(\mathbf{s}_1), \dots, \mathcal{Y}(\mathbf{s}_n)$ debe tener una distribución conjunta gaussiana multivalente. Una función aleatoria $\mathcal{Y}(\mathbf{s})$ con estas propiedades se denomina campo aleatorio gaussiano (véase un ejemplo en la Figura 1.3, en donde (a) representa la realización de un campo aleatorio log-gaussiano, y (b) el patrón generado por el correspondiente proceso de Poisson).



(a) Intensidad log-gaussiana (b) Proceso puntual

Figura 1.3: Fase de generación de un proceso de Cox log-gaussiano

Dado que las variables aleatorias gaussianas pueden tomar valores negativos, un verdadero campo aleatorio gaussiano no puede constituir un modelo para la intensidad. Sin embargo, el campo aleatorio log-gaussiano obtenido mediante exponenciación de un campo aleatorio gaussiano siempre es no negativo, pudiendo usarse como la intensidad de un proceso puntual. Un proceso de Cox log-gaussiano (LGCP) es un proceso de Cox cuya intensidad estocástica es de la forma

$$\Lambda(\mathbf{s}) = \exp(\mathcal{Y}(\mathbf{s})), \quad (1.14)$$

donde $\mathcal{Y}(\mathbf{s})$ es un campo aleatorio gaussiano. Una gran ventaja de $\mathcal{Y}(\mathbf{s})$ es que está completamente especificado por sus momentos de primer y segundo orden:

$$\mu(\mathbf{s}) = E[\mathcal{Y}(\mathbf{s})], \quad (1.15)$$

$$\mathbb{C}(\mathbf{s}, \mathbf{u}) = \text{Cov}(\mathcal{Y}(\mathbf{s}), \mathcal{Y}(\mathbf{u})) = E[\mathcal{Y}(\mathbf{s})\mathcal{Y}(\mathbf{u})] - \mu(\mathbf{s})\mu(\mathbf{u}). \quad (1.16)$$

Dos campos aleatorios gaussianos con las mismas funciones de media y covarianza son, pues, idénticos con respecto a su distribución de probabilidad. Los campos aleatorios gaussianos se pueden ajustar y reescalar de la misma manera que las variables aleatorias gaussianas. Si $\mathcal{Y}_0(\mathbf{s})$ es un campo aleatorio gaussiano con media 0 y función de covarianza $\mathbb{C}_0(\mathbf{s}, \mathbf{u})$, y si $a(\mathbf{s})$ y $b(\mathbf{s})$ son funciones no aleatorias, entonces $\mathcal{Y}(\mathbf{s}) = a(\mathbf{s})\mathcal{Y}_0(\mathbf{s}) + b(\mathbf{s})$ es un campo aleatorio gaussiano con función media $\mu(\mathbf{s}) = b(\mathbf{s})$ y función de covarianza $\mathbb{C}(\mathbf{s}, \mathbf{u}) = a(\mathbf{s})a(\mathbf{u})\mathbb{C}_0(\mathbf{s}, \mathbf{u})$. Esto hace posible construir una amplia gama de campos aleatorios gaussianos no homogéneos, y por lo tanto, procesos de Cox log-gaussianos no homogéneos a partir de una pequeña colección de modelos estacionarios.

En los modelos donde \mathbb{C} es isotrópica, es decir, sólo depende de la distancia $r = \|\mathbf{s} - \mathbf{u}\|$ entre las localizaciones \mathbf{s} y \mathbf{u} , la función de covarianza será de la forma

$$\mathbb{C}(\mathbf{s}, \mathbf{u}) = \mathbb{C}_0(r) = \sigma^2 \gamma\left(\frac{r}{\phi}\right), \quad (1.17)$$

donde $\phi > 0$ es un parámetro de escala, $\sigma^2 > 0$ el parámetro de varianza y $\gamma(r)$ una función de correlación bien definida. La función exponencial es una de las más utilizadas en este contexto, de forma que se tiene en este caso

$$\mathbb{C}_0(r) = \sigma^2 \exp\left(\frac{-r}{\phi}\right). \quad (1.18)$$

Se puede incorporar el efecto de covariables a través de la formulación del campo aleatorio de intensidad.

1.5. Procesos espacio-temporales

1.5.1. Tipo de datos

Según [Sahu \(2022\)](#), los datos se denominan espacio-temporales siempre que cada uno de ellos lleve asociada una ubicación y una marca de tiempo, incluyendo datos que se observan en un solo punto del tiempo o en una sola ubicación en el espacio. La agregación en el tiempo o el espacio reduce la variabilidad de los datos a analizar, y limita el alcance de las inferencias que se pueden hacer. Por ejemplo, no es posible detectar tendencias mensuales simplemente analizando datos agregados anuales.

Una de las primeras tareas en el análisis de datos espacio-temporales es elegir las resoluciones espaciales y temporales de interés para modelizar los datos. Por ello, es importante definir los objetivos inferenciales del estudio, además de decidir sobre las resoluciones espaciales y temporales más altas posibles a las que se deben hacer las inferencias; esto ayuda a determinar si se está sujeto a trabajar con datos diarios, mensuales o anuales. Una resolución demasiado fina, ya sea temporal y/o espacial, puede representar un gran desafío en el procesamiento, modelización y análisis de datos; por lo tanto, es importante mantener un equilibrio.

1.5.2. Procesos puntuales espacio-temporales

El término proceso de puntos o patrón de puntos espacio-temporales hace referencia a que la información principal es sobre una colección de datos que, de algún modo, evolucionan en el espacio y en el tiempo.

Si se considera una región de estudio espacial no discreta $S \in \mathbb{R}^2$ y un marco temporal $T \in \mathbb{R}$, los datos se tratan como una colección de eventos instantáneos, cada uno de los cuales ocurre en una ubicación espacial \mathbf{s} en un tiempo t ; es decir, en t se observa como máximo un suceso en la ubicación \mathbf{s} . Por lo tanto, no se considera que un evento permanezca en \mathbf{s} después de su ocurrencia.

Un proceso puntual espacio-temporal de Poisson $Y(\mathbf{s}, t)$ en $S \times T \subseteq \mathbb{R}^2 \times \mathbb{R}$ satisface las siguientes condiciones:

- En cualquier conjunto acotado $A \subseteq S \times T$, $N(Y \cap A)$ sigue una distribución de Poisson con un número esperado $\int_A \lambda(\mathbf{s}, t) d_2(\mathbf{s}, t)$. Por ejemplo, si se supone $A = A_S \times A_T \subseteq S \times T$ entonces el valor esperado es el número esperado de puntos en la región $A_S \subseteq S$ y en el intervalo de tiempo $A_T \subseteq T$. Nótese que $d_2(\cdot, \cdot)$ indica la medida de Lebesgue en el espacio-tiempo.

- Para cualesquiera k subconjuntos arbitrarios disjuntos de $S \times T$, por ejemplo, A_1, A_2, \dots, A_k , su cardinalidad $N(Y \cap A_1), \dots, N(Y \cap A_k)$ son variables independientes.

Además, para cualquier intervalo de tiempo $[t_1, t_2] \subset T$ la proyección de Y sobre S define un proceso de Poisson espacial en S con función de intensidad $\int_{t_1}^{t_2} \lambda(\mathbf{s}, t) dt$. De manera similar, para cualquier conjunto de $A \subset S$ la proyección de Y sobre T define un proceso de Poisson temporal con función de intensidad $\int_A \lambda(\mathbf{s}, t) ds$ (Moradi y Mateu, 2020).

1.5.3. Características

A continuación se muestran varias definiciones fundamentales que caracterizan a los procesos puntuales espacio-temporales según Diggle et al. (2005) y González et al. (2016).

Un **proceso puntual espacio-temporal** $Y(\mathbf{s}, t)$ es un subconjunto aleatorio numerable de $\mathbb{R}^2 \times \mathbb{R}$, donde un punto (\mathbf{s}, t) corresponde a un evento en la región espacio-temporal, tal que $\mathbf{s} \in \mathbb{R}^2$ ocurre en el tiempo $t \in \mathbb{R}$. Se considera patrón puntual espacio-temporal a una realización del proceso puntual, dada por una colección de puntos $\{\xi\}_{i=1}^n = \{(\mathbf{s}_i, t_i)\}_{i=1}^n \subseteq S \times T$.

Definición. Una vecindad cilíndrica $B[(\mathbf{s}, t), r, h]$ centrada en $(\mathbf{s}, t) \in S \times T$ con un radio espacial $r > 0$ y un radio temporal $h > 0$, se define como

$$\begin{aligned} B[(\mathbf{s}, t), r, h] &= B[\mathbf{s}, r] \times [t - h, t + h] \\ &= \{(\mathbf{u}, l) \in S \times T : \|\mathbf{s} - \mathbf{u}\| \leq r, |t - l| \leq h\}, \end{aligned} \quad (1.19)$$

donde $B[\mathbf{s}, r] = \{\mathbf{u} \in S : \|\mathbf{s} - \mathbf{u}\| \leq r\}$ es una esfera euclídea con centro en $\mathbf{s} \in S$ y radio r , $\|\cdot\|$ representa la distancia euclídea y $|\cdot|$ el valor absoluto. En este contexto, $B_{rh} = B[(\mathbf{0}, 0), r, h]$ define una forma de medir las distancias espacio-temporales.

Definición. Sea $Y(\mathbf{s}, t)$ un proceso puntual espacio-temporal en $S \times T \subseteq \mathbb{R}^2 \times \mathbb{R}$. Se dice que $Y(\mathbf{s}, t)$ es completamente estacionario si el proceso desplazado $(\mathbf{u}, l) + Y(\mathbf{s}, t)$ con $(\mathbf{u}, l) \in S \times T$ tiene la misma distribución que el proceso original $Y(\mathbf{s}, t)$ para cualquier $(\mathbf{s}, t) \in S \times T$.

Definición. Sea $Y(\mathbf{s}, t)$ un proceso puntual espacio-temporal en $S \times T \subseteq \mathbb{R}^2 \times \mathbb{R}$. Se dice que $Y(\mathbf{s}, t)$ es espacialmente isotrópico si, para cualquier

1.5. PROCESOS ESPACIO-TEMPORALES

rotación \mathbf{r} alrededor del origen, el proceso rotado $\mathbf{r}Y(\mathbf{s}, t) = \{(\mathbf{r}\mathbf{s}, t) : (\mathbf{s}, t) \in Y\}$ tiene la misma distribución que $Y(\mathbf{s}, t)$.

Las densidades producto $\lambda^{(k)}$, con $k \geq 1$, son herramientas fundamentales en el análisis estadístico de procesos puntuales. $\lambda^{(k)}$ se pueden definir mediante el teorema de Campbell. Este teorema establece que, dado un proceso puntual espacio-temporal $Y(\mathbf{s}, t)$ en $\mathbb{R}^2 \times \mathbb{R}$, para cualquier función no negativa h en $(\mathbb{R}^2 \times \mathbb{R})^{(k)}$,

$$E \left[\sum_{\substack{\neq \\ \xi_1, \dots, \xi_k \in Y}} h(\xi_1, \dots, \xi_k) \right] = \int_{\mathbb{R}^2 \times \mathbb{R}} \cdots \int_{\mathbb{R}^2 \times \mathbb{R}} h(\xi_1, \dots, \xi_k) \lambda^{(k)}(\xi_1, \dots, \xi_k) \times \prod_{i=1}^k d\xi_i. \quad (1.20)$$

1.5.4. Función de intensidad espacio-temporal

Definición. La medida de intensidad Λ para un proceso puntual espacio-temporal está dada por $\Lambda(A \times B) = E[N(A \times B)]$, $A \times B \subseteq S \times T$. Si la medida de intensidad se puede escribir como

$$\Lambda(A \times B) = \int_A \int_B \lambda(\mathbf{s}, t) ds dt, \quad (1.21)$$

con $A \times B \subseteq S \times T$, donde $\lambda(\mathbf{s}, t)$ es una función no negativa, entonces $\lambda(\mathbf{s}, t)$ se llama función de intensidad de $Y(\mathbf{s}, t)$. Si $\lambda(\mathbf{s}, t) \equiv \lambda > 0$, entonces se dice que $Y(\mathbf{s}, t)$ es homogéneo o estacionario de primer orden con intensidad λ ; de lo contrario, se dice que $Y(\mathbf{s}, t)$ es no homogéneo.

Separabilidad de primer orden

Se habla de separabilidad de primer orden si la función de intensidad de primer orden se puede factorizar en espacio y tiempo; es decir, para $(\mathbf{s}, t) \in \mathbb{R}^2 \times \mathbb{R}$, $\lambda(\mathbf{s}, t)$ se representa como el producto de sus componentes marginales espacial y temporal,

$$\lambda(\mathbf{s}, t) = \lambda_1(\mathbf{s})\lambda_2(t), \quad (1.22)$$

y en su forma continua,

$$\Lambda(A \times B) = \int_A \lambda_1(\mathbf{s}) ds \int_B \lambda_2(t) dt, \quad (1.23)$$

donde $\lambda_1(\cdot)$ y $\lambda_2(\cdot)$ son funciones no negativas y no necesariamente únicas.

Un proceso estacionario espacio-temporal puntual $Y(\mathbf{s}, t)$ es automáticamente separable de primer orden ya que su intensidad $\lambda = \lambda_1 \lambda_2 \geq 0$ es constante. Si $Y(\mathbf{s}, t)$ es estacionario en el espacio, implica que $\lambda(\mathbf{s}, t)$ depende sólo de t ; por tanto, también es separable de primer orden siendo λ_1 una constante no negativa. Del mismo modo, cuando $Y(\mathbf{s}, t)$ es estacionario en el tiempo, lo que implica que $\lambda(\mathbf{s}, t)$ dependa sólo de \mathbf{s} , la separabilidad de primer orden se mantiene con λ_2 y es una constante no negativa. Se definen las funciones marginales de intensidad espacial y temporal como sigue:

$$\begin{aligned}\lambda(\mathbf{s}) &= \lambda_1(\mathbf{s}) \int_T \lambda_2(t) dt, \\ \lambda(t) &= \lambda_2(t) \int_S \lambda_1(\mathbf{s}) ds,\end{aligned}\tag{1.24}$$

por lo que satisface $\lambda(\mathbf{s}, t) \propto \lambda(\mathbf{s})\lambda(t)$ si $Y(\mathbf{s}, t)$ es estacionario. Para $Y(\mathbf{s}, t)$ con intensidad homogénea λ , $\lambda(\mathbf{s})$ y $\lambda(t)$ serían constantes.

Estimación

Considerando que generalmente se dispone de una única realización, al estimar la función de intensidad de primer orden (conjunta espacio-temporal) la tarea es encontrar estimadores para $\hat{\lambda} : S \times T \rightarrow \mathbb{R}$ que sean insesgados.

Si se supone separabilidad de primer orden, se puede encontrar $\hat{\lambda}$ a través de estimadores insesgados de sus componentes marginales, espacial $\hat{\lambda}(\mathbf{s})$ y temporal $\hat{\lambda}(t)$, de la forma

$$\hat{\lambda}(\mathbf{s}, t) = \frac{1}{n} \left(\hat{\lambda}(\mathbf{s}) \hat{\lambda}(t) \right),\tag{1.25}$$

que también constituye un estimador insesgado del número esperado de puntos. Para estimaciones no paramétricas, [Diggle et al. \(2005\)](#) muestran una estimación para $\hat{\lambda}$ mediante un método de suavizado con un núcleo gaussiano, y para la componente temporal mediante un modelo de regresión estándar log-lineal de Poisson, debido a que $\lambda(t)$ representa el valor esperado incondicional del número de casos en un tiempo t_i (véase [Diggle et al., 2005](#), [Taylor et al., 2015](#) y [González et al., 2016](#)). Se pueden consultar otros modelos sobre estimación de $\hat{\lambda}(\mathbf{s})$ y $\hat{\lambda}(t)$ no paramétricos en [Choi y Hall \(2001\)](#) y [Contreras Roza \(2018\)](#). Para un método de suavizado en bases polinomiales, véanse los capítulos 2 y 3, y para un semi-paramétrico véanse [Zhuang y Mateu \(2019\)](#) y el capítulo 4.

Funciones de intensidad condicionadas

La función de intensidad condicionada $\lambda^*(\mathbf{s}, t | \mathcal{H}_t)$ de un proceso puntual espacio-temporal es la tasa esperada de que se produzcan puntos en torno a

1.5. PROCESOS ESPACIO-TEMPORALES

la localización espacio-temporal (\mathbf{s}, t) condicionada a la historia \mathcal{H}_t (conjunto de localizaciones y tiempos de todos los eventos del proceso que ocurren antes del tiempo t) con $t \in T$. Es decir, \mathcal{H}_t es la familia de σ -álgebras generada por los eventos que ocurren en tiempos hasta t , pero sin incluir t (González et al., 2016; Zhuang y Mateu, 2019). Según González et al. (2016) y referencias incluidas, se tiene que

$$\lambda^*(\mathbf{s}, t | \mathcal{H}_t) ds dt = E[\mathcal{N}(ds \times dt) | \mathcal{H}_t], \quad (1.26)$$

con $(\mathbf{s}, t) \in ds \times dt \subseteq S \times T$.

Factor de efecto de borde

El factor de corrección de borde es el producto de un factor de corrección de borde espacial y otro temporal. Los *efectos de borde* surgen porque los puntos que aparecen fuera de los límites de la región de estudio no son tomados en cuenta para estimar los estadísticos de resumen, por lo que producen estimaciones sesgadas.

Los mecanismos y estimadores para corregir el efecto de borde están bien sustentados para el caso espacial, en particular la corrección de bordes isotrópicos (véase Gabriel, 2014; González et al., 2016). Del mismo modo se da la extensión al caso espacio-temporal, en donde los más comunes son: el isotrópico, de bordes modificados y de traslación. En general, se determina un factor de corrección de bordes en función de ciertos pesos. La descripción de cada uno de los métodos puede verse en Gabriel (2014), González et al. (2016).

Medidas de momento de segundo orden

Permiten determinar si los puntos tienden a formar agregaciones o presentar dispersión (regularidad) en un rango de distancias. Esto se logra mediante la función de correlación por pares, medidas de momentos de segundo orden reducidas (funciones L y K) y los estadísticos de resumen basadas en las distancias entre puntos (como las funciones F , G y J).

- **Función de correlación por pares**

Las medidas de interacción espacio-temporal de segundo orden, en particular en presencia de no homogeneidad, están dadas por la función de correlación por pares,

$$g(\xi_1, \xi_2) = \frac{\lambda^2(\xi_1, \xi_2)}{\lambda(\xi_1)\lambda(\xi_2)}, \quad (1.27)$$

con $\xi_1, \xi_2 \in S \times T$. Si los puntos son independientes unos de otros como es el caso del proceso de Poisson, $g(\cdot, \cdot) = 1$. Cuando $g(\cdot, \cdot) > 1$ indica la presencia de atracción entre los puntos (agregación), mientras que $g(\cdot, \cdot) < 1$ indica repulsión (regularidad).

Al igual que en el caso de la separabilidad de primer orden, se dice que la función de correlación por pares es separable si

$$g((\mathbf{s}, t), (\mathbf{u}, l)) = g_1(\mathbf{s}, \mathbf{u})g_2(t, l), \quad (1.28)$$

donde g_1 y g_2 son funciones no negativas (Moller y Ghorbani, 2012).

Definición. El proceso puntual espacio-temporal $Y(\mathbf{s}, t)$ es estacionario de segundo orden ponderado por la intensidad si

$$g((\mathbf{s}, t), (\mathbf{u}, l)) = \bar{g}(\mathbf{s} - \mathbf{u}, t - l), \quad (1.29)$$

para cualquier $(\mathbf{s}, t), (\mathbf{u}, l) \in S \times T$, donde \bar{g} es alguna función no negativa.

Si el proceso también es isotrópico, entonces $\bar{g}(\mathbf{s} - \mathbf{u}, t - l) = g_0(r, h)$, es decir, $g(\cdot, \cdot)$ sólo depende de las distancias $r = \|\mathbf{s} - \mathbf{u}\|$ y $h = |t - l|$, donde g_0 es alguna función no negativa. La función de correlación por pares puede extenderse a órdenes generales $k \geq 2$ de la forma

$$g^k(\xi_1, \dots, \xi_k) = \frac{\lambda^k(\xi_1, \dots, \xi_k)}{\prod_{i=1}^k \lambda(\xi_i)}, \quad (1.30)$$

con $\xi_1, \dots, \xi_k \in S \times T$. Su estimación no paramétrica está dada por

$$\hat{g}(r, h) = \frac{1}{4\pi r} \sum_{i=1}^n \sum_{j=1, j \neq i}^n \frac{\mathfrak{K}_{1\epsilon}(\|\mathbf{s}_i - \mathbf{s}_j\| - r) \mathfrak{K}_{2\delta}(|t_i - t_j| - l)}{\hat{\lambda}(\mathbf{s}_i, t_i) \hat{\lambda}(\mathbf{s}_j, t_j) w_{ij}}, \quad (1.31)$$

con $r > \epsilon$, $t > \delta$, donde $\mathfrak{K}_{1\epsilon}$ y $\mathfrak{K}_{2\delta}$ son funciones núcleo unidimensionales con anchos de banda espacial y temporal ϵ y δ , respectivamente (véanse más detalles sobre funciones núcleo en Baddeley et al., 2015). Finalmente, w_{ij} son pesos (factores de corrección de bordes) que corrigen la pérdida de información relativa a la interacción que se produce entre los puntos cercanos a la frontera de $S \times T$, y sus valores dependerán del método elegido.

Intuitivamente, la probabilidad de observar un par de eventos de $Y(\mathbf{s}, t)$ que ocurren simultáneamente en cada uno de los dos conjuntos infinitesimalmente pequeños con centros (\mathbf{s}, t) , (\mathbf{u}, l) , y volúmenes $d(\mathbf{s}, t)$, $d(\mathbf{u}, l)$, está dada por el producto de una función de las localizaciones

1.5. PROCESOS ESPACIO-TEMPORALES

(\mathbf{s}, \mathbf{u}) y las áreas $(ds, d\mathbf{u})$, y una función que depende de los tiempos (t, l) y las longitudes (dt, dl) , es decir,

$$\left[\hat{\lambda}_1(\mathbf{s}) \hat{\lambda}_1(\mathbf{u}) g_1(\mathbf{s}, \mathbf{u}) ds du \right] \times \left[\hat{\lambda}_2(t) \hat{\lambda}_2(l) g_2(t, l) dt dl \right]. \quad (1.32)$$

- **Medidas de momentos de segundo orden reducidas**

Función K. Representa el número esperado de puntos en una esfera de radio r cuyo centro es un punto en Y no considerado en el conteo. Sea $Y(\mathbf{s}, t)$ un proceso puntual espacio-temporal de intensidad estacionaria reponderada de segundo orden; entonces,

$$K(r, h) = \int 1 [\|\mathbf{s}\| \leq r, t \leq h] g(\mathbf{s}, t) d(\mathbf{s}, t), \quad (1.33)$$

con $r > 0$ y $h > 0$, donde $\|\mathbf{s}\|$ es la distancia euclídea en \mathbb{R}^2 y t es un valor absoluto.

En el caso estacionario, $\lambda K(\mathbf{s}, t)$ puede interpretarse como el número esperado de puntos dentro de un cilindro de radio r y altura $2h$ (para algunas especificaciones sobre la función K , véase [Van Lieshout y Baddeley, 1996](#); [Moller y Ghorbani, 2012](#); [Rodríguez Cortés, 2014](#); [González et al., 2016](#); [Gabriel, 2014](#), entre otros autores).

- **Estadísticos de resumen**

Los estadísticos de resumen basados en las distancias entre puntos, como las funciones F , G y J , son utilizados para cuantificar las interacciones en los procesos espacio-temporales no homogéneos.

Función F. Es la función de distribución de la distancia desde un origen, u otro punto fijo en \mathbb{R}^2 , hasta el punto más cercano en Y .

Función G. Es conocida también como *función al vecino más cercano*, y se interpreta como la función de distribución de la distancia desde un punto típico hasta el vecino más cercano en Y .

Función J. Fue propuesta por [Van Lieshout y Baddeley \(1996\)](#), y está dada de la siguiente forma:

$$J(r) = \frac{(1 - G(r))}{(1 - F(r))}, \quad (1.34)$$

para $F(r) < 1$. Para un proceso Poisson homogéneo en \mathbb{R}^2 , se cumple $J(r) < 1$. En general, para valores pequeños de $r > 0$, $F(r) < G(r)$ (o $J(r) < 1$) indica agregación o clustering, y $F(r) > G(r)$ (o $J(r) > 1$) regularidad ([Moller y Waagepetersen, 2003](#)). (Para más detalles sobre cada una de estas funciones véase [Van Lieshout y Baddeley, 1996](#); [Rodríguez Cortés, 2014](#); [González et al., 2016](#), entre otros autores).

1.5.5. Clasificación

Rodríguez Cortés (2014) y González et al. (2016) exponen una clasificación de los modelos espacio-temporales más recientes. Sin embargo, al identificar que la distribución estacionaria del proceso espacio-temporal INGARCH es equivalente a un proceso auto-excitado dado por Hawkes (1971a), y que los modelos SPINGARCH conservan las propiedades de los INGARCH, se los incluye en la clasificación. Por lo tanto, los modelos espacio-temporales se pueden clasificar de la siguiente forma:

Modelos espacio-temporales empíricos

- Procesos de Poisson homogéneos
- Procesos de Poisson no homogéneos
- Procesos de Neyman-Scott
- Procesos geométricos anisotrópicos de Poisson
- Procesos de inhibición
- Procesos de Strauss
- Procesos de Cox
- Procesos de Cox log-gaussianos
- Procesos estacionarios de Poisson y shot-noise de Cox

Modelos espacio-temporales mecanicistas

- Procesos de Poisson
- Procesos de auto-excitación
 - Procesos de Hawkes semi-paramétricos
 - Procesos Integer Generalized Auto-Regressive Conditionally Heteroscedastic (INGARCH)
 - Procesos Spatial Integer Generalized Auto-Regressive Conditionally Heteroscedastic Correlation (SPINGARCH)
 - Procesos de secuencias de réplicas de tipo epidémico (ETAS)

1.6. Medidas de riesgo

En esta sección se describen algunas definiciones y teoremas elementales necesarios para el análisis y cuantificación de las probabilidades de eventos extremos.

Se toman principalmente las definiciones y teoremas descritos de Salvadori et al. (2007), Malevergne y Sornette (2006), Markose y Alentorn (2005), Abad et al. (2014), McNeil y Frey (2000), Coles (2001), entre otros autores.

1.6.1. Teoría de valores extremos

La teoría de valores extremos o EVT (Extreme Value Theory) es el área de la estadística dedicada a desarrollar modelos y técnicas para estimar (cuantificar) el comportamiento estocástico de los eventos inusuales o raros de uno o varios procesos, es decir, de aquellos valores que pertenecen a las colas de la distribución y están alejados de su media o mediana. La EVT es una herramienta que trata de dar la mejor estimación posible de la cola de la distribución; es útil incluso en ausencia de datos históricos, ya que permite la modelización del comportamiento empírico basándose en el conocimiento de la distribución asintótica.

Tiene una amplia aplicación en medio ambiente, hidrología, meteorología, geología, finanzas, seguros, entre otras áreas.

Valores extremos univariantes

El análisis de valores extremos univariantes se desarrolla a partir de los resultados de Fréchet (1927), Fisher y Tippett (1928) y Von Mises (1936) (Salvadori et al., 2007), que sientan las bases de la teoría probabilística de valores extremos unidimensional. Al modelizar los máximos de una variable aleatoria, la EVT desempeña el mismo papel fundamental que el teorema del límite central al modelizar la suma de variables aleatorias. En ambos casos, la teoría indica cuáles son las distribuciones límite.

Comúnmente se presentan dos enfoques para identificar los extremos en los datos reales.

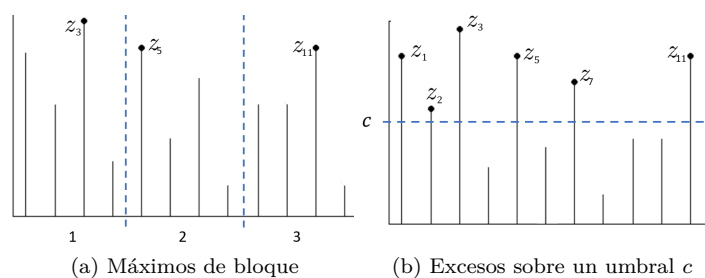


Figura 1.4: Enfoques para identificar valores extremos en datos reales

Máximos por bloques

Es el método tradicional, utilizado para analizar datos con estacionalidad. Considera el máximo o el mínimo que toma la variable en períodos sucesivos,

por ejemplo, meses, años, etc. En la Figura 1.4a se muestran observaciones z_3 , z_5 y z_{11} , que representan máximos de bloque para tres períodos con cuatro observaciones en cada período.

Sea Z_1, \dots, Z_n una secuencia de variables aleatorias i.i.d. con función de distribución común F y M_n el máximo:

$$M_n = \max\{Z_1, \dots, Z_n\}. \quad (1.35)$$

La distribución de M_n podría obtenerse de manera exacta a partir de la distribución de las n variables, teniendo en cuenta las propiedades de independencia (véase Coles, 2001).

La ley límite para los máximos de bloque consiste en la búsqueda de secuencias de constantes $\{b_n; n \geq 1\}$ y $\{a_n; n \geq 1\}$ (véase más detalles sobre su cálculo en Salvadori et al., 2007) tales que la distribución de $M_n^* = \frac{M_n - b_n}{a_n}$ converge a una distribución no degenerada cuando $n \rightarrow \infty$, es decir,

$$\lim_{n \rightarrow \infty} F^n(a_n z + b_n) = G(z). \quad (1.36)$$

El rango completo de distribuciones límite que podrá seguir M_n^* vendrá dado por el teorema de valores extremos de Fisher-Tippett (1928).

Teorema. Si existen sucesiones de constantes $\{a_n > 0\}$, $\{b_n\}$, y alguna función de distribución no degenerada G tal que

$$P\left(\frac{M_n - b_n}{a_n} \leq z\right) = F^n(a_n z + b_n) \rightarrow G(z), \quad (1.37)$$

cuando $n \rightarrow \infty$, para todos los puntos donde G es continua, entonces $G(z)$, con $z \in \mathbb{R}$, pertenece a alguna de las distribuciones siguientes:

1. *Gumbel*:

$$G(z) = \exp\left\{-\exp\left[-\frac{z-b}{a}\right]\right\}, \quad -\infty < z < \infty \quad (1.38)$$

2. *Fréchet*:

$$G(z) = \begin{cases} 0, & z \leq b \\ \exp\left\{-\left(\frac{z-b}{a}\right)^{-\eta}\right\}, & z > b \end{cases} \quad (1.39)$$

3. *Weibull*:

$$G(z) = \begin{cases} \exp\left\{-\left(-\frac{z-b}{a}\right)^\eta\right\}, & z < b \\ 1, & z \geq b \end{cases} \quad (1.40)$$

1.6. MEDIDAS DE RIESGO

con $a > 0$ parámetro de escala, $b \in \mathbb{R}$ parámetro de posición y $\eta > 0$ parámetro de forma.

Las tres distribuciones mencionadas se conocen como **distribuciones de valores extremos**. Este resultado es muy importante, ya que la distribución asintótica de los máximos M_n^* siempre pertenece a una de estas tres distribuciones, independientemente de la distribución original F .

Según Markose y Alentorn (2005), Jenkinson (1955) y Von Mises (1936) propusieron una única representación paramétrica estándar para las tres distribuciones, conocida como **Distribución de Valores Extremos Generalizada** o GEVD (Generalized Extreme Value Distribution). Para ello se establece $\xi = \eta^{-1}$ para la distribución de Fréchet, $\xi = -\eta^{-1}$ para Weibull y el caso límite $\xi = 0$ para la distribución de Gumbel, dando como resultado

$$GEV_{\xi, \mu, \sigma}(z) = \begin{cases} \exp \left\{ - \left[1 + \xi \left(\frac{z - \mu}{\sigma} \right) \right]^{-1/\xi} \right\}, & \xi \neq 0 \\ \exp \left\{ - \exp \left[- \frac{z - \mu}{\sigma} \right] \right\}, & \xi = 0 \end{cases} \quad (1.41)$$

con $\{z : 1 + \xi(z - \mu)/\sigma > 0\}$ y μ parámetro de posición, σ parámetro de escala y ξ parámetro de forma, satisfaciendo $-\infty < \mu < \infty$, $\sigma > 0$ y $-\infty < \xi < \infty$. El parámetro de forma indica el espesor de la cola de la distribución. Si el valor del parámetro es alto, la cola de la distribución es más pesada. Si $\xi < 0$ la variable Z tendrá distribución de Weibull, si $\xi = 0$ distribución de Gumbel y si $\xi > 0$ distribución de Fréchet.

Teorema. Si existen sucesiones de constantes $\{a_n > 0\}$ y $\{b_n\}$ tales que

$$P \left[\frac{M_n - b_n}{a_n} \leq z \right] \rightarrow G(z), \quad (1.42)$$

cuando $n \rightarrow \infty$, para una distribución G no degenerada, entonces G pertenece a la familia de distribuciones GEV,

$$G(z) = \exp \left\{ - \left[1 + \xi \left(\frac{z - \mu}{\sigma} \right) \right]^{-1/\xi} \right\}. \quad (1.43)$$

La representación generalizada es especialmente útil cuando hay que calcular estimaciones de máxima verosimilitud, ya que, en general, no se conoce de antemano el tipo de distribución límite de la muestra.

Este último teorema está estrechamente relacionado con la propiedad de estabilidad maximal.

Definición. Una distribución G se denomina max-estable si, para todo $n = 2, \dots$, existen constantes $a_n > 0$ y b_n tales que,

$$G^n(a_n z + b_n) = G(z), \quad \text{con } z \in \mathbb{R}. \quad (1.44)$$

Es decir, una distribución es max-estable si y solo si es una distribución de GEV. Esto conduce a que al tomar máximos muestrales se genere una distribución idéntica, pero con distintos parámetros de localización y escala.

Excesos sobre un umbral

Estos métodos utilizan los datos de forma más eficiente, concentrándose en las realizaciones que superan un determinado umbral (alto). Por ejemplo, en la Figura 1.4b, las observaciones z_1, z_2, z_3, z_5, z_7 y z_{11} superan el umbral c y constituyen eventos extremos. Se les conoce también como *picos sobre un umbral* o POT (Peaks Over Threshold). Y por ello, es el método más utilizado, a pesar de que, en general, se dispondrá de muy pocas observaciones para umbrales altos.

Sea Z_1, Z_2, \dots una secuencia de variables aleatorias i.i.d. con función de distribución marginal F , y sean \mathcal{Z} las excedencias $\mathcal{Z} = Z - c$. El comportamiento de \mathcal{Z} vendrá dado por la probabilidad condicionada

$$P[Z > c + z | Z > c] = \frac{1 - F(c + z)}{1 - F(c)}, \quad z > 0. \quad (1.45)$$

F y la distribución de las excedencias del umbral deberían ser conocidas; sin embargo, en la práctica no sucede así, por lo que se hallan aproximaciones GEV.

Según el teorema de Pickands (Pickands, 1975), para un umbral suficientemente alto, los excesos de la variable Z sobre c , dado que $Z > c$, tienen aproximadamente **Distribución de Pareto Generalizada** o GPD (Generalized Pareto Distribution); es decir, la función de distribución asintótica de $(Z - c)$, condicionada a $Z > c$, está dada por

$$GPD_{\sigma, \xi}(z) = \begin{cases} 1 - \left\{ \left[1 + \xi \left(\frac{z}{\tilde{\sigma}} \right) \right]^{-1/\xi} \right\}, & \xi \neq 0 \\ 1 - \exp\left(-\frac{z}{\tilde{\sigma}}\right), & \xi = 0 \end{cases} \quad (1.46)$$

con $z \geq 0$, $\left(1 + \frac{\xi z}{\tilde{\sigma}}\right)^{1/\xi} > 0$ y $\tilde{\sigma} = \sigma + \xi(c - \mu)$; $-\infty < \xi < +\infty$ parámetro de forma y $\tilde{\sigma}$ parámetro de escala. Para una GPD estándar se denota GPD_{ξ} .

El comportamiento de la GPD está determinado por ξ , es decir:

- Si $\xi < 0$, la distribución está acotada superiormente, $0 < z < -\frac{\tilde{\sigma}}{\xi}$, y la distribución pertenece al dominio de atracción de Weibull.
- Si $\xi = 0$, la distribución pertenece al dominio de atracción de Gumbel, para $0 < z < +\infty$.
- Si $\xi > 0$, la distribución pertenece al dominio de atracción de Fréchet, para $0 < z < +\infty$.

En lo que refiere a $\tilde{\sigma} = \sigma$, por simplicidad, se utilizará σ en lugar de $\tilde{\sigma}$.

Selección del umbral

Para elegir el umbral adecuado se debe tomar en cuenta que un umbral muy bajo aumentaría el sesgo del modelo, mientras que un umbral demasiado alto generaría un número de excedencias muy pequeño, incrementando su varianza. Para ello, se dispone de varios métodos; sin embargo, aquí se detallan los dos más comunes:

- **Esperanza de la GPD.** Está basado en la esperanza de la GPD, de forma que, si Z sigue esta distribución con parámetros σ y ξ ,

$$E[Z] = \frac{\sigma}{1 - \xi}, \text{ con } \xi > 1. \quad (1.47)$$

Si la GPD es válida para las excedencias del umbral c_0 , para un término arbitrario Z de la serie, tendrá

$$E[Z - c_0 | Z > c_0] = \frac{\sigma_{c_0}}{1 - \xi}, \text{ con } \xi > 1, \quad (1.48)$$

donde σ_{c_0} es el parámetro de escala de la distribución de excedencias del umbral c_0 . Esto quiere decir que, si la GPD es válida para las excedencias de c_0 , será igualmente válida para cualquier umbral $c > c_0$ mediante una sustitución del parámetro de escala por σ_c . Entonces se tiene que

$$E[Z - c | Z > c] = \frac{\sigma_{c_0} + \xi c}{1 - \xi}. \quad (1.49)$$

Se espera que los resultados de la expresión (1.49) varíen linealmente según como la variación de c se ajuste a la GPD apropiada. Esto conduce a

$$\left\{ \left(c, \frac{1}{n_c} \sum_{i=1}^{n_c} (z_{(i)} - c) \right) : c < z_{\text{máx}} \right\}, \quad (1.50)$$

siendo $z_{(1)}, z_{(2)}, \dots, z_{(n_c)}$ las observaciones que exceden el umbral c , y $z_{\text{máx}}$ el máximo de la Z_i , dando como resultado el gráfico de vida residual media (véase la Figura 1.5). El umbral debe estar dentro del intervalo donde el gráfico muestre una tendencia lineal.

- **Ajuste de la GPD.** Consiste en ajustar la GPD en un rango de umbrales, sobre el cual se elegirá c de acuerdo a la presencia de estabilidad de los parámetros estimados. Si una GPD es razonable para modelizar las excedencias de un umbral c_0 , también lo será para $c > c_0$, con parámetros de forma ξ idénticos y parámetros de escala que satisfacen la relación

$$\sigma_c = \sigma_{c_0} + \xi(c - c_0), \text{ siempre que } \xi \neq 0, \quad (1.51)$$

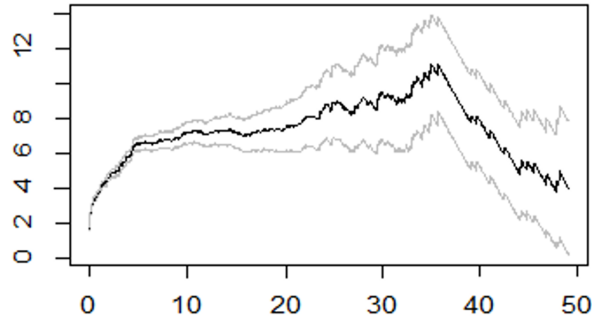


Figura 1.5: Vida residual media

donde σ_c es el valor del parámetro de escala de la GPD para un umbral $c > c_0$. Este problema se puede solucionar reparametrizando σ_c como

$$\sigma^* = \sigma_c - \xi_c, \quad (1.52)$$

considerándola como constante con respecto a c . Las estimaciones de σ^* y de ξ deberían ser constantes por encima de c_0 si c_0 es un umbral válido para que los excesos sean GPD (Coles, 2001). Los intervalos de confianza para ξ se obtendrán a través de la matriz de varianzas-covarianzas, mientras que los intervalos de confianza para $\tilde{\sigma}^*$ requieren el uso del método delta.

También existen modelos semi-paramétricos construidos en torno al estimador de Hill (véase Beirlant et al., 1996; Ciuperca y Mercadier, 2010; Abad et al., 2014; Mora Valencia, 2010).

1.6.2. Valores extremos multivariantes

La teoría matemática de valores extremos multivariantes (MEV) es un campo relativamente novedoso. Diversos aspectos están bien desarrollados y existen análogos de los modelos de bloque y umbral. Como en el caso univariante, estos modelos sólo tienen justificaciones asintóticas, y su idoneidad para cualquier aplicación práctica debe comprobarse con cuidado (Salvadori et al., 2007).

El análisis de los MEV se basa en una ordenación por componentes. Para una serie vectorial i.i.d. $(Z_{i,1}, \dots, Z_{i,p})$, $i = 1, \dots, n$, con función de distribución conjunta F y marginales F_1, \dots, F_p , el vector de máximos por componentes, M_n , viene dado por $M_n = (M_{n,1}, \dots, M_{n,p})$, donde $M_{n,j} = \max\{Z_{1,j}, \dots, Z_{n,j}\}$, para $j = 1, \dots, p$. Bajo condiciones generales, $M_{n,j}$ convenientemente normalizado converge en distribución a un miembro de la clase de distribuciones de valores extremos multivariantes o MEVD (Multivariate Extreme Value Distributions).

Distribuciones de valores extremos multivariantes

Sean $Z_i = (Z_{i,1}, \dots, Z_{i,p})$, con $i = 1, \dots, n$, vectores aleatorios i.i.d. de dimensión p con función de distribución F . Como en el caso univariante, una forma estándar de operar es buscar la existencia de secuencias de constantes $\{a_{n,j} > 0\}$ y $\{b_{n,j}\}$, $\forall j = 1, \dots, p$, y una función de distribución G con marginales no degeneradas, tales que, cuando $n \rightarrow \infty$,

$$\begin{aligned} G(z_1, \dots, z_p) &= \lim_{n \rightarrow \infty} P \left\{ \frac{M_{n,1} - b_{n,1}}{a_{n,1}} \leq z_1, \dots, \frac{M_{n,p} - b_{n,p}}{a_{n,p}} \leq z_p \right\} \\ &= \lim_{n \rightarrow \infty} F^n(a_{n,1}z_1 + b_{n,1}, \dots, a_{n,p}z_p + b_{n,p}) \rightarrow G(z_p), \end{aligned} \quad (1.53)$$

con $Z = (Z_1, \dots, Z_p) \in \mathbb{R}^p$, donde G es la función de MEVD y F pertenece al dominio de atracción maximal de G .

Se han propuesto varios métodos de análisis de excedencias de umbrales, todos ellos basados en la aproximación de F en una región extremal definida usando las condiciones del dominio de atracción especificado. La especificación de las distribuciones multivariantes en términos de cópulas proporciona un enfoque muy útil en este contexto.

1.6.3. Cópulas

Uno de los procedimientos más comunes al analizar eventos extremos multivariantes es el uso de cópulas, por su capacidad de reflejar diversos tipos de dependencia entre dos o más variables aleatorias, independientemente de su distribución marginal; más concretamente, las cópulas permiten representar la estructura de dependencia existente entre varias funciones de densidad univariantes o multivariantes, ya que conducen a una función de densidad conjunta.

Sean X e Y dos variables aleatorias con función de distribución conjunta denotada por H ,

$$H(x, y) = P[X \leq x; Y \leq y], \quad (1.54)$$

con marginales F_1 y F_2 continuas y estrictamente crecientes, por lo que existen sus inversas F_1^{-1} y F_2^{-1} . Entonces, la cópula está dada por

$$\mathcal{C}(u, v) = H(F_1^{-1}(u), F_2^{-1}(v)), \quad \forall u, v \in [0, 1]. \quad (1.55)$$

La función \mathcal{C} es el único objeto que obedece a la propiedad de invariancia bajo mapeo estrictamente creciente y que captura completamente la dependencia total entre X e Y . Es importante destacar las siguientes propiedades:

1. $\mathcal{C}(u, 1) = u$ y $\mathcal{C}(1, v) = v$, $\forall u, v \in [0, 1]$.

-
2. $\mathcal{C}(u, 0) = \mathcal{C}(0, v) = 0, \forall u, v \in [0, 1]$.
 3. \mathcal{C} es 2-creciente, es decir, $\forall u_1 \leq u_2$ y $v_1 \leq v_2$: $\mathcal{C}(u_2, v_2) - \mathcal{C}(u_2, v_1) - \mathcal{C}(u_1, v_2) + \mathcal{C}(u_1, v_1) \geq 0$.

Esta última propiedad es una simple traducción de la no negatividad de las probabilidades, concretamente de

$$P [F_1^{-1}(u_1) \leq X \leq F_1^{-1}(u_2); F_2^{-1}(v_1) \leq Y \leq F_2^{-1}(v_2)] . \quad (1.56)$$

Estas tres propiedades definen el objeto matemático llamado *cópula* introducido por [Sklar \(1959\)](#).

Definición. Una función $\mathcal{C} : [0, 1]^n \rightarrow [0, 1]$ es una *n-cópula* si goza de las siguientes propiedades:

- $\forall u \in [0, 1], \mathcal{C}(1, \dots, 1, u, 1, \dots, 1) = u$.
- $\forall u_i \in [0, 1], \mathcal{C}(u_1, \dots, u_n) = 0$ si al menos una de las u_i es igual a cero.
- \mathcal{C} es *n-creciente*, es decir, el \mathcal{C} -volumen cuyos vértices se encuentran en $[0, 1]^n$ es positivo.

De esta definición se deriva que una cópula no es más que una distribución multivariante con soporte en $[0, 1]^n$ y con marginales uniformes. Se deduce inmediatamente que una suma convexa de cópulas sigue siendo una cópula. Las cópulas son muy útiles para representar distribuciones multivariantes ([Malevergne y Sornette, 2006](#)).

Teorema de Sklar. Dada una función de distribución *n-dimensional* F con marginales continuas F_1, \dots, F_n , existe una única *n-cópula* $\mathcal{C} : [0, 1]^n \rightarrow [0, 1]$ tal que,

$$F(x_1, \dots, x_n) = \mathcal{C}(F_1(x_1), \dots, F_n(x_n)). \quad (1.57)$$

Este teorema proporciona tanto una parametrización de las distribuciones multivariantes como un esquema de construcción de cópulas. En efecto, dada una distribución multivariante F con marginales F_1, \dots, F_n , la función

$$\mathcal{C}(u_1, \dots, u_n) = F(F_1^{-1}(u_1), \dots, F_n^{-1}(u_n)) \quad (1.58)$$

es automáticamente una *n-cópula* de la distribución multivariante F ([Malevergne y Sornette, 2006](#)).

Familias de cópulas

Como se muestra en el teorema de Sklar, para cada distribución multivariante se puede derivar fácilmente una cópula. Sin embargo, solo unas pocas familias de cópulas desempeñan un papel relevante.

A continuación se describen en forma rápida algunas de las familias de cópulas:

- **Cópulas elípticas.** Se derivan de las distribuciones elípticas multivariantes, siendo las más importantes las de Gauss y de Student. Por construcción, se acercan la una a la otra en su parte central, y en su cola sólo cuando aumenta los grados de libertad de la Student. Como consecuencia, a veces es difícil distinguir entre ellas, incluso con pruebas sensibles. Pueden tener comportamientos drásticamente diferentes con respecto a la dependencia entre los extremos (véanse detalles de cada una de estas cópulas en [Malevergne y Sornette, 2006](#); [Salvadori et al., 2007](#), entre otros).
- **Cópulas arquimedianas.** Engloban múltiples cópulas y gozan de ciertas propiedades interesantes. Se han desarrollado diversos modelos para dar cuenta de la dependencia entre múltiples fuentes de riesgo.

Definición. Sea φ una función continua estrictamente decreciente, convexa, de $[0, 1]$ en $[0, \infty]$ tal que $\varphi(1) = 0$. Sea $\varphi^{[-1]}$ la pseudoinvertida de φ , es decir,

$$\varphi^{[-1]}(t) = \begin{cases} \varphi^{-1}(t) & \text{si } 0 \leq t \leq \varphi(0) \\ 0 & \text{si } t \geq \varphi(0). \end{cases} \quad (1.59)$$

Entonces,

$$\mathcal{C}(u, v) = \varphi^{[-1]}(\varphi(u) + \varphi(v)) \quad (1.60)$$

es una cópula arquimediana con generador φ .

La generalización a una n -cópula esta dada por

$$\mathcal{C}(u_1, \dots, u_n) = \varphi^{[-1]}(\varphi(u_1) + \dots + \varphi(u_n)) = \mathcal{C}_n; \quad (1.61)$$

esta formulación se cumple (es decir, \mathcal{C}_n es en realidad una cópula n -arquimediana) si y sólo si $\varphi^{[-1]}$ es n -monótona,

$$(-1)^k \frac{d^k \varphi^{[-1]}(t)}{dt^k} \geq 0, \quad \forall k = 0, 1, \dots, n. \quad (1.62)$$

Cuando esta última relación se cumple para todo $n \in \mathbb{N}$, se dice que $\varphi^{[-1]}$ es completamente monótona. En tal caso, la cópula bivalente arquimediana se puede generalizar a cualquier dimensión.

La complejidad de la estructura de dependencia entre n variables se integra en una función de una sola variable φ . Esto transforma una formulación multidimensional en otra mucho más sencilla de carácter unidimensional. Las más comunes son:

— **Cópula de Clayton.** Es una cópula límite, dada por

$$\mathcal{C}_\theta^{Cl}(u, v) = \max\left([u^{-\theta} + v^{-\theta} - 1]^{-\frac{1}{\theta}}, 0\right), \quad (1.63)$$

con $\theta \in [-1, \infty)$ y generador $\varphi(t) = \frac{1}{\theta}(t^{-1} - 1)$.

— **Cópula de Gumbel.** Describe la dependencia mediante la teoría de valores extremos, y está dada por

$$\mathcal{C}_\theta^G(u, v) = \exp\left(-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{\frac{1}{\theta}}\right), \quad (1.64)$$

con $\theta \in [1, \infty)$ y generador $\varphi(t) = (-\ln t)^\theta$.

— **Cópula de Frank.** Está dada por

$$\mathcal{C}_\theta^F(u, v) = \frac{1}{\theta} \ln \left(1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1} \right), \quad (1.65)$$

con $\theta \in \mathbb{R}$ y generador $\varphi(t) = -\ln \frac{e^{\theta t} - 1}{e^{-\theta} - 1}$.

Un procedimiento general para construir los generadores de cópulas arquimedianas ha sido propuesto por Marshall y Olkin (1988) (véanse más detalles en [Malevergne y Sornette, 2006](#)).

Cópulas de valor extremo

Otra familia de cópulas de uso común son las cópulas de valores extremos. Se derivan de la estructura de dependencia de la distribución GEV multivariante; además, son capaces de representar procesos de max-estabilidad multivariantes.

Definición. Una d -cópula \mathcal{C} que satisface la relación

$$\mathcal{C}(u_1^k, \dots, u_d^k) = \mathcal{C}^k(u_1, \dots, u_d), \quad \forall k > 0, \quad (1.66)$$

se denomina **Cópula de Valor Extremo** o **EVC** (*Extreme Value Copula*).

Además, si F pertenece al dominio de atracción maximal, entonces la función \mathcal{C} es la EVC límite de F .

Una de las familias que satisface la definición de EVC es la cópula de Gumbel-Hougaard, dada por

$$\mathcal{C}_\theta(u)^k = \exp \left\{ - \left[(-k \ln u_1)^\theta, \dots, (-k \ln u_d)^\theta \right]^{1/\theta} \right\} = C_\theta^k(u). \quad (1.67)$$

Esta es la única familia arquimediana de EVC (Salvadori et al., 2007).

NOTA: La expresión (1.66) destaca una característica fundamental de las EVC, más conocida como propiedad de max-estabilidad (véanse más detalles en Salvadori et al., 2007).

1.6.4. Medidas de riesgo

Según Kriele y Wolf (2014), se conoce como riesgo a la posibilidad de que se produzcan *acontecimientos desfavorables*. Cuando se intenta captar el *riesgo* de forma cuantitativa resulta ser un fenómeno de múltiples aspectos.

Matemáticamente consiste en identificar las fluctuaciones, es decir, se tienen en cuenta tanto las variaciones *favorables* como las *desfavorables*. Entonces, sea (Ω, \mathcal{A}, P) un espacio de probabilidad con una σ -álgebra \mathcal{A} y una medida de probabilidad P . Se denota como $\mathcal{M}_{\mathcal{B}}(\Omega, \mathbb{R}^2)$ el espacio de variables aleatorias valoradas en \mathbb{R}^k

$$Z : \Omega \rightarrow \mathbb{R}^k, \omega \mapsto Z(\omega), \quad (1.68)$$

es decir, aspectos medibles con respecto a \mathcal{A} y a la σ -álgebra de Borel.

Definición. Una medida de riesgo es una aplicación

$$\rho : \mathcal{M}(\Omega, \mathbb{R}) \rightarrow \mathbb{R}, Z \mapsto \rho(Z), \quad (1.69)$$

donde $\mathcal{M}(\Omega, \mathbb{R}) \subseteq \mathcal{M}_{\mathcal{B}}(\Omega, \mathbb{R})$ es un subespacio vectorial adecuado (que depende de ρ).

Existen múltiples formas de medir el riesgo (véanse detalles en Kriele y Wolf, 2014); sin embargo, se hace referencia a las más comunes:

Valor en riesgo (VaR)

El valor en riesgo o VaR (Value-at-Risk) es el α -ésimo percentil de la distribución F ,

$$VaR_\alpha(Z) = \inf \{ z \in \mathbb{R} : F_Z(z) \geq \alpha \}, \quad (1.70)$$

donde F_Z es la función de distribución de Z . VaR_α es la mínima pérdida que se produce en el $100(1 - \alpha)\%$ de los peores escenarios.

En el lenguaje de la Estadística, el valor en riesgo representa el cuantil α inferior de la distribución de Z . En el caso especial de que F_Z sea invertible, se tiene que

$$VaR_\alpha(Z) = F_Z^{-1}(\alpha), \quad (1.71)$$

donde $1 > \alpha \geq 0$ y $F_Z^{-1}(\alpha)$ es la inversa de F_Z . Al igual que F , se trata de cantidades teóricas que no se conocen (véanse más detalles en Kriele y Wolf, 2014).

Riesgo de cola

El valor de riesgo de cola, a diferencia del valor en riesgo, también pondera las pérdidas más elevadas.

Definición. El valor del riesgo de cola viene dado por la esperanza condicionada

$$TVaR_\alpha(Z) = E[Z|Z > VaR_\alpha(Z)]. \quad (1.72)$$

Proporciona información sobre la pérdida esperada del $100(1 - \alpha)\%$ de los peores escenarios en que se supera el VaR_α . El valor del riesgo de cola para el mismo nivel de confianza α es siempre mayor que VaR, y en el caso extremo es igual a VaR.

Para variables con distribución continuas Z_1, Z_2 también tiene la propiedad de subaditividad:

$$TVaR_\alpha(Z_1 + Z_2) \leq TVaR_\alpha(Z_1) + TVaR_\alpha(Z_2), \quad (1.73)$$

que expresa intuitivamente que el riesgo en un colectivo diversificado es menor o igual que la suma de los riesgos individuales. Esta propiedad no se cumple en general para variables aleatorias con funciones de distribución que presentan discontinuidades (saltos).

Déficit esperado (ES)

El déficit esperado o ES (Expected Shortfall) está estrechamente relacionado con el riesgo de cola, y muestra subaditividad para todas las variables aleatorias.

Definición. El déficit esperado, también conocido como valor en riesgo promedio (Average Value-at-Risk), viene dado por

$$ES_\alpha(Z) = \frac{1}{1 - \alpha} \int_\alpha^1 VaR_y(Z) dy. \quad (1.74)$$

En el caso continuo, $TVaR_\alpha(Z) = ES_\alpha(Z)$.

Para el caso bivalente, sea la variable aleatoria $L = L_1 + L_2$ la pérdida total en una situación de riesgo, y sea F_L su función de distribución. Entonces se define el valor en riesgo análogamente a la expresión 1.70. (Véanse más detalles sobre medidas de riesgo en [Kriele y Wolf, 2014](#)).

2. Modelos espacio-temporales auto-regresivos para datos de conteo (SPINGARCH)

En este capítulo se muestra el contenido del artículo publicado en **Entropy Journal** que se puede encontrar en <https://www.mdpi.com/1099-4300/24/7/892> (Escudero et al., 2022).

2.1. Introducción

El modelizado de series temporales de recuentos ha recibido una atención considerable y creciente desde los años 50 (Cox, 1955; Bartlett, 1963; Hawkes, 1971b,a; Hawkes y Oakes, 1974) y en las últimas décadas (véase Kedemo y Fokianos, 2002; Jung y Tremayne, 2011; Davis et al., 2021; Xu y Zhu, 2022; Weiß et al., 2022). Se sabe que algunas distribuciones discretas bien conocidas, como Poisson y la binomial negativa (NB), sólo pueden tratar la sobredispersión; sin embargo, las distribuciones Poisson generalizada (GP) y Poisson doble (DP) pueden tratar tanto la sobredispersión como la subdispersión. Estos dos últimos modelos presentan algunas deficiencias o limitaciones. Alternativamente, la clase de modelos basados en la observación denominados modelos Integer Generalized Auto-Regressive Conditionally Heteroscedastic (INGARCH), propuestos por Heinen (2003) y Ferland et al. (2006), muestran flexibilidad para modelizar una amplia gama de casos de sobredispersión y subdispersión, al mismo tiempo que poseen propiedades que los hacen metodológicamente atractivos y útiles en la práctica.

Aunque un modelo clásico de Poisson INGARCH parece proporcionar un marco adecuado para modelizar datos de series temporales de recuentos y se ha aplicado a varios campos, Fokianos y Tjøstheim (2011) señalan que no puede emplearse para modelizar correlación negativa entre recuentos, y puede

incluir exclusivamente covariables que resulten en un término de regresión positivo, ya que, de lo contrario, la media del proceso de Poisson se vuelve negativa. Además, la media condicionada es igual a la varianza condicionada, y esta restricción puede conducir a un rendimiento deficiente de un modelo INGARCH de Poisson en la existencia de posibles observaciones extremas. Para superar estos inconvenientes, se han propuesto dos modelos INGARCH para modelizar la sobredispersión o la subdispersión en el mismo marco. Se trata del modelo DP (Heinen, 2003) y el modelo GP (Zhu, 2012a) (véase también Zhu (2012b) para una propuesta de una distribución Conway-Maxwell (COM) Poisson INGARCH).

En la literatura de estadística espacial, Besag (1974) realiza un primer intento de estructurar las relaciones espaciales para los datos de recuentos, donde la distribución del modelo de datos se especifica condicionalmente dada una vecindad espacial fija. Sin embargo, esto conduce a un modelo estadístico que solo permite una correlación negativa. Augustin et al. (2006) y Kaiser y Cressie (1996) demuestran cómo se podrían hacer modificaciones al modelo estadístico que permitían tanto la correlación negativa como la positiva. La suposición importante en estas clases de modelos es que la distribución de recuentos observados puede especificarse condicionalmente a partir de recuentos observados en vecinos espaciales, la cual es una suposición de Markov en el espacio.

Los datos de recuentos en la literatura estadística espacial se han abordado predominantemente a través de la estructura en un proceso latente, mientras que en series de tiempo han evolucionado de forma diferente. Por ejemplo, el modelo INGARCH de Ferland et al. (2006) y Heinen (2003) es un modelo de series de tiempo para recuentos bajo una distribución de Poisson con la esperanza como función de los dos anteriores. Ferland et al. (2006) demuestran cómo INGARCH(1,1) es análogo a ARMA(1,1) para recuentos. La distribución estacionaria del proceso INGARCH(1,1) también es equivalente a un proceso estocástico dado en Hawkes (1971b), a menudo llamado proceso puntual auto-excitante.

En Reinhart (2018) se dan extensiones al proceso de Hawkes para incorporar la estructura espacio-temporal. Las consideraciones de ramificación en Clark y Dixon (2018) o casi repetición de eventos (Andresen y Malleson, 2015) dan lugar a procesos puntuales espacio-temporales auto-excitantes. Estos modelos son extensiones de los procesos temporales de Hawkes, donde la tasa de eventos depende de la historia del proceso (Reinhart, 2018). Estos procesos se han mostrado beneficiosos para modelizar la dinámica de terremotos, epidemias, incendios forestales, accidentes de tránsito, homicidios, violaciones, terrorismo o delitos, que es el problema motivador de este capítulo. Se pueden encontrar diversos trabajos en este último contexto, por ejemplo, véase Mohler et al. (2011), Hu et al. (2018), Clark y Dixon (2021), Andresen y Malleson (2015).

2.2. DESCRIPCIÓN DE LOS DATOS

Este capítulo está motivado por el análisis de los datos delictivos en la ciudad de Riobamba-Ecuador, proporcionados por tres organismos gubernamentales, con el objetivo de comprender su comportamiento e interacción con la sociedad y coadyuvar con información para las instituciones públicas a optimizar sus acciones. Algunos estudios exploratorios existentes (véanse [Cepa et al., 2018](#) y [Trejo et al., 2013](#)) muestran características de este fenómeno delictivo. Sin embargo, no van más allá en la propuesta de un marco de modelización espacio-temporal.

El desafío general es cómo modelizar adecuadamente la relación espacio-tiempo entre las observaciones, especialmente si se va a realizar cualquier análisis estadístico. Siguiendo la línea de razonamiento de [Clark y Dixon \(2021\)](#), se tiene en cuenta la variación espacial al considerar un modelo Spatial Integer Generalized Auto-Regressive Conditionally Heteroscedastic Correlation (SPINGARCH). Este modelo comparte las propiedades de INGARCH y permite la correlación espacial al agregar un proceso log-gaussiano latente espacialmente correlacionado ([Clark y Dixon, 2018](#)). En este marco de trabajo y en paralelo a [Clark y Dixon \(2021\)](#), se formula una ecuación en diferencias estocástica para la intensidad del proceso espacio-temporal dentro de una clase de modelos espacio-temporales auto-excitados espacialmente correlacionados, que capturan tanto la dependencia debida a la auto-excitación como la dependencia de un proceso espacial subyacente. De hecho, se considera algunas extensiones de [Clark y Dixon \(2021\)](#) para adaptar dicha metodología al contexto de datos de Riobamba en particular. Se observa que el modelo de [Clark y Dixon \(2021\)](#) considera una estructura de regresión lineal en las covariables que también son constantes en el tiempo. Entonces, se propone estructurar la dependencia espacio-temporal de recuentos de delitos de la ciudad de Riobamba-Ecuador mediante una combinación de covariables basadas en la distancia que varían de forma natural tanto en el espacio como en el tiempo. Así pues, se considera bases B-splines dentro de un modelo aditivo generalizado que le permite manejar la variación espacio-temporal y las dependencias no lineales. De hecho, este es otro aspecto que se diferencia con respecto al modelo de [Clark y Dixon \(2021\)](#). La estrategia B-splines también permite combinar covariables que solo varían en el espacio con otras (como las meteorológicas) que solo varían en el tiempo, y con aquellas basadas en distancias que varían tanto en el espacio como en el tiempo. En conjunto esta estrategia es más flexible y se adapta mejor al caso de los datos de Riobamba.

2.2. Descripción de los datos

La inseguridad ciudadana es uno de los principales problemas que afectan al desarrollo de la población en cualquier país. Riobamba es una ciudad ecuatoriana, cabecera del cantón Riobamba y capital de la provincia de Chimborazo (véase la Figura 2.1). Se encuentra en la región interandina,

rodeada de varios volcanes, como el Chimborazo, Tungurahua, Altar y Carihuairazo. Ubicada a 2754 m. sobre el nivel del mar, tiene un clima andino frío con una temperatura promedio de 12°C. Según el censo de 2010, esta ciudad tenía 234170 habitantes y un crecimiento poblacional del 1,06 % hasta 2014.

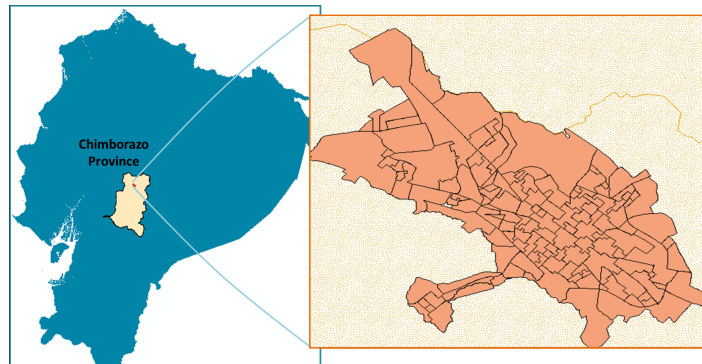
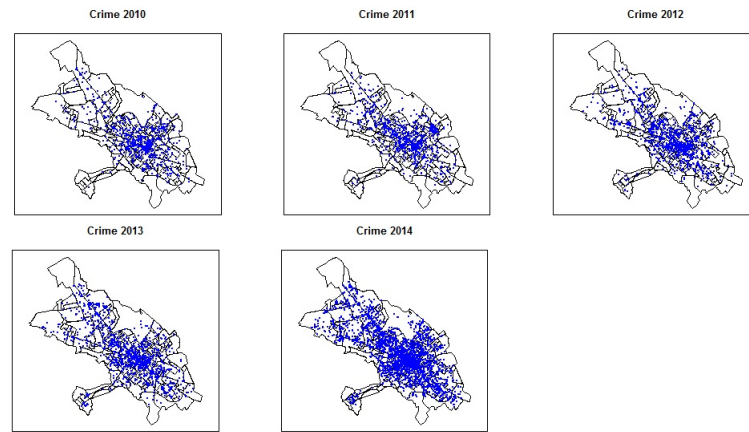


Figura 2.1: Ciudad de Riobamba-Ecuador

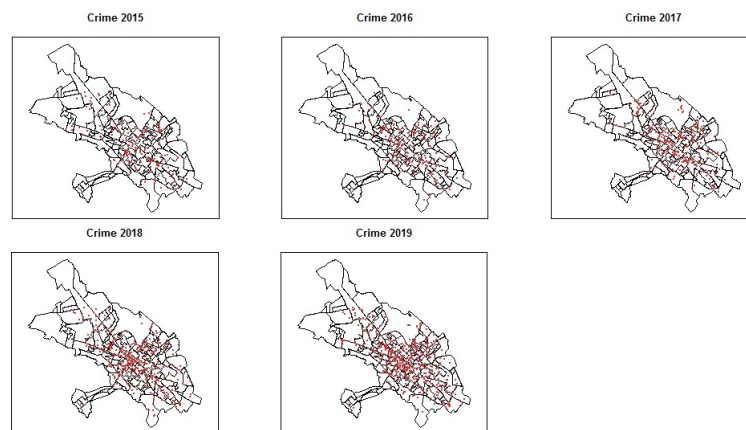
El comercio y turismo son características típicas de la ciudad, considerada como una central de negocios, empleo y educación por ser la tercera ciudad con mayor número de instituciones de educación a nivel superior. Sin embargo, uno de los principales problemas que acechan a esta ciudad son los actos delictivos como asaltos, robos en viviendas y locales comerciales, e intimidación, entre otros, que causan confusión, preocupación e importantes pérdidas a la población (Chávez et al., 2013). En la encuesta nacional ecuatoriana realizada en 2011, la provincia de Chimborazo ocupó el séptimo lugar con un 16,9 % de personas que han sido víctimas de algún delito, el 73,4 % de la población considera que la ciudad es insegura y el 38,0 % experimentó al menos un evento delictivo en su barrio. Según Castro et al. (2015), el gobierno ecuatoriano promovió un conjunto de nuevas políticas para reducir la delincuencia entre 2010 y 2014. Estas políticas involucran a la sociedad civil organizada y a las entidades competentes. A finales de 2014 la tasa de victimización, los homicidios y los robos disminuyen, pero aumenta el número de denuncias de la sociedad, como señal de mayor confianza en las instituciones competentes.

Se utilizan los datos de tres organismos gubernamentales cuya misión y visión es garantizar la seguridad ciudadana y la convivencia social (Unidades Policiales Comunitarias (UPC), Consejo de la Judicatura de Chimborazo (CJCH) y Ministerio del Interior (MI)). La Figura 2.2a muestra los actos delictivos denunciados desde MI para 2010–2014. La Figura 2.2b representa los actos delictivos flagrantes registrados por el CJCH para el periodo 2015–2019, es decir, los delitos cometidos con la detención del agresor en 24 horas, y finalmente, la Figura 2.2c muestra los delitos registrados por la

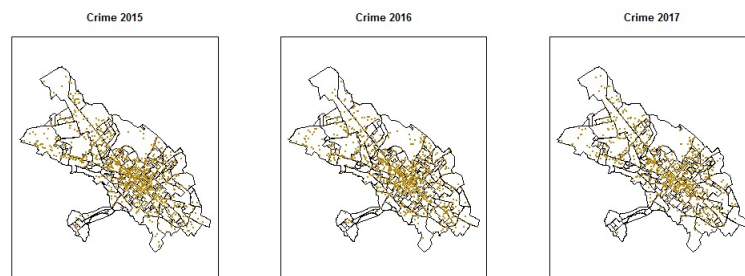
2.2. DESCRIPCIÓN DE LOS DATOS



(a) Delitos registrados por MI (2010-2014)



(b) Delitos Flagrantes registrados por CJCH (2015-2019)



(c) Delitos registrados por UPC (2015-2017)

Figura 2.2: Delitos registrados por tres entidades gubernamentales en la ciudad de Riobamba

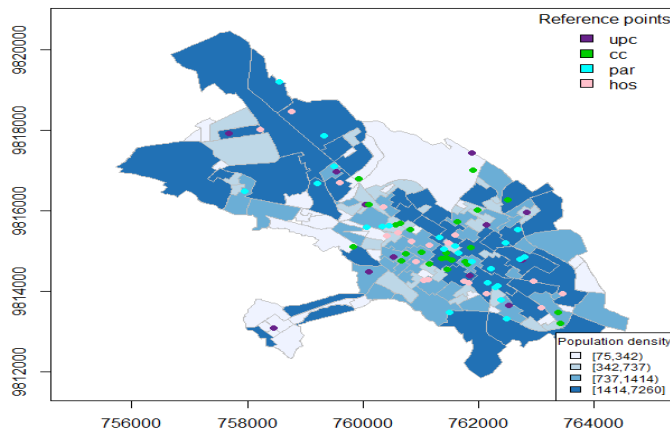


Figura 2.4: Densidad poblacional (azul) y ubicaciones de los puntos de referencia

UPC para el periodo 2015–2017. La información fue proporcionada bajo un contrato de confidencialidad y no está disponible directamente en ningún sitio web; sin embargo, se pueden consultar estadísticas de resumen de delitos subidos por el Ministerio del Interior a partir de 2015 en <http://cifras.ministeriodegobierno.gob.ec/comisioncifras/inicio.php>.

El registro ideal de la información dicta que el MI guarda todos los informes de las otras dos instituciones, como se muestra en la Figura 2.3. Sin embargo, esto dista mucho de ser cierto y al analizar los tres conjuntos de datos se comprobará esta anomalía.



Figura 2.3: Estructura jerárquica para el registro de delitos en la ciudad de Riobamba

La ciudad de Riobamba está dividida en 141 sectores o barrios, como se muestra en la Figura 2.4. Adicionalmente se tienen las ubicaciones geográficas en espacio y tiempo de los delitos. El Servicio Integrado de Seguridad 911 (ECU911) proporcionó como puntos de referencia las ubicaciones de las unidades policiales comunitarias (UPC) y de las cámaras de vigilancia

2.2. DESCRIPCIÓN DE LOS DATOS

instaladas en lugares estratégicos en toda la ciudad (cam) (ECU911, 2020). Se consideran algunos puntos de referencia importantes de la ciudad que describen aquellas áreas con un mayor tráfico vehicular y peatonal como: ubicaciones de parques (par), mercados (cc) (plazas, centros comerciales, supermercados) y hospitales (hos). Como covariable solamente espacial se tiene en cuenta la densidad poblacional estimada a nivel de sector o barrio. En cuanto a las variables temporales se consideran los promedios mensuales de temperatura y precipitación en la ciudad de Riobamba (véase la Figura 2.5), pueden ser descargadas de la página web del GEAA (<http://ceaa.esepoch.edu.ec:8080/redEma/>), dado que existen estudios previos (Andresen y Malleson, 2015) que relacionan los delitos de hurto con este tipo de covariables.

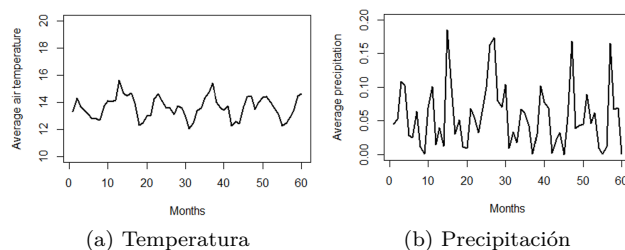


Figura 2.5: Promedios mensuales de temperatura y precipitación en la ciudad de Riobamba

Un primer análisis exploratorio por meses muestra que la mayor cantidad de delitos registrados por MI se centra en enero, junio y octubre, los delitos flagrantes son más elevados en febrero, septiembre y octubre, y los delitos registrados por la policía se incrementan en enero, abril y mayo (véase la Figura 2.6a). Esto es una muestra de comportamiento distintivo entre los tres conjuntos de datos de delitos utilizados, y de que el mes juega un papel importante.

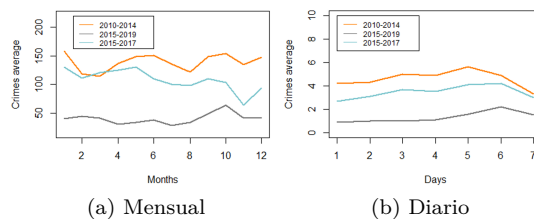


Figura 2.6: Descripción media de eventos por meses y días, para los delitos registrados por MI (2010-2014), CJCH (2015-2019) y UPC (2015-2017)

Al observar los datos por días de la semana (véase la Figura 2.6b), se visualiza mayor número de delitos los viernes y sábados, y de acuerdo a su ubicación espacial (véase la Figura 2.2), existe un alto nivel de casos delictivos en el centro de la ciudad.

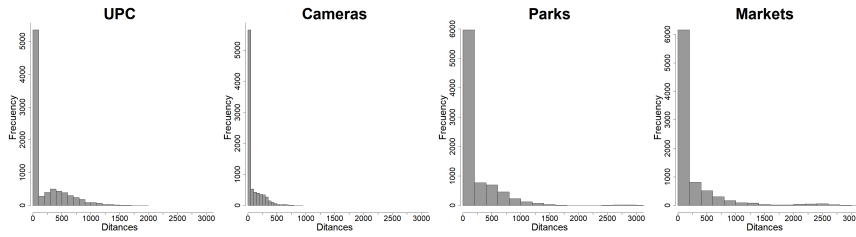


Figura 2.7: Distribución de las covariables

Se tiene en cuenta las distancias al vecino más cercano entre cualquier evento delictivo y la ubicación del punto de referencia correspondiente. Estas distancias informan sobre el vínculo entre un delito, en particular en los puntos de referencia, así como también si los puntos de referencia actúan como atractores o repulsores de los delitos. La distribución de estas distancias se muestra en la Figura 2.7, observándose que las distancias pequeñas entre eventos delictivos y puntos de referencia son mucho más frecuentes que las más grandes, lo que indica que estos puntos de referencia podrían ser fuentes o atractores de delitos.

2.3. Metodología

2.3.1. Modelo

El enfoque metodológico general de esta sección es estructurar la dependencia espacio-temporal de recuentos a través de una combinación de dependencia espacial en un modelo de proceso latente y dependencia temporal en un modelo de datos, con factores exógenos que varían en el espacio y/o el tiempo. Siguiendo a [Clark y Dixon \(2021\)](#) se considera una ecuación en diferencias estocástica para la intensidad del proceso espacio-temporal dentro de una clase de modelos espacio-temporales auto-excitados espacialmente correlacionados para datos de recuento, estos modelos capturan la dependencia del modelo de datos como la de un proceso espacial latente. En particular, el estudio se centra en los modelos SPINGARCH(1,1), que en general, permite al modelizador definir la autocorrelación presente en los datos y la relación media-varianza con mayor flexibilidad.

2.3. METODOLOGÍA

Se denota por $\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n$ un vector de ubicaciones espaciales que permanecen fijas en el tiempo, y t el período de tiempo discreto correspondiente. $N|\mathbf{s}_i|$ indica la vecindad espacial de la ubicación \mathbf{s}_i , $Y(\mathbf{s}_i, t)$ el proceso observado en la ubicación espacial \mathbf{s}_i y tiempo t , y $X(\mathbf{s}_i, t)$ es el proceso latente no observado.

Se utiliza una distribución de Poisson condicional y se establece una estructura espacio-temporal en la covarianza del proceso log-gaussiano latente. El modelo de datos, $Y(\mathbf{s}_i, t)$, se puede definir condicionalmente dado un modelo de proceso $X(\mathbf{s}_i, t)$, que es una suposición crítica en esta técnica. Como resultado, el modelo de proceso es una función tanto de covariables espaciales o temporales observables como de errores espaciales latentes no observables.

En este caso particular, se considera que la intensidad espacio-temporal $\lambda(\mathbf{s}, t)$ proporciona el modelo de proceso, y el modelo completo es una ecuación en diferencias estocástica que opera directamente en la función de intensidad. Por tanto, los recuentos de delitos en el espacio y el tiempo, $Y(\mathbf{s}_i, t)$, son variables aleatorias de Poisson distribuidas condicionalmente para $i = 1, \dots, n$, es decir, $Y(\mathbf{s}_i, t) | \lambda(\mathbf{s}_i, t) \sim \text{Poisson}(\lambda(\mathbf{s}_i, t))$, donde $\lambda(\mathbf{s}_i, t)$ representa la tasa en \mathbf{s}_i en el tiempo t ; por lo tanto, $E[Y(\mathbf{s}_i, t) | \lambda(\mathbf{s}_i, t)] = \lambda(\mathbf{s}_i, t)$.

Bajo el supuesto de que un cambio en la tasa de delitos en una ubicación y período específicos es una función de características geográficas particulares de la ubicación dada por $\alpha_t = (\alpha_{\mathbf{s}_1}, \alpha_{\mathbf{s}_2}, \dots, \alpha_{\mathbf{s}_n})^T$ (también conocido como base de referencia, simplemente una función de factores exógenos, potencialmente variables), junto con otros dos factores, un deterioro natural χ , y victimización repetida η , se tiene un modelo de la forma

$$\frac{(\lambda(\mathbf{s}_i, t + \Delta t) - \lambda(\mathbf{s}_i, t))}{\Delta t} = \exp(\alpha_{\mathbf{s}_i}) - \kappa\lambda(\mathbf{s}_i, t) + \eta Y(\mathbf{s}_i, t), \quad (2.1)$$

para cada $i = 1, \dots, n$, donde $\kappa = (1 - \chi)$ representa estrés en ausencia de victimización repetida, y η captura el cambio esperado debido a acciones repetidas o casi repetidas (Hu et al., 2018; Clark y Dixon, 2018). Con $\Delta t = 1$ en la expresión (2.1), este modelo se convierte en un modelo SPINGARCH(1,1), con $Y(\mathbf{s}_i, t)$ definido condicionalmente en la intensidad $\lambda(\mathbf{s}, t)$ que se puede modelizar utilizando covariables espaciales y temporales observables $\alpha_{\mathbf{s}_i}$, así como errores latentes no observables ϵ_t . De este modo, el modelo final propuesto se define a través de la siguiente estructura jerárquica:

$$\lambda_t = \exp(X_t + \epsilon_t) + \eta Y_{t-1} + \kappa \lambda_{t-1}, \quad (2.2)$$

con

$$X_t \sim \text{Gau} \left(\alpha_t, (I_{n,n} - \zeta C)^{-1} \sigma^2 \right)$$

$$\epsilon_t \sim \text{Gau} \left(0, (I \sigma_\epsilon^2) \right),$$

donde $\lambda_t = (\lambda(\mathbf{s}_1, t), \lambda(\mathbf{s}_2, t), \dots, \lambda(\mathbf{s}_n, t))^T$ es una cadena de Markov en $(\mathbb{R}^+)^n$ y la misma notación se aplica para Y_t y X_t . $I_{n,n}$ es la matriz de identidad, σ^2 es la varianza condicional, y ζ controla la cantidad de dependencia espacial en el modelo no capturado por las covariables en α_t . La estructura espacial a gran escala se explica en el proceso latente X_t mediante el parámetro de regresión espacial α_t , mientras que la estructura espacial a pequeña escala se explica mediante la especificación condicionada de X_t . Para este último se usa un modelo Auto-Regresivo Condicionado (CAR) (a través de vecinos espacialmente adyacentes),

$$X(\mathbf{s}_i, t) | X(\mathbf{s}_j, t), \mathbf{s}_j \in N|\mathbf{s}_i| \sim N(\mu(\mathbf{s}_i, t), \sigma^2) \quad (2.3)$$

con

$$\mu(\mathbf{s}_i, t) = \alpha(\mathbf{s}_i, t) + \zeta \sum_{\mathbf{s}_j \in N|\mathbf{s}_i|} X(\mathbf{s}_j, t) - \alpha(\mathbf{s}_j, t).$$

Aquí C es una matriz con entradas $(i, j) = 1$ si las ubicaciones \mathbf{s}_i y \mathbf{s}_j son vecinas. La variación en el proceso espacio-temporal se obtiene al agregar adicionalmente ruido espacio-temporal, dejando $\epsilon(\mathbf{s}_i, t) \stackrel{iid}{\sim} N(0, \sigma_\epsilon)$.

2.3.2. Inferencia bayesiana de parámetros

El modelo jerárquico definido en la Sección 2.3.1 depende de un conjunto de parámetros en el nivel final de la jerarquía, dado por $\theta = (\eta, \kappa, \alpha, \zeta, \sigma, \sigma_\epsilon)$, de manera similar a un modelo Besag-York-Mollié (BYM) clásico (Morris et al., 2019; Clark y Dixon, 2018; Boulieri et al., 2017) que define un modelo espacial totalmente bayesiano (véase Thamrin y Alimun, 2019).

Siguiendo a Clark y Dixon (2021), el marco inferencial bayesiano consiste en actualizar los parámetros θ según los datos disponibles a través de una densidad a priori $\pi(\theta)$ y una densidad o probabilidad condicionada $\pi(data|\theta)$ para obtener $\pi(\theta|data)$, una densidad a posteriori de θ dados los datos. La distribución conjunta a priori de los parámetros en el modelo se puede expresar como $\pi(\theta) = \pi(\eta|\kappa)\pi(\kappa)\pi(\alpha)\pi(\sigma)\pi(\sigma_\epsilon)\pi(\zeta)$, donde se asume independencia entre los puntos excepto por η y κ debido al condicionamiento $\eta + \kappa < 1$. Haciendo $U(\mathbf{s}_i, t) = Y(\mathbf{s}_i, t) + \epsilon(\mathbf{s}_i, t)$, la distribución condicionada completa para modelizar θ viene dada por

$$\pi(\theta|Y, U) \propto \prod_{t=1}^T \pi(Y_t|\lambda_t)\pi(\lambda_t|\lambda_{t-1}, Y_{t-1}, \theta, U_t)\pi(U_t, \theta)\pi(\lambda_0|\theta)\pi(Y_0|\lambda_0)\pi(\theta), \quad (2.4)$$

y por U se tiene

$$\pi(U|Y, \theta) \propto \prod_{t=1}^T \pi(Y_t|\lambda_t)\pi(\lambda_t|\lambda_{t-1}, Y_{t-1}, \theta, U_t)\pi(U_t|\theta)\pi(\lambda_0|\theta)\pi(Y_0|\lambda_0). \quad (2.5)$$

2.3. METODOLOGÍA

Para cualquier inferencia sobre los parámetros, MCMC (Markov Chain Monte Carlo) debe tomar muestras de la densidad de estado latente total U , lo que requiere una evaluación de

$$\log(U|\alpha_t, \sigma, \sigma_\epsilon, \zeta) \propto \frac{-T \times n}{2} \log(2\pi) + \frac{1}{2} \log |\Sigma_f^{-1}(\theta)| - \frac{1}{2} (X_t - \alpha_t)^T \Sigma_f^{-1}(\theta) (X_t - \alpha_t). \quad (2.6)$$

Dado que la estructura de vecindad es constante para todos los períodos, se puede interpretar $\Sigma_f(\theta)$ como la matriz de covarianza espacio-temporal completa $(I_{T,T} \otimes I_{n,n} - C)^{-1} \sigma^2 + I_{n \times T} \sigma_\epsilon^2$.

La dispersión de la estructura de covarianza significa que los únicos cálculos de $\frac{1}{2} (X_t - \alpha_t)^T \Sigma_f^{-1}(\theta) (X_t - \alpha_t)$ que deben ocurrir son para vecinos espaciales. Por lo tanto, el desafío es el cálculo del determinante en $\log |\Sigma_f^{-1}(\theta)|$. La estructura específica de $\Sigma^{-1}(\theta) = (I - C)(1/\sigma^2)$ hace que $\log |\Sigma^{-1}(\theta)| = \frac{n}{2 \log \sigma^2} + \log |I_{n,n} - \zeta N|$, donde N es la matriz de vecindad o adyacencia. Esto a su vez se puede reescribir como $\log |\Sigma_f^{-1}(\theta)| = \frac{n \times T}{\log \sigma^2} + T \log |\Sigma^{-1}(\theta)|$, resultando finalmente en

$$\log |\Sigma_f^{-1}(\theta)| = T \log |\Sigma^{-1}(\theta)| \propto \frac{n \times T}{\log \sigma^2} + T \sum_{j=1}^n (1 - \zeta \chi_j), \quad (2.7)$$

con χ_j los valores propios de la matriz de vecindad que dependen únicamente de su estructura y no de los parámetros.

2.3.3. Estimación de la función de factores exógenos

La acción de las covariables deterministas que dependen del espacio o del espacio-tiempo es altamente no lineal sobre las respuestas. Por lo tanto, se utiliza un modelo aditivo generalizado (GAM) que admite la estimación de suavidad integrada abordando la falta de linealidad (Gasparrini et al., 2017; Taylan et al., 2010; Goetz et al., 2011). GAM da como resultado un método analítico más eficiente que los modelos lineales más clásicos (Ravindra et al., 2019). La relación entre cada predictor x_i y la media de la variable respuesta $g(u)$ es indirecta, porque se calcula utilizando la función $f(x_i)$ (generalmente funciones splines que tienen bases polinómicas; Taylan et al., 2010), dada de la forma

$$g(u) = \beta_0 + \sum_{i=1}^p f_i(x_i). \quad (2.8)$$

También se puede tener versiones multivariantes por instantes temporales. Por ejemplo,

$$g_t(u) = \beta_0 + f_t(x_1, x_2), \quad (2.9)$$

con f_t una superficie espacial suavizada en el tiempo t -ésimo. Esta superficie suavizada para cada t se puede escribir como $f_t(x_1, x_2) = \mathbf{B}_s \psi^{(t)}$, donde $\mathbf{B}_s = \mathbf{B}_1 \otimes \mathbf{B}_2$ es una base B-spline k -dimensional de dimensión $I \times k_1 k_2$ derivada del producto de Kronecker por fila de los B-splines de base marginal para \mathbf{B}_1 , \mathbf{B}_2 y $\psi^{(t)} = (\psi_1^{(t)}, \dots, \psi_{k_1 k_2}^{(t)})^T$. Se debe tener en cuenta que k_1 y k_2 son los números de columnas de las bases marginales \mathbf{B}_1 y \mathbf{B}_2 , respectivamente, y dependen del número de nodos y grado de polinomios utilizados para generar estas bases (Vicente et al., 2021). Aquí, se utiliza el criterio de validación cruzada generalizado (GCV) para estimar los parámetros que proporcionan el grado de suavizado.

2.4. Resultados

La ciudad de Riobamba está dividida en $m = 141$ barrios (véase la Figura 2.4), cuyos centroides se denotan por $\{\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_{141}\}$. Se halla la matriz de adyacencia o vecindad necesaria para determinar el proceso latente espacial; esta es una matriz dispersa que permite optimizar los costos computacionales (Solarte et al., 2013).

La unidad temporal en este trabajo es el mes del año, por lo que se considera el número de delitos por barrio en cada mes. Como los tres conjuntos de datos tienen períodos de tiempo diferentes, se establecen los instantes de tiempo para MI y CJCH como $t \in \{1, 2, \dots, 60\}$, y para UPC $t \in \{1, 2, \dots, 35\}$.

Como ya se dijo en la Sección 2.2, se consideran las distancias del vecino más cercano de cada delito a las unidades de policía comunitaria (UPC), a las cámaras de vigilancia (cam), a los mercados (cc), a los parques (par) y a los hospitales (hos). Estas distancias son tomadas en promedio por sector, proporcionando matrices de dimensión $m \times n$. La densidad de población (pob) se incorpora al modelo como una covariable espacial de dimensión $m \times 1$. Inicialmente se consideran dos variables meteorológicas (véase la Figura 2.5), con temperaturas promedio mensuales que oscilan entre 12 y 15 °C y precipitaciones entre 0.00 y 0.15 mm. Así, aunque son consideradas en otros estudios, en la región particular no son influyentes en promedios mensuales en los eventos delictivos.

Se prueban todas las combinaciones posibles de un GAM multivariante, y se encuentra que los GAM univariantes proporcionan los mejores resultados. Por lo tanto, se utiliza un modelo aditivo generalizado univariante con B-splines cúbicos (denotado por $\hat{f}_i^{[3]}$) que permite la incorporación de relaciones no lineales entre cada covariable y la variable de respuesta. El modelo GAM completo tomando las covariables deterministas o factores exógenos

2.4. RESULTADOS

propuestos inicialmente es el siguiente:

$$\alpha_t = \hat{\beta}_0 + \hat{f}_1^{[3]}(UPC) + \hat{f}_2^{[3]}(cam) + \hat{f}_3^{[3]}(cc) + \hat{f}_4^{[3]}(par) + \hat{f}_5^{[3]}(hos) + \hat{f}_6^{[3]}(pob). \quad (2.10)$$

Finalmente se muestra el modelo representativo con cada conjunto de datos:

$$\alpha_t^{MI} = \hat{\beta}_0 + \hat{f}_2^{[3]}(cam) + \hat{f}_4^{[3]}(par),$$

$$\alpha_t^{CJCH} = \hat{\beta}_0 + \hat{f}_3^{[3]}(cc) + \hat{f}_6^{[3]}(pob),$$

y

$$\alpha_t^{UPC} = \hat{\beta}_0 + \hat{f}_1^{[3]}(UPC) + \hat{f}_2^{[3]}(cam).$$

La Figura 2.8 representa el modelo ajustado correspondiente con B-splines para cada conjunto de datos. Se observa cómo el modelo propuesto alcanza a caracterizar las variaciones de datos reales.

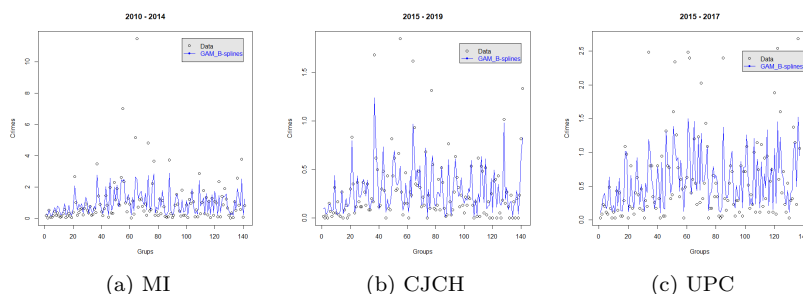


Figura 2.8: B-splines cúbicos de las covariables exógenas para los delitos de los tres grupos de datos

Una vez estimado el parámetro α_t y manteniendo fijo $\zeta = 0.99$, los parámetros restantes $\theta = (\eta, \kappa, \sigma, \sigma_\epsilon)$ se estiman utilizando un marco bayesiano como se explica en la Sección 2.3. Se usan distribuciones Beta informativas para η y κ , y Cauchy para σ y σ_ϵ .

Se ejecutan tres cadenas de Markov de 70000 iteraciones cada una por parámetro y para cada uno de los tres conjuntos de datos. Las primeras 10000 iteraciones se descartan como período de preparación y se toman las muestras cada 100 iteraciones para eliminar cualquier posible autocorrelación. Las Figuras 2.9a, 2.10a y 2.11a muestran las cadenas MCMC para los cuatro parámetros y para los tres conjuntos de datos, visualizándose la convergencia y estabilidad de estas cadenas. Las distribuciones a posteriori de cada uno

de los parámetros se muestran en las Figuras 2.9b, 2.10b y 2.11b. También se muestra la autocorrelación de los parámetros simulados de la distribución a posteriori, confirmando la ausencia de autocorrelación (véanse las Figuras 2.9c, 2.10c y 2.11c). Las Tablas 2.1–2.3 muestran los estadísticos de resumen de las estimaciones y de diagnóstico para las distribuciones a posteriori.

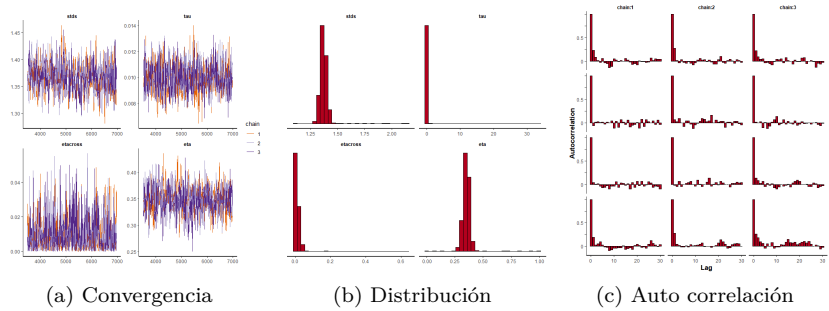


Figura 2.9: Inferencia bayesiana para los registros delictivos de MI (2010–2014), donde $\eta = eta$ $\kappa = etacross$, $\sigma = 1/tau$ y $\sigma_\epsilon = stds$

Tabla 2.1: Distribuciones a posteriori de los parámetros para los registros delictivos de MI (2010-2014)

Post-param	mean	sd	2.5 %	50.0 %	97.5 %	n_eff	Rhat	H
η	0.35	0.03	0.29	0.35	0.40	685	1.00	2.99
κ	0.01	0.01	0.00	0.01	0.04	920	1.00	3.01
$1/\sigma^2$	0.01	0.00	0.01	0.01	0.01	1028	1.00	2.69
σ_ϵ^2	1.37	0.03	1.31	0.037	1.42	618	1.01	2.95

Teniendo en cuenta que κ (coeficiente de deterioro) pondera el valor esperado (es decir, la intensidad) mientras que η (coeficiente de victimización) pondera los datos u observaciones en sí, se espera que η sea mayor para los delitos de MI ya que a partir de 2015 se incrementan las intervenciones policiales en respuesta a un mayor número de denuncias. Además, κ es mayor para la UPC, lo que indica que el modelo pondera más la intensidad esperada, dando más importancia a lo esperado que a los datos reales. Conviene recordar que el tamaño efectivo de la muestra (n_eff) y una medida de equilibrio en cadena (Rhat) son el número de extracciones independientes en la muestra y los estadísticos de diagnóstico sobre la convergencia de la cadena, respectivamente. Rhat iguales o cercanos a 1 son indicativos de convergencia (Vuong et al., 2020).

2.4. RESULTADOS

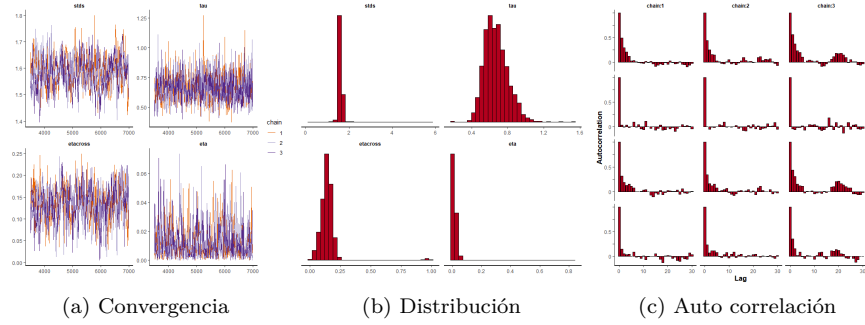


Figura 2.10: Inferencia bayesiana para los registros delictivos de CJCH (2015–2019), donde $\eta = eta$, $\kappa = etacross$, $\sigma = 1/tau$, $\sigma_\epsilon = stds$

Tabla 2.2: Distribución de parámetros para los delitos de CJCH (2015–2019)

Post-param	mean	sd	2.5 %	50.0 %	97.5 %	n_eff	Rhat	H
η	0.01	0.01	0.00	0.01	0.50	530	1.00	2.98
κ	0.13	0.04	0.04	0.14	0.21	342	1.01	2.86
$1/\sigma^2$	0.66	0.12	0.45	0.65	0.93	1057	1.00	3.07
σ_ϵ	1.59	0.06	1.47	1.59	1.70	293	1.01	2.68

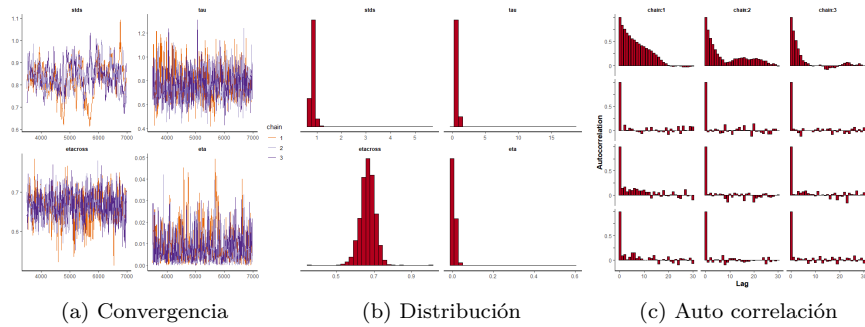


Figura 2.11: Inferencia bayesiana para los registros delictivos de UPC (2015–2017), donde $\eta = eta$, $\kappa = etacross$, $\sigma = 1/tau$, $\sigma_\epsilon = stds$

Para completar la información, también se halla la entropía de Shannon para cada uno de los parámetros asociados a cada uno de los tres conjuntos de datos. Aprovechando la salida del MCMC para cada parámetro, para el que se tiene una muestra a posteriori de tamaño 1050, se discretiza su longitud

de rango en un número de bins igual al valor entero más cercano a la raíz cuadrada del tamaño de la muestra, y se calcula la entropía H a partir de estos intervalos.

Tabla 2.3: Distribución a posteriori de los parámetros para delitos de UPC (2015–2017)

Post-param	mean	sd	2.5 %	50.0 %	97.5 %	n_eff	Rhat	H
η	0.01	0.01	0.00	0.01	0.03	635	1.00	2.89
κ	0.67	0.04	0.59	0.67	0.74	505	1.00	2.96
$1/\sigma^2$	0.77	0.14	0.53	0.75	1.08	881	1.00	2.91
σ_ϵ	0.83	0.07	0.68	0.83	0.96	79	1.02	2.70

El valor de H se muestra en las Tablas 2.1–2.3, y refleja la incertidumbre asociada a cada parámetro. Una H pequeña indica una pequeña incertidumbre en la estimación del parámetro y, por tanto, una mayor confiabilidad en su valor. De hecho, los parámetros estimados con menor incertidumbre son η y $1/\sigma^2$ para el período 2015–2019, κ para 2010–2014, y $1/\sigma_\epsilon^2$ para 2015–2017.

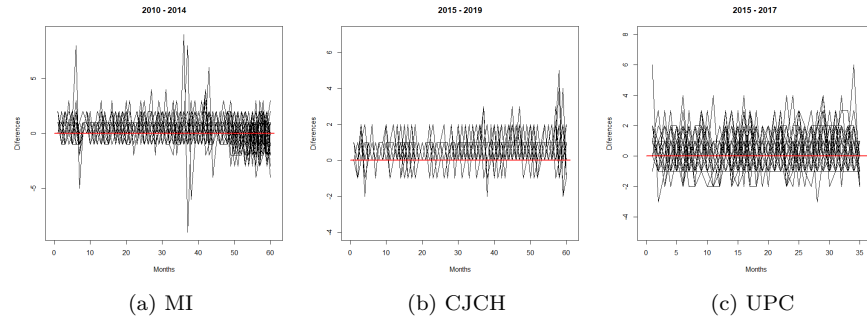


Figura 2.12: Diferencias espacio-temporales entre delitos reales y predicciones

Como herramienta de bondad de ajuste, se hallan los errores cuadráticos medios temporales de predicción (MSPE) (véase la Tabla 2.4), que informan de una medida de las diferencias entre los valores de las predicciones y los reales, observando que el SPINGARCH con B-splines cúbicos muestra los valores de MSPE más bajos. También se calculan las diferencias entre las predicciones y los valores reales en el desplazamiento espacio-temporal (véase la Figura 2.12), siendo los MSPE 0.45 (MI), 0.20 (CJCH), y 0.41 (UPC), manteniéndose pequeños en términos generales.

2.4. RESULTADOS

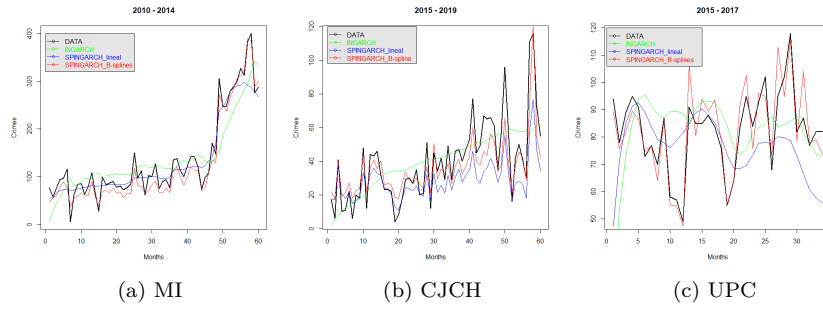


Figura 2.13: Delitos reales y predicciones

Tabla 2.4: Errores cuadráticos medios temporales (MSPE)

DATA	INGARCH	SPINGARCH_Lineal	SPINGARCH_B-splines
MI	1990.97	891.33	364.66
CJCH	351.81	301.39	122.93
UPC	508.88	373.66	56.24

Como ilustración final, se compara la predicción temporal de SPINGARCH(1,1) con las covariables modelizadas a través de B-splines, SPINGARCH(1,1) con covariables estáticas en el tiempo y modelizadas linealmente, y modelo INGARCH(1,1). Los resultados comparativos se muestran en la Figura 2.13, y se observa que SPINGARCH(1,1) con covariables suavizadas con B-splines proporciona los mejores predicciones, ya que se acercan más a los datos de delitos reales.

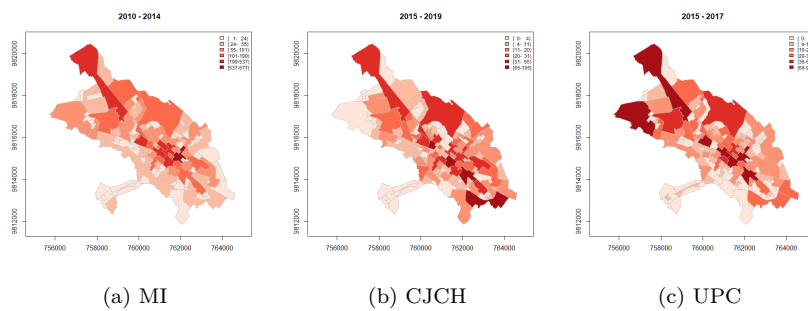


Figura 2.14: Predicciones espaciales de delitos para los tres grupos de datos

Para este modelo, también se informa de las predicciones espaciales para los tres conjuntos de datos sobre delitos en la ciudad de Riobamba, lo que ilustra

que el modelo espacio-temporal es lo suficientemente flexible para proporcionar predicciones temporales y también predicciones espaciales.

2.5. Discusión y conclusiones

Este capítulo formula un modelo estadístico que contiene tanto dependencia espacial latente como dependencia temporal en forma de una ecuación en diferencias estocástica para la intensidad espacio-temporal. Este modelo es consistente con las creencias comunes sobre cómo la violencia y el crimen evolucionan en el espacio y el tiempo. De hecho, el modelo propuesto es un modelo espacio-temporal auto-excitado correlacionado espacial y temporalmente, que captura tanto la dependencia en los datos como la dependencia de un proceso espacial latente en la línea que lo hacen los modelos INGARCH(1,1). Otro aspecto del modelo presentado es que el efecto de las covariables exógenas se añade mediante B-splines no lineales, lo que mejora los modelos anteriores que solo tienen formas lineales en las covariables y no varían en el tiempo.

Se sigue un marco inferencial bayesiano, ya que es flexible y puede manejar la estimación de un gran número de parámetros con estructuras complejas, como las consideradas en este apartado en el espacio y el tiempo. Se ha logrado estimar la estructura de vecindad en el espacio y el comportamiento de auto-regresión en el tiempo.

Al analizar los datos delictivos en la ciudad de Riobamba, mediante una extensa búsqueda e inspección preliminar se detectan las covariables más influyentes basadas en la distancia y cómo se incorporan al modelo de predicción. Destacamos algunas diferencias entre los tres tipos de conjuntos de datos. En el caso de los registros generales (conjunto de datos de 2010–2014), las distancias mínimas a las cámaras de vigilancia y a los parques son relevantes, porque a través de la vigilancia de estas cámaras se prevé o se da por hecho que suceda un evento delictivo, y en lugares, como los parques, hay una mayor protección policial sobre todo los fines de semana. En el caso de los delitos flagrantes (2015–2019), las covariables relevantes son las distancias a la variable mercados (plazas, centros comerciales y supermercados) y la densidad de población, ya que un mayor movimiento poblacional contribuye a la alerta y denuncia inmediata de los hechos delictivos. Por último, para los registros policiales (2015–2017), las distancias a cámaras y UPC tienen una influencia más representativa, puesto que la mayoría de las víctimas acuden a la policía en primera instancia solicitando ayuda, independientemente de que el hecho delictivo registrado sea denunciado legalmente o no. Los resultados de la estimación muestran mayor número de delitos en el sector 65, denominado *San Alfonso*, un posible motivo pudiera ser que en el periodo 2010-2014 el mayor mercado de la ciudad se encontraba en aquel sector. Sin embargo, para los registros de delitos flagrantes (2015–2019), las predicciones muestran

2.5. DISCUSIÓN Y CONCLUSIONES

mayor número de eventos en las áreas 37, 55, 76 y 141 que corresponde a los sectores de *La Dolorosa*, *La Merced*, *La Estación* y *Tubasec*, respectivamente, mientras que para los expedientes policiales (2015–2017) se destacan otros nueve sectores (véase la Figura 2.14, 2015–2017).

Estos resultados pueden ser utilizados por las entidades gubernamentales encargadas de la seguridad ciudadana para optimizar recursos mejorando la planificación, el despliegue de unidades policiales o el patrullaje y verificación aleatoria. Además, proporcionan información valiosa para las entidades gubernamentales encargadas de la seguridad ciudadana, para optimizar los recursos mejorando la planificación, el despliegue de las unidades policiales o el patrullaje y la verificación aleatoria. Los datos utilizados en este estudio, aunque vienen en forma de coordenadas espacio-temporales, tienen algunas limitaciones. Una de ellas es que los datos proporcionados por las entidades gubernamentales no tienen información detallada sobre el delito y las características de los hechos. Si tuvieran alguna información adicional, se podrían haber utilizado modelos más complejos con información de marcas. Otro aspecto es que el INEC (institución gubernamental que proporciona y pone a disposición los datos censales) tiene como unidad mínima de estudio las parroquias y no lleva estadísticas por distritos y/o por sectores o barrios de las ciudades. Esto obliga a desagregar de forma empírica los datos aportando burdas aproximaciones de población.

3. Modelos espacio-temporales de Cox log-gaussianos (LGCP)

3.1. Introducción

Los modelos de procesos de Cox log-gaussianos (LGCP) fueron propuestos por [Cox \(1955\)](#). Toman el nombre de procesos de Poisson doblemente estocásticos porque surgen como un proceso de Poisson no homogéneo con una medida de intensidad aleatoria log-gaussiana ([Diggle et al., 2013](#)). Esta medida de intensidad se suele especificar mediante una función conocida como proceso o superficie de intensidad Λ ([Brix y Diggle, 2001](#)). Tiene una estructura multiplicativa, lo que permite incorporar información de covarianza mediante componentes deterministas temporales y espaciales, y un término residual modelizado por un proceso de ruido ([Møller y Díaz-Avalos, 2010](#)). Una elección adecuada de la función de correlación espacio-temporal proporciona una estimación del proceso gaussiano subyacente confiable.

El desarrollo de herramientas computacionales forma parte importante de los avances en la modelización de los LGCP. La estimación de parámetros con métodos de Markov Chain Monte Carlo (MCMC) e Integrated Nested Laplace Approximation (INLA) proporciona un enfoque bastante preciso para ajustar modelos gaussianos latentes complejos, y lo más relevante es que se obtienen respuestas similares ([Serra et al., 2014](#)). Sin embargo, la literatura muestra con más frecuencia el uso de MCMC para hallar la distribución a posteriori del campo gaussiano no observado o de la función de intensidad aleatoria, donde el campo gaussiano se aproxima mediante una función escalonada obtenida por discretización.

La ventaja de los modelos discretizados es que pueden ser fácilmente explorados, utilizando el muestreo de Gibbs, pero no convergen a un resultado razonable, ya que los lados de los cuadrantes tienden a cero ([Rathbun y Cressie, 1994](#)). Sin embargo, [Moller et al. \(1998\)](#) muestran que los LGCP pueden ser simulados exactamente sin ningún problema con los efectos

de borde, y las a posteriori aproximadas convergen a las a posteriori exactas cuando el tamaño de las celdas para la discretización tiende a cero (Waagepetersen, 2004). También se incluyen la escala espacial, los efectos de borde, la pre-elección en los modelos LGCP (Illian et al., 2012), y una función de verosimilitud aproximada para llevar a cabo la inferencia espacio-temporal no estacionaria. El método utiliza modelos basados en la convolución para capturar la estructura de correlación espacio-temporal (Rodrigues y Diggle, 2012).

Debido a su potencial continúan surgiendo contribuciones. Por ejemplo, el establecimiento de funciones de covarianza separables y no separables, sobre una región espacial acotada y cruzada con tiempo circular, con el propósito de modelizar la periodicidad de la componente determinista temporal (Shirota y Gelfand, 2017). Con respecto a la estimación de parámetros de segundo orden, Siino et al. (2018) utilizan el método de contraste mínimo, específicamente para la clase de los LGCP espacio-temporales homogéneos. Por otro lado, Sørbye et al. (2019) realizan una combinación de campo estructurado espacialmente escalado con un campo aleatorio no estructurado. Para ello emplean un parámetro de precisión común y un hiperparámetro adicional, este último identifica la fracción de varianza explicada por el término estructurado espacialmente. El escalado adecuado hace que el análisis sea invariante a la resolución de la cuadrícula. Kuronen et al. (2022) muestran un LGCP jerárquico, en el que un conjunto de puntos afecta a otro, pero no a la inversa. Hesselund et al. (2020) exponen un método de validación cruzada para la selección de modelos y un algoritmo para la inferencia regularizada. Estos pueden utilizarse para obtener modelos dispersos para funciones de correlación por pares cruzados.

La principal herramienta computacional que usan Taylor et al. (2013) para los LGCP es MCMC, e incluyen la lectura y conversión de datos, la estimación de los componentes y parámetros clave del modelo, la especificación de las cantidades de salida y simulación, el cálculo de los valores esperados de Monte Carlo, el post-procesamiento y la simulación de conjuntos de datos. Utilizan un enfoque bayesiano para llevar a cabo la inferencia. En este contexto, se han desarrollado contribuciones sobre nuevos métodos computacionales de inferencia para mejorar la teoría general de convergencia (Simpson et al., 2016) y las aproximaciones discretas que superan una limitación inherente a los modelos espaciales basados en estructuras de Markov (Johnson et al., 2019).

Los avances computacionales han hecho de los LGCP modelos prácticos y adaptables a múltiples áreas, en las que el riesgo relativo se observa de forma incompleta a través de la lente de los datos de patrones puntuales o de recuentos agregados, por ejemplo, en medio ambiente (Siino et al., 2018; Nasirzadeh et al., 2021), presencia de plantas o animales (Møller y Díaz-Avalos, 2010; Myllymäki y Penttinen, 2009; Illian et al., 2012; Serra et al., 2014; Sørbye et al., 2019; Flagg y Hoegh, 2022), salud (Beneš et al., 2002; Rostami et al.,

3.2. DESCRIPCIÓN DE LOS DATOS

2017; Johnson et al., 2019), accidentes de tráfico (Tang et al., 2022), análisis de medios sociales (Ilhan y Kozat, 2020), entre otros.

En el ámbito delictivo la literatura de los LGCP es reducida a comparación con otro tipo de modelos, como por ejemplo, el modelo auto-excitado de Hawkes. Entre los pocos ejemplos destaca Shirota y Banerjee (2019), que aplican LGCP a datos de delitos en San Francisco e investigan la recuperación de la superficie de intensidad. González y Mateu (2021) utilizaron la metodología LGCP para eventos delictivos como homicidio, robo de automóviles, robo en tiendas, robo de motocicletas y otros tipos de robos.

En este capítulo se modelizan datos de delitos registrados durante el período 2010–2014 en la ciudad de Riobamba-Ecuador, en la línea de Taylor et al. (2013) y González y Mateu (2021) pero capturando la no linealidad de las covariables espaciales basadas en distancias. Se utiliza modelos lineales generalizados con base B-splines y efectos de variables meteorológicas para la componente temporal. Se realizan predicciones para seis semanas consecutivas al periodo analizado.

3.2. Descripción de los datos

La ciudad de Riobamba tiene aproximadamente 225741 habitantes. Está situada en el hemisferio sur (longitud 761780 y latitud 9815149), que corresponde a una localización en el centro del Ecuador, por lo que toma el nombre de *Corazón de la patria* (véanse la Figura 3.1 y en el capítulo 2 más detalles sobre esta ciudad). Esta característica en particular, hace de Riobamba una ciudad que concentra actividades gubernamentales, comerciales, laborales y educativas; dando lugar a un movimiento considerable en las actividades diarias de la población, y con ello provocación en el incremento en las acciones delictivas como robos a personas, viviendas y locales comerciales, asaltos, intimidaciones, hurtos, estafas, entre otros tipos de delitos (INEC, 2012), que causan confusión, preocupación y pérdidas importantes a la población en general (Chávez et al., 2013).

Para el estudio con datos reales de este capítulo, se utilizan los registros de eventos delictivos del período 2010–2014 (véase la Figura 3.2), proporcionados por las autoridades del Ministerio del Interior a mediados del 2015 bajo restricciones de confidencialidad. Los resúmenes y estadísticos descriptivos sobre delitos a partir del 2015 pueden ser descargados de la página web <http://cifras.ministeriodegobierno.gob.ec/comisioncifras/inicio.php>.

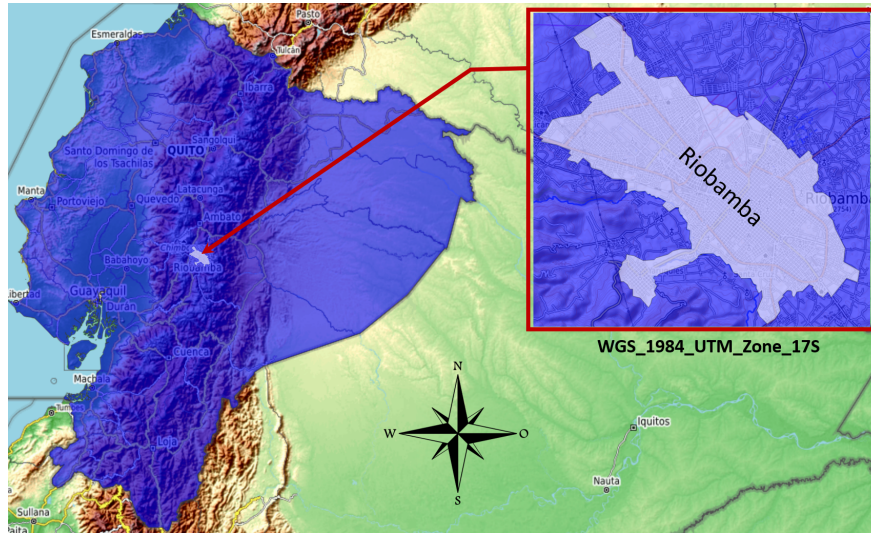


Figura 3.1: Ubicación geográfica de la ciudad de Riobamba-Ecuador



Figura 3.2: Delitos registrados durante el período 2010-2014 en la ciudad de Riobamba

Los actos delictivos en la ciudad de Riobamba presentan comportamientos acordes al calendario y horarios de las actividades comunes de la población, es decir, este fenómeno tiende a incrementarse de acuerdo al movimiento de

3.2. DESCRIPCIÓN DE LOS DATOS

la población al realizar sus actividades laborales, comerciales, recreativas, familiares, y de acuerdo al calendario de celebraciones o días festivos de la ciudad.

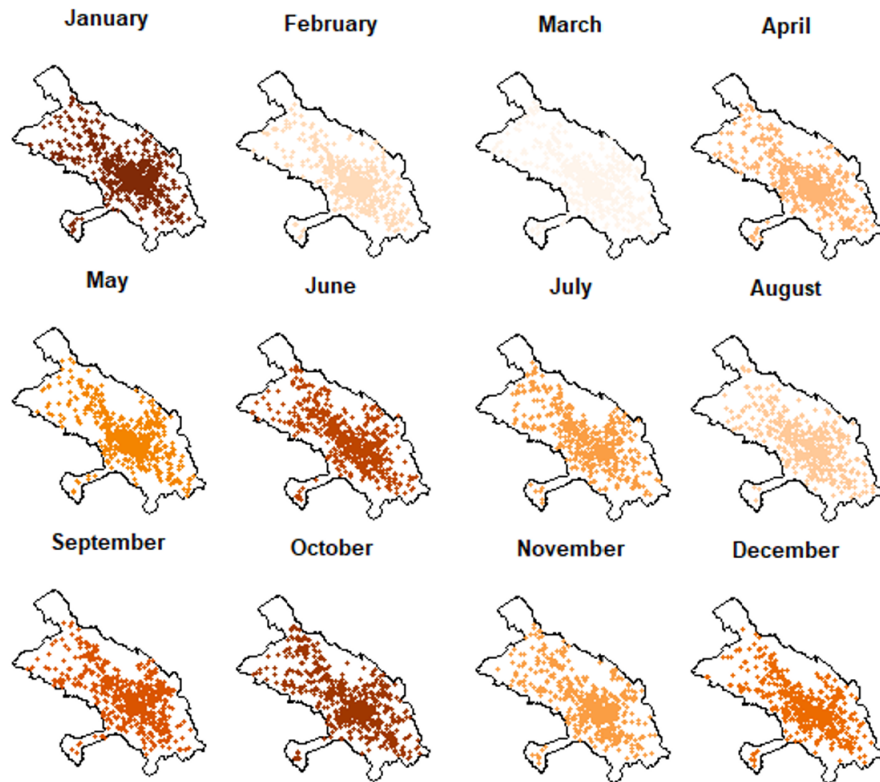


Figura 3.3: Eventos delictivos mensuales en la ciudad de Riobamba

La Figura 3.3 muestra los registros delictivos clasificados por meses. El degradado de colores indica mayor (color intenso) y menor (color claro) número de eventos. Se identifica el mayor número de eventos en el siguiente orden (de mayor a menor): enero, octubre y junio, seguidos de septiembre, diciembre y mayo, con una tendencia decreciente a partir de julio, noviembre, abril, agosto y febrero, y el menor número de registros en marzo. En general, en el año 2010 se ha registrado el menor número de delitos (aproximadamente 902 hechos), mientras que en 2014 superan los 3000.

Como referencias importantes para la modelización, se toman las distancias más cortas desde el evento delictivo hasta los lugares estratégicos como las unidades de policía comunitaria, las cámaras de vigilancia instaladas en lugares estratégicos de la ciudad (ECU911, 2020), áreas con mayor tráfico

de vehículos y peatones como parques, supermercados (plazas, mercados y supermercados) y hospitales, ya que los delitos se concentran en las áreas industriales, comerciales y recreativas, y los delincuentes actúan de acuerdo a su reloj biológico (Zhuang y Mateu, 2019).

Inicialmente se consideran como covariables temporales, los promedios semanales de la temperatura del ambiente, la sensación térmica, la radiación solar global, la temperatura del suelo y la humedad relativa (variables descargadas de <http://ceaa.esPOCH.edu.ec:8080/redEma/>). Esto debido a que el clima tiene un impacto en los robos (Andresen y Malleson, 2015), sin embargo, se identifica que sólo las tres primeras covariables temporales son significativas, por lo que, se utilizan estas tres variables en promedios semanales (véase la Figura 3.5).

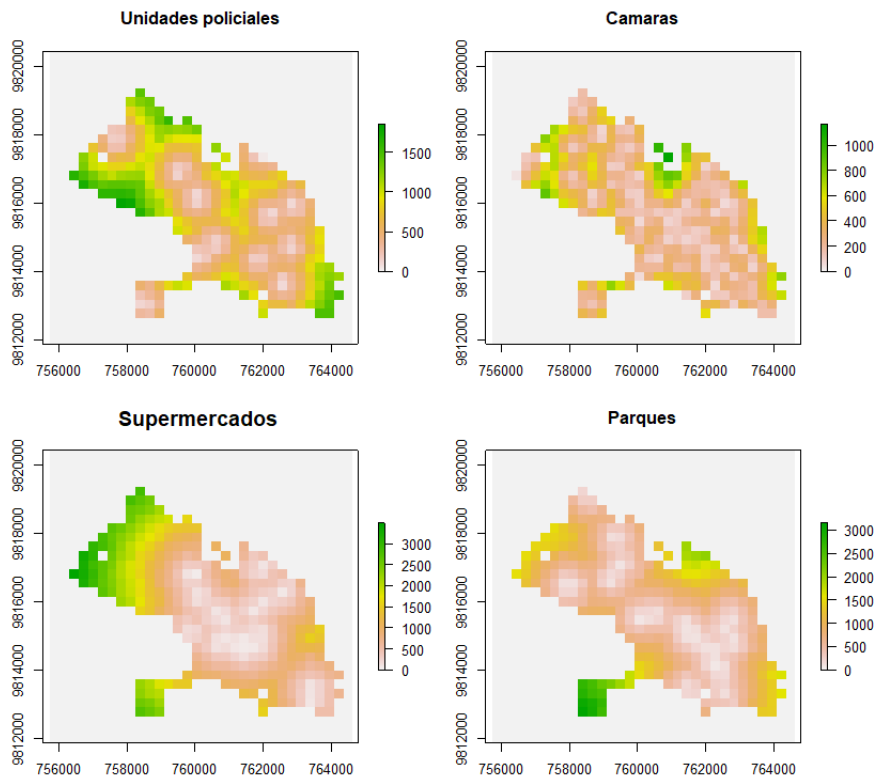


Figura 3.4: Covariables espaciales de la ciudad de Riobamba

Se utilizan como covariables espaciales las distancias a UPC, cámaras, supermercados y parques. Para el modelo final no se usan las ubicaciones de los hospitales porque no contribuyen a captar la varianza no explicada. Las

3.2. DESCRIPCIÓN DE LOS DATOS

UPC se sitúan en la zona central de la ciudad y algunas en los alrededores. Las cámaras de vigilancia muestran una distribución que abarca casi toda el área de estudio. Los supermercados se concentran en la parte más interna de la ciudad, y los parques están a lo largo de las calles principales desde el Sureste hasta el Noroeste (véase la Figura 3.4). Además, el color más claro indica el punto de referencia en cada covariable. A medida que avanza la decoloración se identifican las distancias más lejanas al lugar donde se ha cometido algún tipo de delito.

Se identifica que la mayoría de delitos registrados sucedieron a pequeñas distancias con respecto a los UPC, cámaras de vigilancia, supermercados y parques. Se cree que esto es debido al alto movimiento peatonal, aprovechando los delincuentes esta circunstancia, es decir, son puntos de atracción para los actos delictivos.

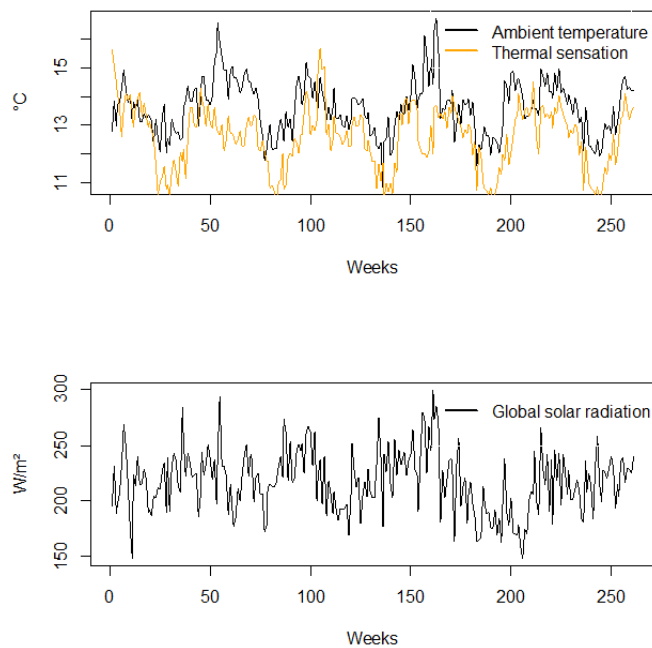


Figura 3.5: Promedio semanal de las covariables temporales

Las covariables temporales temperatura del aire, sensación térmica y radiación solar global muestran una periodicidad aproximadamente anual (véase la Figura 3.5), con los valores mínimos y máximos de $10.86^{\circ}C$ y $16.71^{\circ}C$, $9.63^{\circ}C$ y $15.68^{\circ}C$, y $148.26W/m^2$ y $644.91W/m^2$, respectivamente.

3.3. Metodología

Si un evento delictivo está dado en una ubicación \mathbf{s}_i y en un intervalo de tiempo t_i con $i = 1, \dots, n$, entonces un conjunto de eventos dados en $\mathbf{s} \in S \subseteq \mathbb{R}^2$ y tiempo $t \in T \subseteq \mathbb{R}$ se denomina patrón puntual, y el modelo estocástico que conduce los datos se denomina proceso puntual espacio-temporal.

Según un proceso de Cox log-gaussiano espacio-temporal no homogéneo, es decir un proceso de Poisson con intensidad estocástica $\Lambda(\mathbf{s}, t)$, el número de eventos $Z(\mathbf{s}, t)$ que surgen en cualquier $\mathbf{s} \in S$ durante $[t_1, t_2] \in T$ es una distribución de Poisson condicionada, dada por

$$Z_{\mathbf{s}, [t_1, t_2]} \sim \text{Poisson} \left\{ \int_S \int_{t_1}^{t_2} \Lambda(\mathbf{s}, t) d\mathbf{s} dt \right\}. \quad (3.1)$$

La función de intensidad $\Lambda(\mathbf{s}, t)$ se descompone en el producto

$$\Lambda(\mathbf{s}, t) = \lambda(\mathbf{s}, t) \exp \{ \mathcal{Y}(\mathbf{s}, t) \}, \quad (3.2)$$

donde la función $\mathcal{Y}(\mathbf{s}, t)$ es un proceso gaussiano continuo, tanto en el espacio como en el tiempo.

Considerando la separabilidad de primer orden, es decir $\lambda(\mathbf{s}, t) = \lambda(\mathbf{s})\mu(t)$, se tiene

$$\Lambda(\mathbf{s}, t) = \lambda(\mathbf{s})\mu(t) \exp \{ \mathcal{Y}(\mathbf{s}, t) \}, \quad (3.3)$$

donde $\mu(t)$ es la componente determinista temporal, $\lambda(\mathbf{s})$ la componente determinista espacial, y $\exp(\mathcal{Y}(\mathbf{s}, t))$ un campo de intensidad aleatorio log-gaussiano que modeliza la dependencia y la variación espacio-temporal de los eventos (componente estocástica espacio-temporal).

3.3.1. Componente determinista temporal

Para estimar la componente determinista temporal $\mu(t)$, se utilizan diversos modelos. En este capítulo se utiliza un modelo lineal generalizado (GLM) con estructura dada en [Diggle et al. \(2005\)](#) como sigue:

$$\hat{\mu}(t) = \delta_c + \delta_t + \delta_m + \alpha_1 \cos(wt) + \beta_1 \sin(wt) + \alpha_2 \cos(2wt) + \beta_2 \sin(2wt), \quad (3.4)$$

donde, δ_c denota el efecto de las covariables meteorológicas, δ_t el efecto de tendencia, δ_m el efecto del mes del año, α y β son los parámetros de regresión armónica, y w la periodicidad en las tasas de incidencia.

3.3.2. Componente determinista espacial

Para estimar la componente determinista espacial $\lambda(\mathbf{s})$, se utiliza un modelo aditivo generalizado (GAM) (Gasparrini et al., 2017; Siino et al., 2018) con B-splines univariantes (Durbán, 2015), de la forma

$$\ln(\hat{\lambda}(\mathbf{s})) = \beta_0 + \sum_{i=1}^p f_i(x_i(\mathbf{s})), \quad (3.5)$$

donde β_0 es la ordenada en el origen, $x_i(\mathbf{s})$ es una covariable espacial, $f_i(x_i(\mathbf{s}))$ es una función suave de las covariables definida en bases polinómicas, y p es el número de covariables.

Los GLM y GAM incorporan las relaciones no lineales existentes entre las covariables temporales y espaciales, y logran un rendimiento predictivo robusto (Goetz et al., 2011).

3.3.3. Componente estocástica espacio-temporal

Si $\mathcal{Y}(\mathbf{s}, t)$ es estacionario de segundo orden, la función de covarianza mínimamente parametrizada esta dada por

$$\mathbb{C}(\mathcal{Y}(\mathbf{s}_1, t_1), \mathcal{Y}(\mathbf{s}_2, t_2)) = \sigma^2 \gamma(\|\mathbf{s}_1 - \mathbf{s}_2\|, |t_1 - t_2|), \quad (3.6)$$

donde $\gamma(\cdot)$ es la función de correlación espacial isotrópica del campo aleatorio gaussiano, $\|\cdot\|$ es la distancia en \mathbb{R}^2 (norma euclídea), y $|t_1 - t_2|$ es la distancia en \mathbb{R} . Según Moller et al. (1998), bajo estacionariedad, si $\|\mathbf{s}_1 - \mathbf{s}_2\| = r$ y $|t_1 - t_2| = h$, la función de correlación por pares se expresa como

$$g(r, h) = \exp\{\sigma^2 \gamma(r, h)\}. \quad (3.7)$$

En una especificación separable, la función de covarianza del campo aleatorio gaussiano tiene la forma $\mathbb{C}(r, h; \phi, \theta) = \sigma^2 \gamma_1(r; \phi) \gamma_2(h; \theta)$. Nótese que esta estructura separable no implica necesariamente que la propiedad de separabilidad de segundo orden se cumpla, como se muestra en Moller y Ghorbani (2012) y Siino et al. (2018). Una estructura separable muy común es la familia exponencial (Brix y Diggle, 2001; González y Mateu, 2021), de la siguiente forma:

$$\mathbb{C}(r, h; \phi, \theta) = \sigma^2 \exp\left\{-\frac{r}{\phi}\right\} \exp\left\{-\frac{h}{\theta}\right\}, \quad (3.8)$$

donde σ es el parámetro de escala de intensidad logarítmica, ϕ es el parámetro de escala espacial y θ el parámetro de escala temporal; los dos últimos parámetros son las tasas a las que la función de correlación disminuye en

el espacio y en el tiempo. La media del proceso $\mathcal{Y}(\mathbf{s}, t)$ es $-\frac{\sigma^2}{2}$, de modo que, $E[\exp\{\mathcal{Y}(\mathbf{s}, t)\}] = 1$.

Para llevar a cabo la inferencia bayesiana aproximada para el campo gaussiano no observado o modelos gaussianos latentes, se puede utilizar el algoritmo de Monte Carlo con Cadenas de Markov (MCMC) o la aproximación integrada anidada de Laplace (INLA), en combinación con el Algoritmo de Langevin Ajustado por Metrópolis (MALA). Tanto MCMC (véase más detalles en [Taylor et al., 2013](#)) como INLA (véase más detalles en [Flagg y Hoegh, 2022](#)) proporcionan un enfoque rápido y bastante preciso para ajustar modelos gaussianos latentes complejos ([Taylor et al., 2013](#)), reforzando la inferencia espacio-temporal bayesiana para los campos aleatorios gaussianos ([Serra et al., 2014](#)). Sin embargo, se utiliza MCMC en combinación con el MALA para estimar la media y la varianza de $\exp(\mathcal{Y}(\mathbf{s}, t))$ para cada centroide de las celdas de la malla espacial y algún intervalo temporal, además de, utilizar muestras de las distribuciones predichas del proceso discretizado Y_{t_i, t_j} dado Z_{t_i, t_j} ([Brix y Diggle, 2001](#)), como

$$\pi(Y_{t_i: t_j} | Z_{t_i: t_j}) \propto \pi(Z_{t_i: t_j} | Y_{t_i: t_j}) \pi(Y_{t_i: t_j}), i < j, \quad (3.9)$$

donde las observaciones Z son los recuentos de celdas en una malla de tamaño $M \times N \equiv 2^m \times 2^n$ para números enteros positivos m y n , que se extiende a una malla de tamaño $2M \times 2N$ para el cálculo ([Moller et al., 1998](#)). Para evaluar $\pi(Y_{t_i: t_j})$ en la expresión (3.9), los parámetros del proceso Y deben ser conocidos o estimados a partir de los datos. La estimación de σ , ϕ y θ se realiza mediante contraste mínimo ([Moller et al., 1998](#); [Davies y Hazelton, 2013](#)), aunque también se puede realizar mediante un enfoque bayesiano o por métodos ad-hoc ([Zhai y Lafferty, 2017](#)).

Se utiliza la función g (véase la expresión 3.7) ([Baddeley et al., 2000](#); [Taylor et al., 2013](#)) para estimar los parámetros de covarianza espacial σ y ϕ del modelo, y la correlación del parámetro temporal θ .

En la inferencia para el proceso de Cox log-gaussiano, aplicando MALA, la simulación Monte Carlo de $\pi(Y_{t_i: t_j} | Z_{t_i: t_j})$, se vuelve más eficiente cuando se trabaja con una transformación lineal de Y , parcialmente determinada por la matriz C , donde $C^{1/2} = Q\mathcal{X}^{1/2}Q^*$, \mathcal{X} es una matriz diagonal de valores propios de C , Q es una matriz unitaria con entradas dadas por la transformada discreta de Fourier, y Q^* denota la transposición hermitiana de Q .

Si $\Gamma_t = \Lambda^{-1/2}Q\left(Y_t - \frac{\sigma^2}{2}\right)$, el objetivo de interés viene dado por,

$$\pi(\Gamma_{t_i: t_j} | Z_{t_i: t_j}) \propto \left[\prod_{t=t_i}^{t_j} \pi(Z_t | Y_t) \right] \left[\pi(\Gamma_{t_i}) \prod_{t=t_i+1}^{t_j} \pi(\Gamma_t | \Gamma_{t-1}) \right], \quad (3.10)$$

donde $\left[\prod_{t=t_i}^{t_j} \pi(Z_t | Y_t) \right]$ corresponde a $(Z_{t_i: t_j} | Y_{t_i: t_j})$, y el último término de la derecha de (3.10) es la densidad conjunta, $\pi(\Gamma_{t_i: t_j})$, por la propiedad de

Markov. Como Y y por tanto Γ es un proceso de Ornstein-Uhlenbeck, la densidad de transición $\pi(\Gamma_t|\Gamma_{t-1})$ tiene una expresión explícita como densidad gaussiana (Brix y Diggle, 2001). Dado que el gradiente de la densidad de transición también puede escribirse explícitamente, el algoritmo de Metrópolis-Hastings con una propuesta de tipo Langevin es un método MCMC natural y eficiente para el muestreo de la densidad de predicción (Roberts y Tweedie, 1996; Moller et al., 1998),

$$q(\Gamma, \Gamma') = N\left[\Gamma'; \Gamma + \frac{1}{2}\nabla \log\{\pi(\Gamma|Z)\}, h^2 I\right], \quad (3.11)$$

donde $N(y; \mu_y, \sigma_y)$ denota una densidad gaussiana con media μ_y y varianza σ_y evaluada en y , I es la matriz identidad y $h > 0$ es un parámetro de escala (Metropolis et al., 1953; Hastings, 1970).

3.3.4. Modelos de predicción

Para predecir los delitos después del último tiempo de observación, es necesario calcular la intensidad en $t_j + k$,

$$\Lambda(\mathbf{s}; t_j + k) = \lambda_0(\mathbf{s})\mu_0(t_j + k) \exp\{Y(\mathbf{s}, t_j + k)\}, \quad (3.12)$$

donde k es un entero positivo, y es necesario inferir la distribución de $Y(\mathbf{s}, t_j + k)$ dada la información observada. La independencia condicional del modelo conduce a

$$[Y(\mathbf{s}, t_j + k) | Z] = \int [Y(\mathbf{s}, t_j + k) | Y(\mathbf{s}, t_j)] [Y(\mathbf{s}, t_j) | Z] dY(\mathbf{s}, t_j). \quad (3.13)$$

Las estimaciones de predicciones de la media y la varianza vienen dadas por

$$E[Y(\mathbf{s}, t_j + k) | Z] = \exp\{-\theta k\} E[Y(\mathbf{s}, t_j) | Z] + (1 - \exp\{-\theta k\}) \mu \quad (3.14)$$

$$V[Y(\mathbf{s}, t_j + k) | Z] = (\exp\{-\theta k\})^2 V[Y(\mathbf{s}, t_j) | Z] + (1 - \exp\{-2\theta k\}) \sigma^2. \quad (3.15)$$

Con ello, se generan muestras aleatorias gaussianas con media según (3.14) y varianza según (3.15) para cada \mathbf{s}_i . Para evaluar el procedimiento de predicción, se calcula el error cuadrático medio (MSE) y el RMSE.

3.3.5. Modelo para la identificación de excedencias

En la misma línea que Diggle et al., 2005; Rodrigues y Diggle, 2012; González y Mateu, 2021, la probabilidad de excedencias puede expresarse como la esperanza, de la siguiente forma,

$$P[\exp\{Y(\mathbf{s}, t)\} > c | Z] = E_{[Y_{t_i:t_j} | Z]}(I\{\exp(Y_{t_i:t_j}) > c\}). \quad (3.16)$$

En la práctica, (3.16) permite identificar cualquier anomalía, es decir, si en algún lugar específico, la intensidad delictiva supera un umbral estipulado.

3.4. Resultados

Para el modelo LGCP con estructuras separables, se estima la componente determinista temporal $\mu(t)$ con soporte $t \in \{1, 2, \dots, 255\}$, que corresponde a 255 semanas.

Tabla 3.1: Coeficientes estimados mediante GLM para la componente temporal

Covariables	Coef.	Std.Error	val. z	Pr(> z)	Signif.
(Intercept)	81.44	10.02	8.13	0.00	***
Temperature	-0.47	0.02	-2.35	0.02	*
Solar radiation	0.002	0.00	4.69	0.00	***
Thermal sensation	0.06	0.02	3.67	0.00	***
Linear trend	-60.75	7.31	-8.31	0,00	***
February	0.09	0.05	1.65	0.10	.
March	-0.13	0.06	-2.31	0.02	*
April	-0.16	0.06	-2.90	0.01	**
May	-0.12	0.06	-2.16	0.03	*
June	-0.14	0.06	-2.56	0.01	*
July	-0.04	0.05	-0.77	0.44	
August	0.10	0.05	1,97	0.05	*
September	-0.09	0.06	-1.54	0.12	
October	0.075	0.05	1.42	0.16	
November	-0.03	0.06	-0.58	0.56	
December	0.04	0.05	0.74	0.46	
Tuesday	-0.10	0.04	-2.39	0.02	*
Wednesday	-0.14	0.04	-3.36	0.00	***
Thursday	-0.04	0.04	-0.96	0.34	***
Friday	0.01	0.04	0.19	0.85	
Saturday	-0.08	0.04	-1.81	0.07	.
Sunday	-0.11	0.04	-2.59	0.01	**
Sine	67.24	7.44	9.04	0.00	***
Sine2	5.37	0.98	5.50	0.00	***
Cosine	-95.32	11.78	-8.09	0.00	***
Cosine2	16.46	1.84	8.94	0.00	***

Se incluyen en el modelo lineal generalizado (GLM) los efectos de covariables temporales como temperatura del ambiente (δ_{at}), sensación térmica (δ_{st}), radiación solar global (δ_{rg}) (véase la Figura 3.5), los efectos del día de la semana (δ_d) tomando el lunes como base, los efectos del mes del año (δ_m) tomando enero como base, los parámetros de regresión armónica ($\alpha = 1/8$ y $\beta = 1/8$), la periodicidad semanal de las tasas de incidencia (w) y la pendiente

3.4. RESULTADOS

para la tendencia global (γ_t).

La Tabla 3.1 muestra los coeficientes estimados con el valor AIC más bajo ($AIC = 2047.9$), siendo las variables más significativas (***) la radiación solar global, la sensación térmica, la tendencia, el miércoles, el jueves, y tres términos que modelizan el componente periódico. Las variables radiación solar global, sensación térmica, los meses de febrero, agosto, octubre, diciembre, viernes y tres componentes periódicas influyen en el aumento en la tasa de eventos delictivos, mientras que la temperatura ambiental, la tendencia y los meses de marzo, abril, mayo, junio, julio, septiembre, noviembre, martes, sábado y domingo tienden a reducirla (véase la Tabla 3.1).

La serie temporal de los sucesos delictivos semanales originales muestra un crecimiento acelerado a partir de la semana 200 aproximadamente (véase la Figura 3.6), alcanzando frecuencias que duplican el número de eventos identificados en semanas anteriores. Se cree que posiblemente se deba a las mejoras en el sistema de registro de actos delictivos, cambios en las autoridades, mejoras en las políticas de vigilancia y seguridad, aumento del personal policial (Castro et al., 2015), o cualquier otra razón. Sin embargo, se decide incluir estos datos en el análisis para identificar si los modelos propuestos son capaces de captar la variabilidad de esta naturaleza.

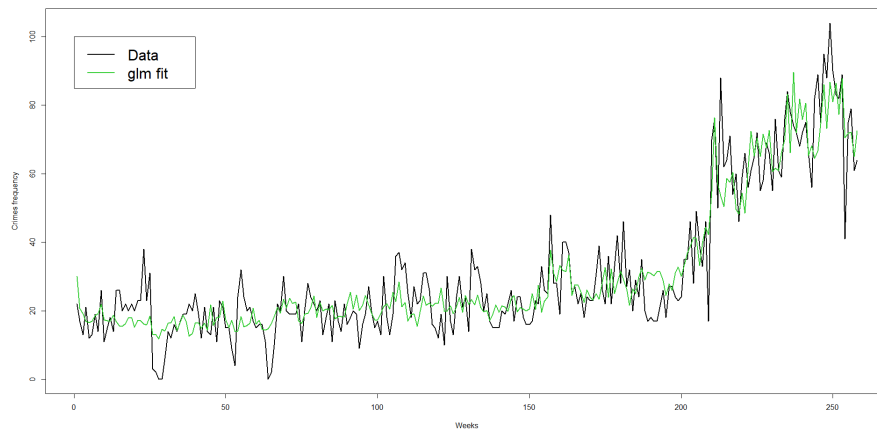


Figura 3.6: GLM para la componente determinista temporal $\mu(t)$

El GLM (línea verde de la Figura 3.6) trata de alcanzar los extremos de variación de los eventos delictivos originales; sin embargo, no los alcanza, siendo insuficiente por sí mismo.

Para la componente determinista espacial $\lambda(\mathbf{s})$ se define una malla de 32×32 , dando un área aproximada de $996m^2$ en cada celda; sin embargo,

pueden elegirse mallas con mayor o menor dimensión en función de la capacidad para cubrir su coste computacional.

Tabla 3.2: Coeficientes estimados mediante GAM con B-splines para la componente espacial

Covariates	Coefficients	S.E.	CI95.lo	CI95.hi	Zval	Ztest
(Intercept)	-9.57	0.04	-9.66	-9.49	-218.41	***
bs(im.UPC)1	7.22	0.17	6.89	7.55	43.08	***
bs(im.UPC)2	-2.54	0.14	-2.81	-2.27	-18.46	***
bs(im.UPC)3	4.52	0.15	4.22	4.81	29.84	***
bs(im.cam)1	3.85	0.22	3.41	4.29	17.17	***
bs(im.cam)2	-6.45	0.27	-6.98	-5.93	-24.10	***
bs(im.cam)3	2.26	0.27	1.73	2.78	8.45	***
bs(im.cc)1	-3.15	0.14	-3.42	-2.88	-22.52	***
bs(im.cc)2	-0.97	0.22	-1.40	-0.53	-4.38	***
bs(im.cc)3	-1.29	0.22	-1.72	-0.87	-6.00	***
bs(im.par)1	-0.47	0.17	-0.81	-0.13	-2.73	**
bs(im.par)2	-4.63	0.26	-5.15	-4.11	-17.56	***
bs(im.par)3	0.20	0.18	-0.15	0.56	1.11	

La relación de cada predictor con la media de la variable respuesta no es directa, sino a través de la función $f_i(X_i(\mathbf{s}))$ indicada en la expresión (3.5). La estimación GAM con coeficientes B-splines puede verse en la Tabla 3.2, donde las covariables espaciales fueron altamente significativas para cada B-spline. Las distancias a UPC y cámaras de vigilancia en su mayoría aportan a incrementar la intensidad espacial delictiva, mientras que las distancias a supermercados y parques aportan en la reducción.

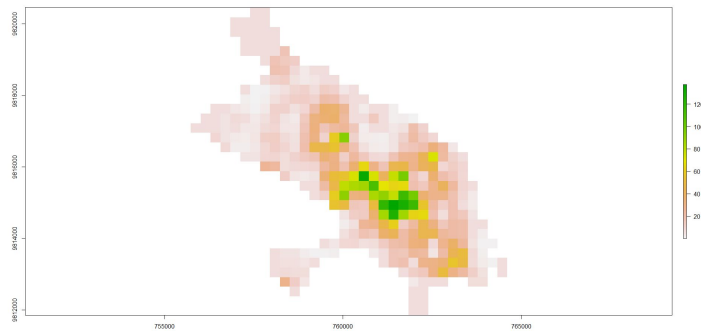
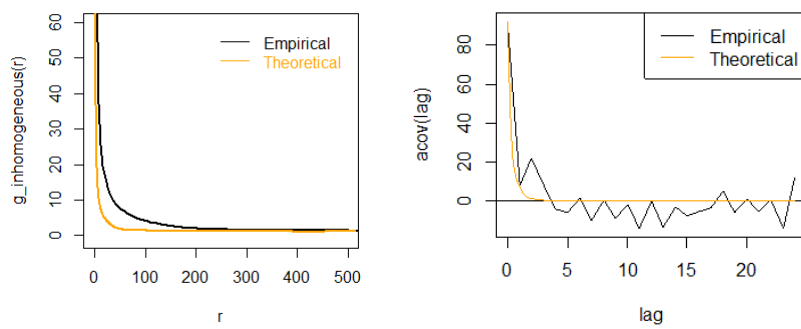


Figura 3.7: GAM para la componente determinista espacial $\lambda(\mathbf{s})$

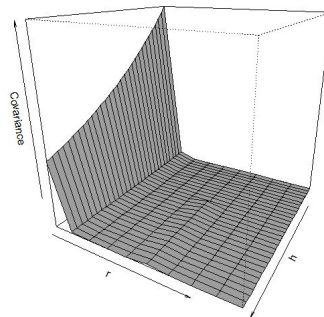
3.4. RESULTADOS

En la Figura 3.7 los colores verdes intensos significan una mayor frecuencia de actos delictivos en esas áreas (centro de la ciudad) y muestran alrededor de 120 eventos, mientras que las áreas de color claro representan aproximadamente 20 eventos delictivos. La zona del centro de la ciudad es el lugar preferible por los delincuentes para cometer cualquier acto delictivo.



(a) Estimación de σ y ϕ

(b) Estimación de θ



(c) Autocovarianza espacio-temporal.

Figura 3.8: Estimación de los parámetros de escala espacial y temporal mediante la función de correlación por pares

La Figura 3.8 muestra las funciones de correlación por pares y la covarianza del proceso estocástico evaluado sobre el conjunto de datos. Los parámetros de escala espacial σ y ϕ se estiman mediante la función g (véase la Figura 3.8b). La función de autocorrelación de los recuentos de eventos por

unidad de tiempo estiman el parámetro de escala temporal θ (véase la Figura 3.8b) y la función de covarianza espacio-temporal (véase la Figura 3.8c). Para un proceso de Poisson espacio-temporal la covarianza es cero y la función de correlación por pares es 1.

Mediante el método de contraste mínimo se obtiene que $\hat{\phi} = 107.21$, $\hat{\sigma} = 1.89$ y $\hat{\theta} = 1.02$. Las funciones de covarianza y autocovarianza exponenciales muestran que las curvas empíricas tienen un comportamiento similar a las teóricas, es decir, el ajuste con covariables tanto temporales como espaciales es adecuado; la Figura 3.8 corrobora lo expuesto.

Para modelizar la componente estocástica, se utiliza un procedimiento MCMC con 300000 iteraciones en el algoritmo MALA, un período de burn-in de 30000, y una frecuencia de 100 para almacenar los cálculos y predecir los delitos desde la semana 256 hasta la 261. Las predicciones se calculan mediante simulaciones de las medias y varianzas estimadas (expresiones 3.14 y 3.15).

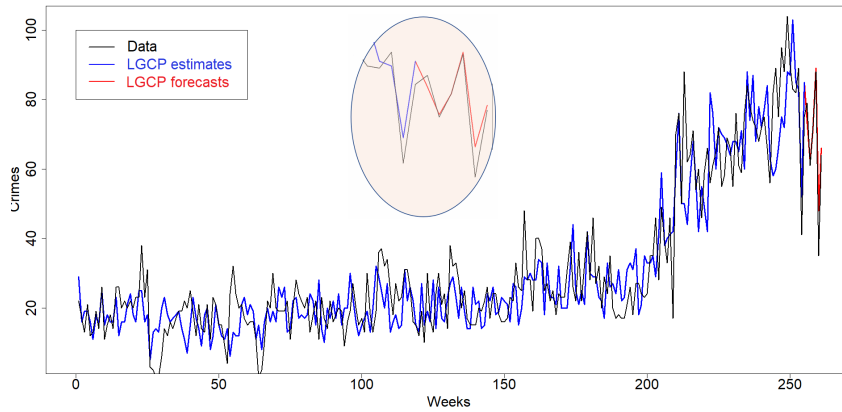
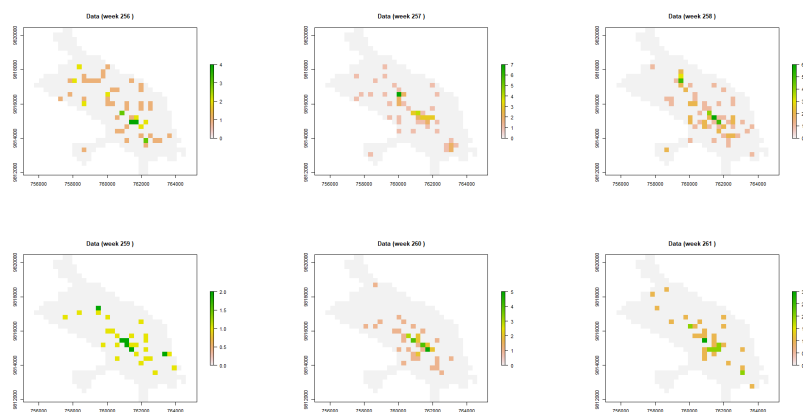


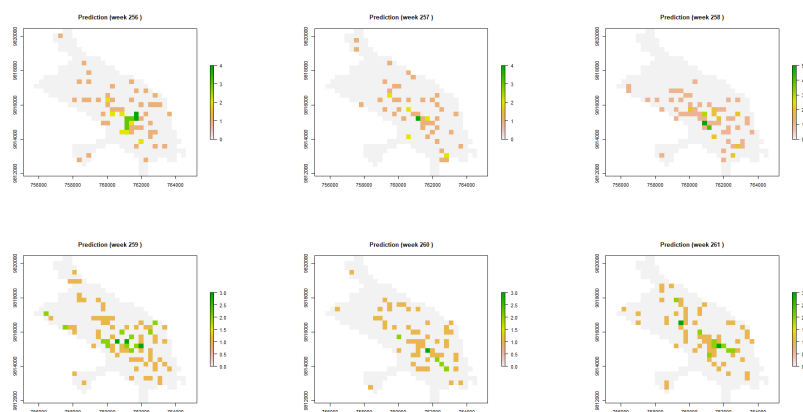
Figura 3.9: Series temporales de los delitos reales y predicciones

En el ajuste temporal, la contribución de la modelización LGCP (véase la Figura 3.9) es muy notable con respecto a la componente temporal determinista, es decir, el modelo alcanza los extremos más altos y más bajos de la serie temporal de los eventos originales. Las predicciones (línea roja) tienen un comportamiento bastante aproximado a los datos reales. Por ejemplo, en la ampliación que se muestra dentro de la Figura 3.9, se puede observar la proximidad entre los delitos reales y la predicción con más detalle. Para las semanas 256 a 261 tienen un comportamiento muy similar. La media de las diferencias es de 7 delitos (véase la Figura 3.11).

3.4. RESULTADOS



(a) Datos originales.



(b) Predicciones con LGCP.

Figura 3.10: Delitos originales y predicciones para las semanas [256, 261]

En la Figura 3.10 se observan los delitos tanto para los datos originales como para las predicciones. Los colores claros muestran el menor número de delitos, mientras que los tonos verdes intensos identifican los valores máximos de eventos encontrados en cada semana. En la Figura 3.10a se muestran algunas localizaciones que llegan hasta un máximo de 7 y 2 eventos, específicamente para las semanas 257 y 259, respectivamente, mientras que en las predicciones con LGCP se observan hasta un máximo de 5 delitos en las semanas 258 y 3 desde la semana 259 hasta la 261 (véase la Figura 3.10b).

La distribución de los delitos en los datos originales y las predicciones muestra una tendencia a la aglomeración hacia el centro de la ciudad. Se hallan

los errores cuadráticos medios (MSE) desde el enfoque temporal y espacial. El MSE para las predicciones en el periodo de observación es 105.88 y su correspondiente raíz cuadrada (RMSE) de 10.29. La Figura 3.11 representa el comportamiento de las diferencias entre delitos reales y predicciones en función del tiempo. Los puntos sobre la línea horizontal en cero identifican sobrestimación, mientras que los puntos por debajo corresponden a las cifras de delincuencia subestimadas. Los valores del MSE en función del espacio para las seis semanas de predicción son 1.31, 0.69, 0.72, 0.65, 0.83, 0.96, respectivamente.

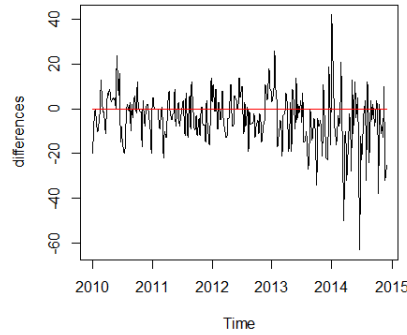


Figura 3.11: Diferencias temporales entre los delitos originales y predicciones con LGCP

Para identificar las superficies de excedencia, se fija como umbrales los valores correspondientes a los cuantiles 90 %, 75 % y 50 % de la distribución marginal estimada de $\exp\{Y(s; t)\}$, mostrando mayor concentración en el centro de la ciudad (véase la Figura 3.12). En la Figura 3.12 el color azul intenso indica los lugares con el menor número de eventos y el rojo las áreas con las más altas frecuencias de delitos. Con estos resultados se identifica que gran parte de los sectores Loma de Quito, La Estación, Santa Rosa y La Concepción son los que requieren mayor vigilancia policial.

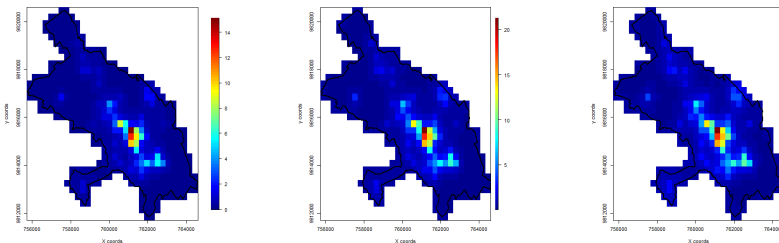


Figura 3.12: Probabilidad de excedencias a posteriori

3.5. Discusión y conclusiones

Los LGCP son procesos de Poisson no homogéneos con intensidades aleatorias log-gaussianas. Su función de intensidad es multiplicativa, permitiendo incorporar información de primer orden mediante componentes deterministas temporales y espaciales e información de segundo orden mediante una componente estocástica. La elección de la función de correlación espacio-temporal debe ser cuidadosa para obtener una estimación adecuada del proceso gaussiano subyacente.

La motivación para la aplicación de la estructura LGCP espacio-temporal es el fenómeno delictivo de la ciudad de Riobamba. Al analizar los datos de delitos de esta ciudad se ha encontrado que el mes de enero posee mayor número de registros, lo que se cree coherente debido a las festividades de los últimos días del mes de diciembre y los primeros días de enero. Para la estimación de la intensidad del proceso se considera una estructura separable y función de covarianza de forma exponencial, tanto para la componente temporal como para la espacial (Brix y Diggle, 2001).

Siguiendo la propuesta de González y Mateu (2021) aplicada a los delitos de la ciudad de Riobamba, proporciona los siguientes valores para los parámetros estimados: $\sigma^2 = 1.91$ y $\phi = 235.77$ (metros) para la varianza y dependencia espacial, respectivamente, y $\theta = 1.07$ (semanas) para la dependencia temporal. El cálculo del MSE para las respuestas espaciales de las seis semanas de predicciones proporciona valores de 1.58, 1.16, 0.98, 1.46, 1.68 y 1.31, respectivamente. La propuesta presentada en este capítulo incorpora una componente temporal mediante un GLM con efectos de los promedios semanales de temperatura del ambiente, sensación térmica y radiación solar global, los días de la semana, meses del año y ciclicidad semanal a través de las funciones seno y coseno; una componente espacial mediante un GAM con bases B-splines, con efectos de covariables basadas en distancias a UPC, cámaras de vigilancia, supermercados y parques; y una componente estocástica mediante algoritmo MCMC con iteraciones MALA (Diggle et al., 2005; Taylor et al., 2013; González y Mateu, 2021). Se obtiene valores estimados para los parámetros del modelo como sigue: $\sigma^2 = 1.89$ para la varianza espacial, $\phi = 107.21$ (metros) para la dependencia espacial y $\theta = 1.02$ (semanas) para la dependencia temporal. Los valores de MSE de las predicciones de seis semanas son 1.31, 0.69, 0.72, 0.65, 0.83 y 0.96, respectivamente. Como se puede observar, las características incluidas en nuestra propuesta reducen los errores en las predicciones.

La identificación de excedencias, es decir, las ocurrencias de eventos delictivos que superan a los cuantiles 90 %, 75 % y 50 %, corresponden a gran parte de los sectores denominados *Loma de Quito*, *La estación*, *Santa Rosa* y *La Concepción*, por lo que estos serían los lugares que requieren mayor atención policial.

4. Modelos espacio-temporales semiparamétricos auto-excitados de Hawkes

4.1. Introducción

Los Self-exciting Hawkes Processes (SEHP) son modelos propuestos por [Hawkes \(1971b\)](#). Surgen como un resultado particular de los procesos temporales de Hawkes, pero con las mismas propiedades de segundo orden de un proceso doblemente estocástico. [Hawkes y Oakes \(1974\)](#) demostraron que cualquier proceso puntual estacionario auto-excitado con intensidad finita se puede interpretar como un proceso de conglomerados de Poisson. Son ampliamente utilizados para pronosticar la tasa de eventos como una función del espacio, el tiempo y su historia previa.

Los SEHP se dividen en dos subprocesos, uno de background (centros de conglomerados) y otro de triggering (desencadenante). El primero es considerado como un proceso de Poisson, que puede ser no homogéneo en el espacio y/o no estacionario en el tiempo, y el segundo está compuesto por el efecto excitante de todos los eventos ocurridos en el pasado. En otras palabras, una vez que ocurre un evento en el proceso, sin importar si es un evento de background o un evento excitado por otros, excita un proceso de su propia descendencia directa de acuerdo con algunas reglas de probabilidad. Más concretamente, los SEHP capturan el comportamiento triggering y de background de forma natural. Estas características permiten identificar algún tipo de dependencia auto-excitada ([Hawkes, 1971a](#); [Hawkes y Oakes, 1974](#); [Ogata, 1988](#)).

A partir de su introducción se han desarrollado múltiples herramientas potentes como, por ejemplo, la desclasificación y reconstrucción estocástica basada en modelos ([Zhuang et al., 2002](#)), en la que se supone que la función triggering tiene una forma paramétrica, pero el background debe estimarse

de forma no paramétrica a partir de los eventos de background, y a la vez se requiere determinar si cada suceso fue desencadenado por el background, por lo que es un procedimiento iterativo. En cuanto a la desclasificación estocástica independiente del modelo, [Marsan y Lengliné \(2010\)](#) demuestran que puede considerarse como un algoritmo de maximización de valores esperados.

Más adelante [Chiodi y Adelfio \(2011\)](#) introducen el enfoque predictivo basado en probabilidad directa, en el que, en lugar de maximizar directamente la verosimilitud, consideran incrementos log-verosimilitud, utilizando los eventos pasados para predecir el evento futuro. Este enfoque se convirtió en un método semiparamétrico con el desarrollo de [Adelfio y Chiodi \(2015\)](#). Entre otros avances, se han realizado procedimientos de simulación para la estimación de máxima verosimilitud ([Ozaki, 1979](#); [Cui et al., 2019](#)); metodologías de estimación totalmente no paramétrica para conocer la forma de la función de triggering espacio-temporal y las tendencias temporales de la tasa de background ([Mohler et al., 2011](#)); incorporación de las características espaciales, la casi repetición, los efectos triggering, métodos de inferencia y herramientas de diagnóstico ([Reinhart y Greenhouse, 2017](#)); generalización del método de reconstrucción estocástica no paramétrica, para estimar cada componente en la tasa de background y la respuesta triggering que aparece en la intensidad condicionada del modelo ([Zhuang y Mateu, 2019](#)).

Los SEHP han sido ampliamente utilizados en campos donde se observa la agrupación espacio-temporal de eventos de conteo ([Reinhart, 2018](#)). Por ejemplo, en medio ambiente ([Lieshout et al., 2012](#); [Pratiwi et al., 2017](#); [Reinhart, 2018](#)), salud ([Chiang et al., 2022](#)), finanzas ([Ait-Sahalia et al., 2015](#); [Hawkes, 2018](#); [Moreno Trujillo, 2019](#)), delitos como se describe más adelante, entre otros.

[Reinhart \(2018\)](#) describe algunas aplicaciones en datos sobre delitos; sin embargo, se mencionan otros ejemplos, como el modelo Next Hit Predictor (NHP), que encuentra la ubicación más probable del próximo delito en serie a través de un modelo de riesgo cuidadosamente diseñado, habiendo demostrado resultados prometedores en décadas de datos de delitos en serie ([Li y Wang, 2018](#)). [Zhu y Xie \(2022\)](#) capturan los vínculos de los incidentes de la delincuencia multivariante marcada, tratando de extraer del texto libre registrado marcas del incidente, con el propósito de incluir la noción de *modus operandi*. Se introduce un novedoso marco de inferencia basado en transformaciones aleatorias y descenso de gradiente para aprender el proceso, para lo cual [Ilhan y Kozat \(2020\)](#) integran los cálculos del núcleo espacial por transformaciones aleatorias basadas en características de Fourier y lo aplican para incluir la predicción de delitos. [Liu y Vanhatalo \(2020\)](#) diseñan un método de regularización de gráficos para integrar de forma eficaz la estructura espacial previa para el aprendizaje de la matriz de influencia entre los diferentes lugares. [Park et al. \(2021\)](#) proponen un algoritmo marcado para estimar la tasa de background no paramétrica en función de las

4.2. DESCRIPCIÓN DE LOS DATOS

covariables demográficas, utilizando datos de delitos violentos relacionados con las bandas. [Yuan et al. \(2021\)](#) desarrollan un método no paramétrico para la reconstrucción de la red a partir de conjuntos de datos espacio-temporales multivariantes, utilizando tanto la información temporal como espacial, y no asumen una forma paramétrica específica de la dinámica de la red. [Okawa et al. \(2021\)](#) amplían una red neuronal convolucional (CNN) combinándola con la convolución de núcleo continuo, y diseñan la intensidad condicionada del proceso de Hawkes basada en el modelo de red neuronal ampliado que acepta imágenes. Realizan varias aplicaciones, entre ellas a datos de conflictos armados y protestas.

Entre los más recientes, [Jun y Cook \(2022\)](#) muestran una generalización de la estructura de la función de intensidad espacio-temporal, permitiendo la no-separabilidad, la no-estacionariedad y la intensidad cruzada (entre grupos). Utilizan muestras de datos de terrorismo en Afganistán para el análisis univariante y de Nigeria para el bivalente. Mencionan que sus modelos superan a los modelos estándar del proceso de Hawkes.

En este contexto, motivados por el potencial de los SEHP, se modeliza el fenómeno delictivo de la ciudad de Riobamba-Ecuador, como complemento a los resultados obtenidos con el LGCP (véase el capítulo 3), con el propósito de estimar el background y cuantificar el efecto triggering.

4.2. Descripción de los datos

Se utilizan los datos de eventos delictivos descritos en la Sección 3.2 del capítulo 3, con el fin de aprovechar el potencial de cada una de sus estructuras estocásticas y conseguir una caracterización más completa del fenómeno delictivo de la ciudad de Riobamba-Ecuador.

Para el modelizado con SEHP es importante tener en cuenta la unidad temporal, puesto que de ello depende el planteamiento de su periodicidad. Para los delitos registrados en la ciudad de Riobamba-Ecuador, se toma como unidad temporal el día de 24h, es decir, se determina que como mínimo se puede reconstruir la periodicidad semanal y tendencia a largo plazo (mientras más pequeña sea la unidad temporal, más detallados serán los resultados).

En la Figura 4.1a se muestra una descripción del historial de los datos de delitos mediante un degradado de colores, es decir, en función del tiempo. Los puntos en color rojo representan los primeros sucesos del 2010 y en magenta los últimos sucesos del 2014. En la distribución acumulativa (véase la Figura 4.1b) se puede observar crecimiento acelerado aproximadamente a partir del día 1500, lo que indica mayor frecuencia delictiva en el último año del periodo de estudio.

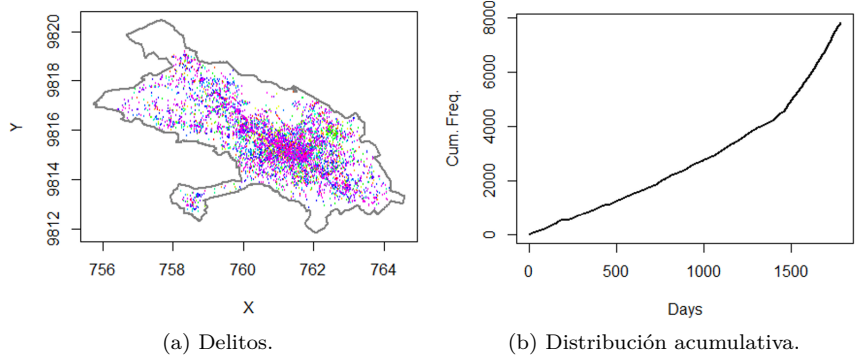


Figura 4.1: Distribución de los eventos delictivos en la ciudad de Riobamba-Ecuador y su función acumulativa

Las Figuras 4.2a y 4.2b despliegan color intenso (magenta) en las coordenadas y días en los que se ha registrado un mayor número de eventos delictivos, lo que contrasta con la Figura 4.1b descrita anteriormente. Además, se identifica que los eventos son algo más frecuentes alrededor de las longitudes 9815(1000) y 9816(1000), y latitud 762(1000).

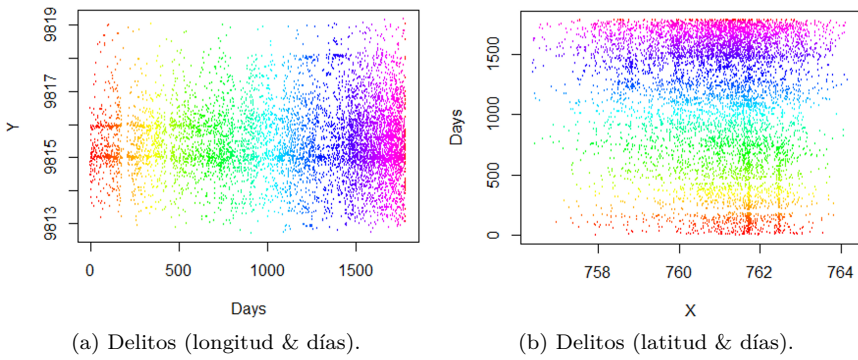


Figura 4.2: Historial de eventos delictivos en la ciudad de Riobamba-Ecuador

4.3. Metodología

El proceso de Hawkes describe los mecanismos de excitación entre una serie de eventos que ocurren en un dominio de tiempo o espacio-tiempo continuo,

4.3. METODOLOGÍA

que se explican detalladamente en Hawkes (1971b), Hawkes (1971a), Zhuang y Mateu (2019), Schoenberg et al. (2019), Chiang et al. (2022).

Hawkes menciona que el proceso auto-excitado con la intensidad condicionada $\Lambda(t)$ puede escribirse como

$$\Lambda(t) = \mu + \int_{-\infty}^{t-} g(t-l)d\mathcal{N}(\tau), \quad (4.1)$$

donde μ es la tasa de ocurrencia de eventos espontáneos (también llamados eventos de background), $g(t)$ es la tasa de ocurrencia de la descendencia directa generada por un evento ocurrido antes de t (denotado por t_-), y \mathcal{N} es un proceso puntual denotado por $\mathcal{N} = \{t_i, i \in T\}$. Es importante señalar que $l < t$ y μ y g son no negativos.

Para el proceso de Hawkes espacio-temporal la función de intensidad condicionada viene dada por

$$\Lambda(t, x, y) = \mu(t)\mu(x, y) + \int_{-\infty}^{t-} \int_S g(t-l, x-u, y-v)d\mathcal{N}(u \times v \times l), \quad (4.2)$$

donde $\mu(t)$ es la componente temporal de la tasa de background, $\mu(x, y)$ representa la tendencia espacial de la tasa de background en (x, y) , y $g(t-l, x-u, y-v)$ el subproceso triggering por un evento ocurrido previamente en (u, v) en el tiempo l , siendo (x, y) y (u, v) localizaciones de eventos dados en un área $S \in \mathbb{R}^2$.

La función de intensidad condicionada espacio-temporal para los delitos de la ciudad de Riobamba, utilizando el método de reconstrucción estocástica para la estimación no paramétrica de μ_t , μ_t , μ_W , μ_S y g , está dada por

$$\begin{aligned} \Lambda(t, x, y) = & \mu_0 \mu_t(t) \mu_W(t) \mu_s(x, y) \\ & + A \int_{-\infty}^{t-} \int_S g(t-l) h(x-u, y-v) d\mathcal{N}(u \times v \times l), \end{aligned} \quad (4.3)$$

donde μ_0 y A son los coeficientes de relajación a estimar, $\mu_t(t)$ y $\mu_W(t)$ son el término de tendencia y la periodicidad semanal en los componentes temporales de la tasa de background, respectivamente, $\mu_s(x, y)$ representa la tendencia espacial de la tasa de background, y g y h son funciones de densidad de probabilidad, es decir, $\int_0^\infty g(l)dl = 1$, $\int_S h(u, v)dudv = 1$. La función de respuesta de background espacio-temporal está separada en una componente espacial y otra temporal, para evitar la estimación no paramétrica de una función en tres dimensiones.

Las componentes periódicas de la tasa de background formuladas en la expresión (4.3) no se pueden estimar directamente mediante el método de

reconstrucción estocástica, por lo que se utiliza el método residual (véase [Baddeley et al., 2005](#); [Zhuang, 2006](#); [Zhuang y Mateu, 2019](#)), es decir, mediante la propiedad de la intensidad condicional

$$\left[\int_{[T_1, T_2] \times S} f(t, x) d\mathcal{N}(dt \times dx) \right] = E \left[\int_{T_1}^{T_2} \int_S f(t, x) \Lambda(t, x) dt dx \right], \quad (4.4)$$

donde $[T_1, T_2]$ es un intervalo de tiempo, S es el área, \mathcal{N} es el proceso espacio-temporal con intensidad $\Lambda(t, x)$ y $f(t, x)$ es un proceso predecible. Siempre que exista la integral en cualquier lado o que f no sea negativa.

4.3.1. Componente de background

Dada una realización de \mathcal{N} , con t días y localizaciones (x, y) , la tasa de background esta compuesta por el término de tendencia a largo plazo $\mu_t(t)$, la periodicidad semanal $\mu_W(t)$ y la tendencia espacial $\mu_S(x, y)$. Según [Zhuang y Mateu \(2019\)](#), la reconstrucción de cada término está dada como sigue:

$$\hat{\mu}_t(t) \propto \sum_i w_i^{(t)} I(t_i \in [t - \Delta_t, t + \Delta_t]), \quad (4.5)$$

$$\hat{\mu}_w(t) \propto \sum_i w_i^{(W)} I(t_i \in [t + 7 - \Delta_t, t + 7 + \Delta_t]),$$

$$\hat{\mu}_s(x, y) \propto \sum_i \varphi_i I(x_i \in [x - \Delta_x, x + \Delta_x]) I(y_i \in [y - \Delta_y, y + \Delta_y]),$$

donde $w_i^{(t)}$ y $w_i^{(W)}$ son las ponderaciones global y semanal de los datos para todos los i y j , φ_i es la probabilidad de que el evento i sea un evento de background y Δ_t , Δ_x y Δ_y son números positivos pequeños. Los valores de las ponderaciones son la clave para la reconstrucción de los términos de la tasa de background.

Las ponderaciones están dadas por

$$w_i^t = \frac{\mu_t(t_i) \mu_S(x_i, y_i)}{\Lambda(t_i, x_i, y_i)}, \quad (4.6)$$

$$w_i^W = \frac{\mu_W(t_i) \mu_S(x_i, y_i)}{\Lambda(t_i, x_i, y_i)},$$

y

$$\varphi_i = \frac{\mu_0 \mu_t(t_i) \mu_W(t_i) \mu_S(x_i, y_i)}{\Lambda(t_i, x_i, y_i)},$$

NOTA: $w_i^{(t)}$ en su forma general es $w^{(t)}(t, x, y) = f(t, x, y)$. Sustituyendo en (4.4) se obtienen los términos de la componente de background (véase algunos detalles en [Zhuang y Mateu, 2019](#)).

4.3.2. Componente de triggering

Según la reconstrucción de [Zhuang y Mateu \(2019\)](#), la estimación de g se realiza mediante

$$\hat{g}(t) \propto \sum_{i,j} \rho_{ij} I(t_j - t_i \in [t - \Delta_t, t + \Delta_t]), \quad (4.7)$$

donde ρ es la componente de excitación, con $i < j$, dada por

$$\rho_{ij} = \frac{g(t_j - t_i)h(x_j - x_i, y_j - y_i)}{\lambda(t_j, x_j, y_j)}, \quad (4.8)$$

y h se estima de la forma

$$\hat{h}(x, y) \propto \sum_{i,j} \rho_{ij} \times I(x_j - x_i \in [x - \Delta_x, x + \Delta_x]) I(y_j - y_i \in [y - \Delta_y, y + \Delta_y]). \quad (4.9)$$

donde Δ_x y Δ_y son números positivos pequeños.

4.3.3. Estimación de coeficientes

Una vez estimados μ_t , μ_w , μ_s , g y h se actualizan los coeficientes μ y A por máxima verosimilitud.

La función de verosimilitud está dada por

$$\log(L) = \sum_{i=1}^n \log \{ \Lambda(t_i, x_i, y_i) \} - \int_0^T \int \int_S \Lambda(t, x, y) dx dy dt, \quad (4.10)$$

y sus ecuaciones de máximos son

$$\sum_{i=1}^n \frac{\mu_t(t_i) \mu_w(t_i) \mu_s(x_i, y_i)}{\Lambda(t_i, x_i, y_i)} - U = 0, \quad (4.11)$$

$$\sum_{i=1}^n \frac{\sum_{j:t_j < t_i} g(t_j - t_i) h(x_j - x_i, y_j - y_i)}{\Lambda(t_i, x_i, y_i)} - G = 0, \quad (4.12)$$

donde

$$G = \sum_i \int_{t_i}^T \int \int_s g(t - t_i) h(x - x_i, y - y_i) dx dy dt, \quad (4.13)$$

$$U = \int_0^T \int \int_S \mu_t(t) \mu_w(t) \mu_s(x, y) dx dy dt.$$

Las ecuaciones de los máximos se resuelven mediante el siguiente sistema:

$$A^{(2)} = \frac{n - \sum_{i=1}^n \varphi_i}{G}, \quad (4.14)$$

$$\mu_0^{(2)} = \frac{n - A^{(2)}G}{U}, \quad (4.15)$$

donde n es el número de eventos, y φ_i es la probabilidad de background, dada por

$$\varphi_i = \frac{\mu_0 \mu_t(t_i) \mu_w(t_i) \mu_s(x_i, y_i)}{\mu_0 \mu_t(t_i) \mu_w(t_i) \mu_s(x_i, y_i) + A \sum_{j: t_j < t_i} g(t_j - t_i) h(x_j - x_i, y_j - y_i)}. \quad (4.16)$$

Como se puede observar en la expresión (4.16), para estimar φ_i es necesario conocer los coeficientes de relajación y los coeficientes también necesitan a φ_i , por lo que se trata de un procedimiento iterativo. Se asumen valores iniciales en μ_0 y A y se obtiene φ_i para todo i , dado que ya se han estimado los coeficientes μ_t , μ_w , μ_s , g y h ; se actualiza los coeficientes de relajación μ_0 y A , y se vuelve a recalcular φ_i hasta que se alcance la convergencia.

4.3.4. Suavizado y corrección del efecto de borde

Se suavizan las estimaciones para que los resultados de las reconstrucciones sean robustos y de este modo asegurar la convergencia. Para ello se utilizan funciones de núcleo.

Si \mathfrak{K} es el núcleo gaussiano dado por $\mathfrak{K}(x; \omega) = \frac{1}{\sqrt{2\pi\omega}} \exp\left(-\frac{x^2}{2\omega^2}\right)$, con ω el ancho del núcleo, $\lfloor x \rfloor$ representa el mayor entero no mayor que x , y h el ancho de banda, con h_x para el espacio y h_t para el tiempo, las ecuaciones de estimación suavizadas se convierten en:

$$\begin{aligned} \hat{\mu}_t(t) &\propto \sum_i w_i^{(t)} \mathfrak{K}(t - t_i; \omega_t) \\ \hat{\mu}_W(t) &\propto \sum_i w_i^{(W)} \sum_{k=0}^{\lfloor T/7 \rfloor} \mathfrak{K}(t - t_i + 7 \lfloor t_i/7 \rfloor - 7k; \omega_W), \\ \hat{\mu}_S(x, y) &\propto \sum_i \varphi_i \mathfrak{K}(x - x_i; \omega_x) \mathfrak{K}(y - y_i; \omega_y), \\ \hat{g}(t) &\propto \sum_{i,j} \rho_{ij} \mathfrak{K}(t - t_j + t_i; \omega_g), \\ \hat{h}(x, y) &\propto \sum_{i,j} \rho_{ij} \mathfrak{K}(x - x_j + x_i; \omega_{h_x}) \mathfrak{K}(y - y_j + y_i; \omega_{h_y}), \end{aligned}$$

4.3. METODOLOGÍA

Para corregir el efecto de borde, cada una toma la forma

$$\begin{aligned}\hat{\mu}(t) &\propto \sum_i w_i^{(t)} \frac{\mathfrak{K}(t - t_i; \omega_t)}{\int_0^T \mathfrak{K}(u - t_i; \omega_t) du}, \\ \hat{\mu}_W(t) &\propto \sum_i w_i^{(W)} \frac{\sum_{k=0}^{T/7} \mathfrak{K}(t - t_i + 7\lfloor t_i/7 \rfloor - 7k; \omega_W)}{\int_0^T \mathfrak{K}(u - t_i; \omega_W) du}, \\ \hat{\mu}_S(x, y) &\propto \sum_i \varphi_i \frac{\mathfrak{K}(x - x_i; \omega_x) \mathfrak{K}(y - y_i; \omega_y)}{\int \int_S \mathfrak{K}(u - x_i; \omega_x) \mathfrak{K}(v - y_i; \omega_y) dudv}, \\ \hat{g}(t) &\propto \frac{\sum_{i,j} \rho_{ij} \times \mathfrak{K}(t - t_j + t_i; \omega_g)}{\int_0^{T-t_i} \mathfrak{K}(u - t_j; \omega_g) du} \frac{1}{\sum_i I(t_i + t \leq T)},\end{aligned}\tag{4.17}$$

$$\hat{h}(x, y) \propto \frac{\sum_{i,j} \rho_{ij} \times \mathfrak{K}(x - x_j + x_i; \omega_{h_x}) \mathfrak{K}(y - y_j + y_i; \omega_{h_y})}{\int \int_S \mathfrak{K}(\mu - x_j + x_i; \omega_{h_x}) \mathfrak{K}(\mu - y_j + y_i; \omega_{h_y}) dx dy} \frac{1}{\sum_i I\{(x_i + x, y_i + y) \in S\}},\tag{4.18}$$

respectivamente. Los denominadores en las expresiones (4.17) y (4.18) son el número de veces que se observa el efecto triggering en el desfase temporal t o espacial (x, y) (Zhuang y Mateu, 2019).

4.3.5. Bondad de Ajuste

Para identificar la bondad de ajuste del modelo (análisis residual), se utilizan los tiempos transformados según la propuesta de Ogata (1988) y la reformulación de Zhuang y Mateu (2019):

$$t_i \rightarrow \tau_i = \int_0^{t_i} \int_S \Lambda(t, x, y) dx dy dt.\tag{4.19}$$

La expresión (4.19) transforma el proceso puntual $\mathcal{N} = \{(t_i, x_i, y_i), i = 1, 2, \dots, n\}$, en un proceso estacionario de Poisson con tasa unitaria (proceso de Poisson estándar) $\mathcal{N}' = \{\tau : i = 1, \dots, n\}$ denominada secuencia temporal transformada. Dado que $\Lambda(t, x, y)$ es desconocida en el análisis de datos reales, es posible sustituirla por $\hat{\Lambda}(t, x, y)$, que es una buena aproximación del modelo verdadero, y obtener una secuencia temporal transformada que es aproximadamente un proceso de Poisson estándar. Por tanto, si dicha secuencia no se desvía significativamente del proceso de Poisson estándar, se puede concluir que el modelo se ajusta bien a los datos.

Dadas las características de \mathcal{N}' , las bandas de confianza se pueden construir mediante los estadísticos relacionados con el proceso de Poisson (Zhuang y Mateu, 2019). La curva de frecuencia acumulativa $(\hat{\tau}_i = \int_0^{t_i} \int_S \hat{\Lambda}(t, x, y) dx dy dt, i)$ siempre conecta $(0, 0)$ y (T, n) , donde $\hat{\Lambda}(t, x, y)$ es el modelo estimado a partir de los datos en $[0, T]$ utilizando la estimación de máxima verosimilitud y $n = N([0, T] \times S)$. Para cada entero positivo k , si $k < n$, el intervalo de confianza para τ_k es el mismo que kZ , donde Z una variable aleatoria bajo una distribución beta con parámetro $(k + 1, n - k + 1)$; cuando $k > n$, τ_k puede aproximarse mediante una distribución gamma con parámetro de forma $k - n$ y parámetro de escala 1 (véase más detalles en Schoenberg, 2002).

4.4. Resultados

El análisis de datos se inicia con la presentación de la variable tiempo en forma continua mediante el uso de fechas julianas. Como se menciona en la sección 4.2, es importante identificar la unidad temporal según la granularidad de los datos. Al disponer de los registros diarios de los eventos delictivos, se realiza la reconstrucción de la periodicidad semanal y la tendencia a largo plazo. Como los delincuentes actúan según las actividades recurrentes de la localidad no se considera la periodicidad de la tasa de background.

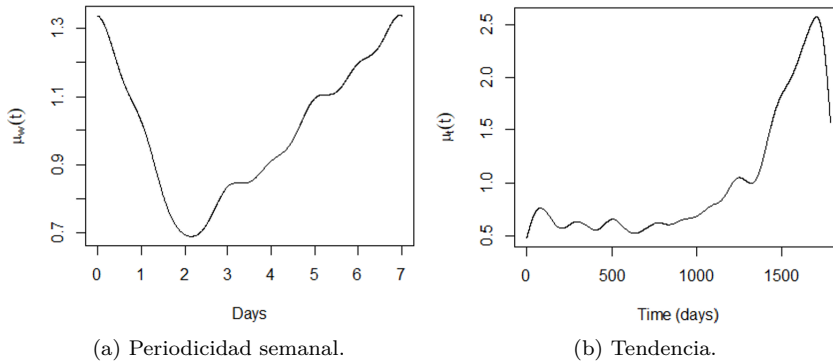


Figura 4.3: Periodicidad semanal y tendencia a largo plazo

En la Figura 4.3a se muestra la periodicidad semanal, en donde se observa un comportamiento estrictamente decreciente en domingo y lunes (domingo inicio de semana). A partir del martes un crecimiento lento, pero a medida que avanza al fin de semana su pendiente aumenta con mayor velocidad hasta alcanzar el valor máximo en el día sábado. Por la forma de la curva entre los

4.4. RESULTADOS

días se intuye que posiblemente exista algún patrón de delincuencia en función de las horas del día. El crecimiento acelerado de los eventos en los fines de semana (especialmente viernes y sábados) se cree que es un efecto debido al incremento de peatones que circulan por la ciudad, es decir, en esos días se presentan mayor número de personas en los lugares públicos y los delincuentes aprovechan para actuar.

En cuanto a la tendencia a largo plazo (véase la Figura 4.3b) es evidente que en el último año (2014) los delitos aumentan rápidamente, llegando a ser el año con mayor número de eventos registrados. Los años 2010 y 2012 presentan mayor variabilidad que el 2013. Los datos muestran un crecimiento delictivo lento hasta aproximadamente los últimos meses del 2013. Las políticas de seguridad que el gobierno ecuatoriano ha promovido a partir del 2010 al parecer se reflejan especialmente en el 2014, y como consecuencia, se incrementa el número de denuncias de forma acelerada (Castro et al., 2015).

La Figura 4.4 muestra las funciones de respuesta temporal y espacial. La función de respuesta temporal (véase la Figura 4.4a) indica que una vez que ha ocurrido un delito es probable que se desencadene otro evento en el tiempo t , mientras que la función de respuesta espacial (véase la Figura 4.4b) indica la distancia a la que posiblemente suceda dicho evento. Es decir, una vez que se ha producido un delito es probable que se desencadene otro suceso aproximadamente a un radio de 175m en los 3.5 días posteriores.

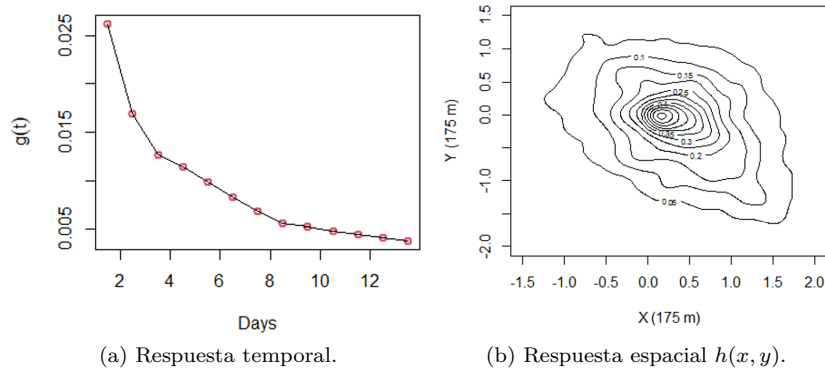
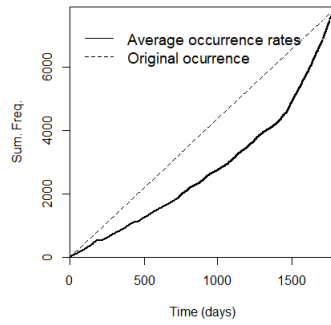


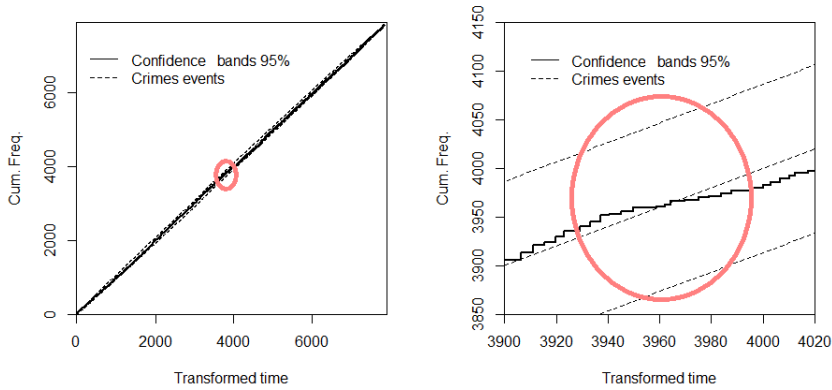
Figura 4.4: Funciones de respuesta temporal y espacial

Las componentes de background y triggering se estiman no paramétricamente utilizando el método de reconstrucción estocástica, y los coeficientes de relajación por el método de máxima verosimilitud. Por esta razón se conoce el método como semiparamétrico. Los valores de los parámetros de relajación que cuantifican la tasa de background y el efecto triggering son $\hat{\mu} = 0.027$ y $\hat{A} = 0.752$ con $\log(L) = -898.98$, en donde $\hat{A} = 0.75$ implica que el 75 % del

fenómeno delictivo de la ciudad de Riobamba puede explicarse por el efecto triggering, es decir, los delincuentes son reincidentes, y de alguna manera no son atrapados por el personal policial. Mientras que la tasa de background contribuye tan solo un 3%.



(a) Tiempo original (t)



(b) Tiempo transformado (τ)

Figura 4.5: Tasa de ocurrencia en el dominio del tiempo transformado

La Figura 4.5 muestra la frecuencia acumulada de los eventos delictivos y los cambios en la tasa de ocurrencia en el dominio del tiempo transformado (eventos delictivos en una secuencia estacionaria), donde los tiempos de ocurrencia originales no muestran cambios en el promedio de ocurrencias (véase la Figura 4.5a). En cuanto a los tiempos transformados se observan cambios alrededor de los tiempos 500, 1000, 2000, 3960, 4700 y en los últimos registros, que corresponden a los eventos delictivos de julio a diciembre de 2014

4.4. RESULTADOS

(véase la Figura 4.5b). El cambio más notable se produjo aproximadamente en el tiempo transformado 3960, por lo que se muestra una ampliación de éste a la derecha de la Figura 4.5b para identificarlo. El comportamiento delictivo en toda la ciudad muestra similitud en los primeros 4 años de datos, mientras que en el último año se produce un crecimiento acelerado.

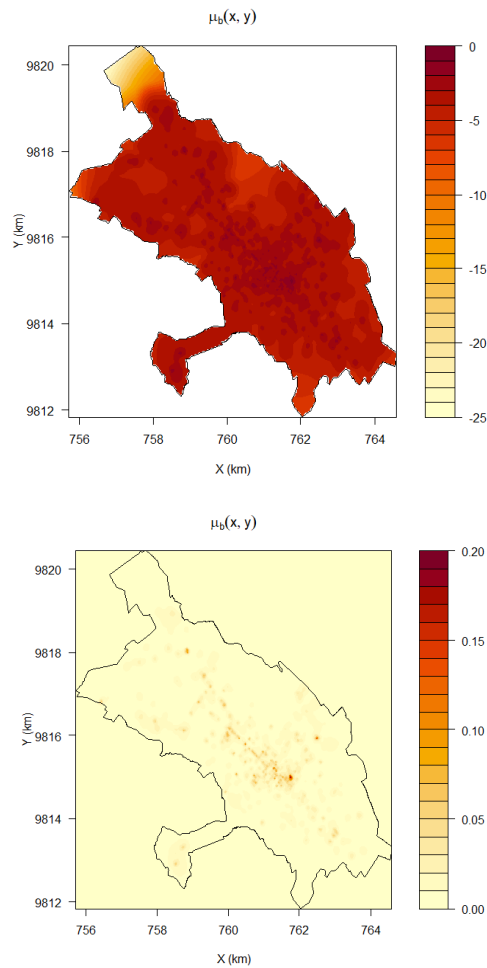


Figura 4.6: Tasa de background $\mu_s(x, y)$

La tasa de background $\mu_b(x, y)$ (véase la Figura 4.6) muestra valores elevados que se concentran en el centro de la ciudad (color rojo intenso). Una característica peculiar en la exponencial de la tasa de background es la visualización de la distribución de los eventos en forma de estructuras lineales

a lo largo de la ciudad. Este comportamiento promueve la idea de que las calles principales de la ciudad también podrían ser un factor de atracción para los delitos (véase la Figura 4.6), supuesto particular que sería relevante tomar en cuenta para futuros estudios.

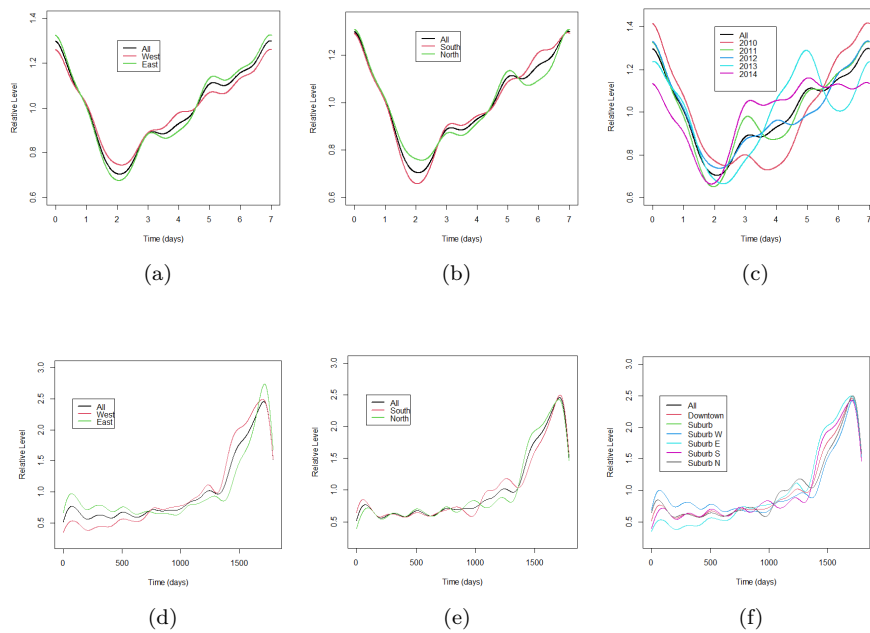


Figura 4.7: Periodicidad semanal y tendencia a largo plazo, clasificadas según los puntos cardinales y los años

Se aprovecha la reconstrucción de la periodicidad semanal y de la tendencia a largo plazo del background para identificar el comportamiento delictivo según los puntos cardinales de la ciudad y los cinco años de registros estudiados.

Su comportamiento semanal está estrechamente relacionado con las actividades comerciales, laborales y recreativas de la población. Por ejemplo, los días domingos y lunes muestran comportamientos planos, tanto en las periodicidades según los puntos cardinales (véase las Figuras 4.7a y 4.7b) como anuales (véase la Figura 4.7c); entre los posibles motivos, sería debido a que la mayoría de locales comerciales se mantienen cerrados, las personas tienden a mantenerse en casa junto a sus familiares y/o acuden a visitar las zonas recreativas familiares. A partir del día martes hasta su valor máximo en sábados el fenómeno delictivo crece de forma acelerada y con mayor variabilidad, especialmente entre años.

4.4. RESULTADOS

Con respecto a la tendencia a largo plazo, al Oeste y al Norte de la ciudad la frecuencia delictiva es predominante (véase las Figuras 4.7d y 4.7e, respectivamente), pero con mayor variabilidad entre el Este y el Oeste. De alguna manera, en los dos primeros años, en el Norte y el Sur los eventos son menos variables hasta aproximadamente los últimos meses del 2013; a partir de ahí la variabilidad aumenta hasta unos meses antes de finalizar el 2014, ya que vuelven a ser similares hasta su culminación (véase la Figura 4.7f).

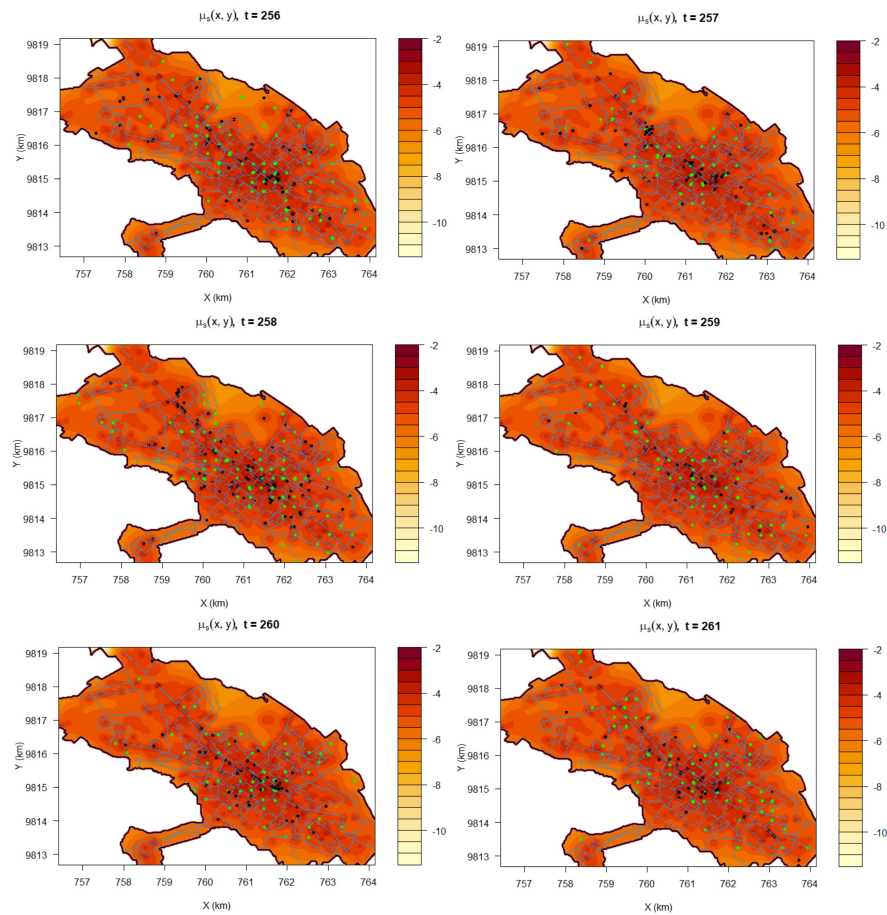


Figura 4.8: Eventos delictivos reales (negro) y predicciones (verde) para las seis semanas, sobre el background $\mu_s(x, y)$

Es interesante visualizar las posibles ubicaciones de las predicciones delictivas para las seis semanas consecutivas al conjunto de datos analizado, los cuales se obtienen mediante simulaciones basadas en las intensidades de las

predicciones obtenidas con el proceso de Cox log-gaussiano (véase el capítulo 3) sobre la tasa de background resultante del SEHP (véase la Figura 4.8). Las líneas azules muestran los límites de las divisiones de los sectores o barrios dentro del perímetro de la ciudad, los puntos negros corresponden a las ubicaciones de los delitos reales registrados y los verdes a las predicciones. Se puede observar que las predicciones no están aleatoriamente alejadas de los eventos reales, y en su mayoría están sobre las áreas sombreadas del background.

4.5. Discusión y conclusiones

La *teoría del patrón delictivo* menciona que la delincuencia es un fenómeno que tiene un comportamiento basado en patrones. Es decir, tanto en el espacio como en el tiempo los delitos no se distribuyen de forma aleatoria ni uniforme, dando lugar a la existencia de sectores concretos que concentran gran parte de las infracciones, así como también reincidencia de sujetos que cometen múltiples delitos (Higgins y Swartz, 2018).

Son muchos los métodos estadísticos empleados con el propósito de describir los patrones delictivos; sin embargo, en su mayoría llegan a identificar los sectores con alta frecuencia delictiva, pero no cuantifican la reincidencia. Los modelos de procesos puntuales espacio-temporales auto-excitados de Hawkes poseen características relevantes que permiten predecir la tasa de eventos en función del espacio, el tiempo y el comportamiento histórico de eventos anteriores, dando lugar a una cuantificación de la reincidencia mediante la estimación del efecto desencadenante (triggering).

En este capítulo se ha realizado la modelización de los datos de eventos delictivos registrados durante 2010–2014 en la ciudad de Riobamba. El análisis se inicia con la descripción del comportamiento de los delitos mediante gráficos de colores degradados para visualizar la evolución de este fenómeno a través del tiempo. La estimación paramétrica mediante la reconstrucción estocástica de la periodicidad semanal y la tendencia temporal a largo plazo facilita la identificación del comportamiento delictivo desde el punto de vista de los diferentes días de la semana, punto cardinales y por años, así como también la obtención de las componentes de background y triggering. La estimación de los parámetros de relajación por máxima verosimilitud facilita el algoritmo iterativo para la estimación de los parámetros de relajación, es decir, la cuantificación del efecto de reincidencia.

Zhuang y Mateu (2019) estabilizan su modelizado con $\hat{\mu} = 0.77$ y $\hat{A} = 0.029$ en datos de robos, con un valor de log-verosimilitud $\log(L) = -920.81$. Esto indica que aproximadamente un 3% de los robos desencadenan otro evento (reincidencia), es decir, el proceso está dominado por el background. Mientras que en la ciudad de Riobamba sucede lo contrario, ya que se obtiene un efecto

4.5. DISCUSIÓN Y CONCLUSIONES

triggering de $\hat{A} = 0.752$ y background de $\hat{\mu} = 0.027$, con un valor de log-verosimilitud $\log(L) = -898,98$. Esto significa que el efecto triggering domina el proceso, ya que puede explicar el 75% de los datos. Es decir, los eventos son cometidos repetitivamente por los mismos delincuentes (reincidencias de sujetos), posiblemente debido a la falta de seguimiento policial a los diferentes casos, no denuncia por temor a las represalias, o, de alguna manera, los sujetos consiguen evadir el ámbito legal de la localidad. En cuanto al background, tan solo contribuye aproximadamente en un 3%.

Las funciones de respuesta espacial y temporal reconstruidas en la componente de background y triggering implican que, una vez que se ha producido un delito, es probable que desencadene otro en aproximadamente los 3.5 días siguientes en un radio aproximado de 175m de distancia. Estos resultados son confiables debido a que los tiempos transformados que han permitido identificar el cambio en la tasa de ocurrencia del background a largo plazo se mantienen dentro del intervalo de confianza. La tasa de background ha permitido visualizar el efecto de agrupación en diferentes áreas de la ciudad, especialmente en las áreas con mayor movimiento comercial, laboral y recreativo.

La simulación de eventos en base a las intensidades predichas con el proceso de Cox log-gaussiano (LGCP) permite visualizar los eventos reales y las predicciones sobre el background estimado por el SEHP. Es decir, en forma conjunta, proporciona información más completa sobre el fenómeno delictivo. Sus estructuras estocásticas permiten caracterizar el movimiento delictivo en forma diaria, predicciones semanales, tendencia a largo plazo y patrones de comportamiento en la reincidencia de eventos, tanto en el tiempo como en el espacio, proporcionando resultados de gran interés para la sociedad en general.

Una desventaja que se ha considerado especialmente relevante en los SEHP es el coste computacional, puesto que el algoritmo iterativo requiere de un tiempo considerable para alcanzar la convergencia.

5. Medidas de riesgo en LGCP bivariantes

5.1. Introducción

La delincuencia es un fenómeno social negativo y una realidad en la mayoría de los países del mundo, por ello los esfuerzos se concentran en la búsqueda de estrategias para su control y/o prevención.

El fenómeno delictivo tiene un comportamiento basado en patrones, dando lugar a la existencia de sectores concretos que concentran gran parte de las infracciones (Higgins y Swartz, 2018). El desafío es identificar aquellos sectores que condensan anomalías delictivas y cuantificar la peligrosidad en términos de probabilidad o riesgo. La tarea es compleja debido a la dificultad para predecir válidamente conductas humanas. Sin embargo, los modelos de procesos de Cox log-gaussianos (LGCP) muestran estructuras estocásticas que se aproximan a capturar patrones de comportamiento en las actividades humanas, haciendo de estas buenas herramientas que proporcionan estimaciones confiables para el fenómeno delictivo. Un primer acercamiento a la identificación de eventos extremos en modelos LGCP univariantes lo muestran Diggle et al. (2005), Rodrigues y Diggle (2012), Taylor et al. (2013) y González y Mateu (2021), en donde identifican la relación entre el riesgo actual y el esperado que supera un umbral estipulado, pero no hacen referencia al caso bivalente ni multivariante.

Las medidas de riesgo VaR (Value-at-Risk) y ES (Expected Shortfall) son una posible forma de cuantificar el riesgo, y tanto la EVT (Extreme Value Theory) como las cópulas contribuyen a mejorar la precisión de su estimación (Hsu et al., 2012; Karmakar, 2017; Yu et al., 2018; Manosalva, 2020).

En este contexto, este capítulo se concentra, en primer lugar, en la modelización de las componentes dentro del flujo y entre flujos de los hurtos y robos a personas, permitiendo descomponer la variación espacial de los dos tipos de eventos en una variación asociada a un tipo en particular. Se incluyen efectos de covariables espaciales y estimación de los parámetros del campo latente con enfoque bayesiano. Y en segundo lugar, en la estimación de las medidas de riesgo VaR y ES dada una estructura de dependencia mediante la cópula extrema de Gumbel-Hougaard (Salvadori et al., 2007), cuyas marginales se

refieren a eventos que exceden un umbral dado, siendo aproximados mediante la distribución de Pareto Generalizada (GPD) (Castellanos y Cabras, 2007; Abad et al., 2014).

La EVT se enfoca a formas de analizar aquellos fenómenos con comportamientos fuera del patrón común. En particular, para el caso bivalente se puede abordar mediante cópulas (Salvadori et al., 2007). Esto es debido a que toda distribución multivalente admite una representación en términos de una cópula y un conjunto de distribuciones marginales (Joe, 2015). El propósito de utilizar cópulas es encontrar familias paramétricas flexibles para ser utilizadas como descriptores de las estructuras de dependencia evidentes en los datos (Joe, 2015; Henderson et al., 2021). Las cópulas permiten simular fácilmente distribuciones bivariantes y analizar su estructura. La representación de Sklar (1959) sugiere que la construcción de la distribución conjunta de la cópula sea modelizada mediante la estimación de las funciones de distribución acumulativas marginales y la estimación de la cópula de forma independiente.

Para la estimación de las funciones de distribución acumulativas marginales, De Haan (1970) y los resultados de inferencia estadística propuestos por Pickands (1975) motivaron el desarrollo de modelos basados en excedencias de umbrales. Esto hizo que la EVT evolucionara desde los enfoques de máximos por bloques y de pico sobre un umbral (POT).

En la actualidad el método POT es uno de los más utilizados, debido a que evita el problema de pérdida de datos. Sin embargo, la forma de seleccionar un umbral adecuado es compleja, dado que la estimación de valores extremos depende del umbral. Si el umbral es demasiado bajo, la cola no alcanza el criterio de convergencia, lo que provocará un gran sesgo, y si el umbral es demasiado alto, se tendrán pocos datos, dando lugar a una alta varianza y, como consecuencia resultados poco fiables. Por lo tanto, la elección de un umbral adecuado implica un equilibrio entre el sesgo y la varianza (Yang et al., 2018). Entre los métodos más comunes para la elección del umbral está el basado en la esperanza de la GPD, que da origen al gráfico de vida residual media. La estimación de los parámetros de la GPD por máxima verosimilitud ha sido considerada por Smith (1985) y se han encontrado varias inconsistencias debido al tamaño de las muestras (Hosking y Wallis, 1987; Grimshaw, 1993) y estimaciones del parámetro de forma fuera del intervalo $(-1/2, 1/2)$ (Gómez, 2008). En consecuencia, la inferencia bayesiana es una buena alternativa, especialmente con a priori de Jeffreys debido a que proporcionan buenos resultados (Castellanos y Cabras, 2007).

Para la modelización del parámetro de dependencia de la cópula se suele seleccionar basándose en la estimación de máxima verosimilitud. Como resultado, se espera que el parámetro de la cópula extrema de Gumbel-Hougaard capture la dependencia localizada entre los dos patrones de delitos. Es importante tener en cuenta que un volumen pequeño de datos no proporciona suficientes observaciones para formar una estructura de dependencia distintiva (Henderson et al., 2021).

5.2. DESCRIPCIÓN DE LOS DATOS

Las medidas de riesgo VaR desde el punto de vista delictivo se refieren a la probabilidad de identificar delitos iguales o que excedan valores predeterminados, para una localización dada, es decir, cualquier anomalía que justifique algún tipo de intervención policial sin previo aviso. El periodo de tiempo y el nivel de confianza (el cuantil) son los dos parámetros principales que deben elegirse de forma adecuada al objetivo general de la medición del riesgo. El cuantil suele ser muy pequeño; por ejemplo, el 1% de los peores resultados. Sin embargo, se puede usar el valor que estipule el modelo de gestión de riesgo interno de la oficina de policía o ley regulatoria del país para controlar la exposición al riesgo; la cifra típica es de alrededor del 5%.

En este contexto, se propone un procedimiento para identificar excedencias espaciales y sus regiones asociadas, así como también cuantificar la peligrosidad en términos del VaR y ES localizado en patrones puntuales bivariantes modelizados con LGCP. Para ello se consideran los delitos de hurtos (sustracción de alguna cosa sin empleo de fuerza o violencia) y robos a personas (sustracción de algo con empleo de fuerza, violencia o intimidación a las personas) registrados en la ciudad de Riobamba-Ecuador.

5.2. Descripción de los datos

Se consideran los datos de eventos delictivos registrados durante el año 2014 en la ciudad de Riobamba-Ecuador. Se dispone de múltiples tipos de delitos; sin embargo, se toman los datos de hurto y robo a personas, eventos que comúnmente son tratados como si fuesen un solo tipo de delitos. En la Figura 5.1 se observa un mayor número de registros de hurtos (707 puntos verdes) que robos a personas (350 puntos celestes). Ambos patrones puntuales muestran tendencia a agruparse en el centro de la ciudad.

Como referencias importantes para la modelización mediante LGCP se tienen en cuenta algunas covariables; en particular, se toman las distancias más cortas entre cada evento y las unidades de policía comunitaria (UPC), las cámaras de vigilancia instaladas por toda la ciudad (ECU911, 2020) y las áreas con mayor tráfico vehicular y peatonal (parques, mercados, plazas, supermercados, hospitales y cajeros automáticos).

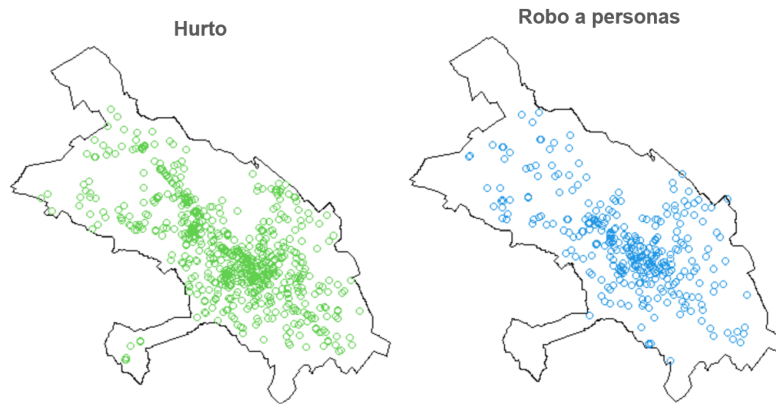


Figura 5.1: Hurtos y robos a personas registrados durante 2014 en la ciudad de Riobamba

Se toman aquellas referencias como puntos estratégicos, debido a que la delincuencia se mueve en función de las actividades diarias de la población (Zhuang y Mateu, 2019), además del interés de capturar la mayor variabilidad posible.

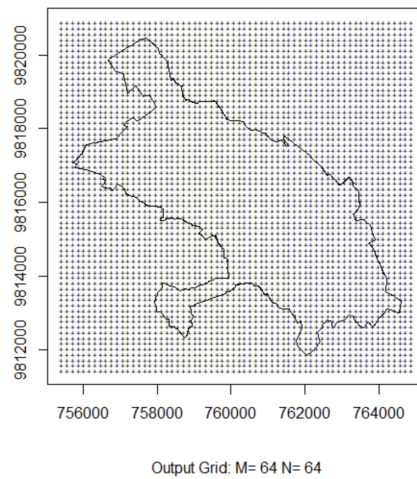


Figura 5.2: Malla computacional

Mediante una malla computacional de dimensiones 64×64 , con un ancho de celda de 150m (véase la Figura 5.2), se construye una superposición de polígono a polígono para inferir los valores de las covariables en la malla (véase la Figura 5.3).

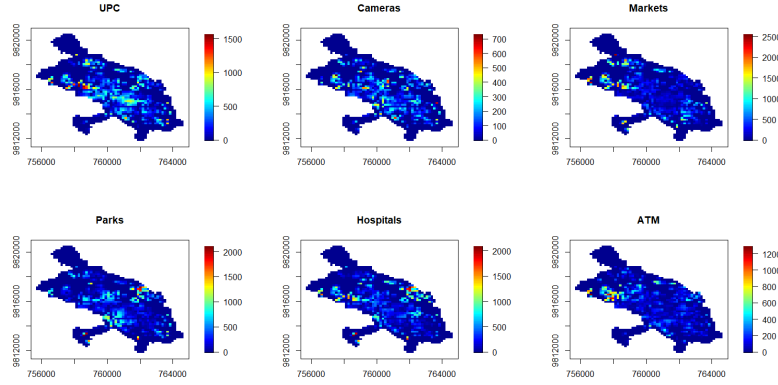


Figura 5.3: Covariables espaciales ajustadas a la malla computacional determinada

La Figura 5.3 muestra en color degradado las distancias entre los eventos registrados en cada localización y los puntos de referencia. El color rojo indica mayor distancia y el color azul las más pequeñas. La covariable *mercados* agrupa a centros comerciales, plazas y mercados, por sus características comerciales algo semejantes.

5.3. Metodología

5.3.1. Modelo LGCP bivalente

Un proceso de Cox con intensidad log-gaussiana multivariante para un patrón de puntos multitipo tiene la forma

$$Z_i(\mathbf{s}) \sim \text{Poisson} \{ \Lambda_i(\mathbf{s}) \} \quad (5.1)$$

$$\Lambda_i(\mathbf{s}) = C_A \lambda_i(\mathbf{s}) \exp \{ X(\mathbf{s})_i \beta_i + \mathcal{Y}_i(\mathbf{s}) + \mathcal{Y}_{k+1}(\mathbf{s}) \}, \text{ con } i = 1, 2, \dots, k, \quad (5.2)$$

donde $Z_i(s)$ es el número de eventos de tipo i en la celda de la malla computacional que contiene al punto \mathbf{s} , k representa los tipos de eventos, C_A es el área de celda en la malla, $\lambda(\mathbf{s})$ actúa como un offset, $X(\mathbf{s})$ es un vector de valores de las covariables con efectos asociados β , e $\mathcal{Y}_i(\mathbf{s})$ con $i = 1, \dots, k + 1$ son procesos gaussianos latentes con parámetros en escala transformada $\eta_i \{ \log(\sigma_i), \log(\phi_i) \}$ (Taylor et al., 2015).

Como se modelizan dos tipos de eventos, $k = 2$, las intensidades estarán dadas de la forma

$$\Lambda_1(\mathbf{s}) = C_A \lambda(\mathbf{s}) \exp\{X(\mathbf{s})_1 \beta_1 + \mathcal{Y}_1(\mathbf{s}) + \mathcal{Y}_3(\mathbf{s})\} \quad (5.3)$$

y

$$\Lambda_2(\mathbf{s}) = C_A \lambda(\mathbf{s}) \exp\{X(\mathbf{s})_2 \beta_2 + \mathcal{Y}_2(\mathbf{s}) + \mathcal{Y}_3(\mathbf{s})\}. \quad (5.4)$$

El modelo para cada uno de los dos flujos de eventos es un proceso de Cox log-gaussiano, cuya parte estocástica está compuesta de dos elementos, una componente dentro del flujo y otra entre los flujos. La idea principal es descomponer la variación espacial de los dos tipos de eventos en una variación asociada a un tipo particular, \mathcal{Y}_i con $i = 1, 2$, y una común \mathcal{Y}_{k+1} . Así, aunque \mathcal{Y}_1 e \mathcal{Y}_2 muestren un patrón espacial individual, el proceso \mathcal{Y}_3 capta las áreas de alta o baja intensidad que son comunes a los dos tipos.

La estimación de los parámetros se realiza mediante aproximaciones, debido a que el método de contraste mínimo no es capaz de identificar parámetros en un componente aleatorio específico, posiblemente debido al bajo número de recuentos en los datos originales (Taylor et al., 2015). Las aproximaciones obtenidas se refieren a los procesos estocásticos $\mathcal{Y}_1 + \mathcal{Y}_3$ e $\mathcal{Y}_2 + \mathcal{Y}_3$ y la función de correlación es exponencial e incorrelada:

$$\begin{aligned} \text{Var}[\mathcal{Y}_1 + \mathcal{Y}_3] &= \text{Var}[\mathcal{Y}_1] + \text{Var}[\mathcal{Y}_3] \\ &= \sigma_1^2 \exp\left\{\frac{-r}{\phi_1}\right\} + \sigma_3^2 \exp\left\{\frac{-r}{\phi_3}\right\} \\ &= 2\sigma^2 \exp\left\{\frac{-r}{\phi}\right\}; \end{aligned} \quad (5.5)$$

siguiendo el mismo razonamiento para \mathcal{Y}_2 incorrelada con \mathcal{Y}_3 , se tiene

$$\text{Var}[\mathcal{Y}_2 + \mathcal{Y}_3] = \sigma_2^2 \exp\left\{\frac{-r}{\phi_2}\right\} + \sigma_3^2 \exp\left\{\frac{-r}{\phi_3}\right\} = 2\sigma^2 \exp\left\{\frac{-r}{\phi}\right\}, \quad (5.6)$$

bajo el supuesto de que r es la distancia euclídea, y por ejemplo, $\sigma = \sigma_1 = \sigma_3$ y $\phi = \phi_2 = \phi_3$. Esto implica que la varianza de \mathcal{Y}_1 o \mathcal{Y}_3 es aproximadamente la mitad de la varianza estimada y su parámetro de escala es aproximadamente igual al parámetro de escala estimado (Taylor et al., 2015).

5.3.2. Estimación de parámetros para el LGCP bivariante

Para estimar los parámetros del modelo LGCP se utiliza un enfoque bayesiano, debido a que proporciona un procedimiento de inferencia flexible y a menudo soluciones prácticas, especialmente para la inferencia que surge de

5.3. METODOLOGÍA

modelos estadísticos complejos. Además, no hace ninguna distinción formal entre la estimación de β , η y la predicción de Y , y de este modo incorpora naturalmente la incertidumbre de los parámetros en la inferencia predictiva.

Se definen las probabilidades a priori gaussianas multivariantes para β y una gaussiana multivariante en la escala logarítmica para los parámetros positivos σ y ϕ , es decir, $\beta \sim N(\mu_\beta, \Sigma_\beta)$ y $\eta = \{\log(\sigma), \log(\phi)\} \sim N(\mu_\eta, \Sigma_\eta)$.

La muestra $\{\beta^{(j)}, \eta^{(j)}, Y^{(j)}\}_{j=1}^N \sim \pi(\beta, \eta, Y|Z)$, que en si misma ya es compleja, se obtiene mediante los métodos Gibbs y MALA.

La idea es simular la distribución estacionaria de $\pi(\beta, \eta, Y|Z)$ a partir de una cadena de Markov. Una vez inicializada la cadena en 0, $\{\beta^{(0)}, \eta^{(0)}, Y^{(0)}\}$, el i -ésimo paso del algoritmo consiste en extraer un candidato $\{\beta^*, \eta^*, Y^*\}$ para la densidad propuesta $q(\beta^*, \eta^*, Y^*|\beta^{(i-1)}, \eta^{(i-1)}, Y^{(i-1)})$ y aceptarlo, es decir, estableciendo $\{\beta^{(i)}, \eta^{(i)}, Y^{(i)} = \beta^*, \eta^*, Y^*\}$, con probabilidad

$$\min \left\{ 1, \frac{\pi(\beta^*, \eta^*, Y^*|Z)}{\pi(\beta^{(i-1)}, \eta^{(i-1)}, Y^{(i-1)}|Z)} \frac{q(\beta^{(i-1)}, \eta^{(i-1)}, Y^{(i-1)}|\beta^*, \eta^*, Y^*)}{q(\beta^*, \eta^*, Y^*|\beta^{(i-1)}, \eta^{(i-1)}, Y^{(i-1)})} \right\}. \quad (5.7)$$

El diseño de q es una mezcla de paseo aleatorio y núcleos de la propuesta de Langevin. Se usa MALA para aprovechar la información del gradiente en $\pi(\beta, \eta, Y|Z)$. Se trabaja con una transformada de Y , específicamente Γ , donde $Y = \sum_n^{1/2} \Gamma + \mu\eta$, y un subíndice η que denota la dependencia de los parámetros del proceso latente.

Sea $\zeta = \{\beta, \eta, \Gamma\}$ un MALA completo. Se utiliza el núcleo

$$q(\zeta^{(i^*)}|\zeta^{(i-1)}) = N \left[\zeta^{i^*}; \zeta^{(i-1)} + \frac{h^2}{2} \Sigma_{opt} \nabla \log \left\{ \pi(\zeta^{(i-1)}|Z) \right\}, h^2 \Sigma_{opt} \right], \quad (5.8)$$

donde h es una constante de escala. La elección ideal para Σ_{opt} sería la inversa de la matriz de información de Fisher, pero debido a su alta dimensionalidad es preferible utilizar una aproximación a la inversa negativa de dicha matriz (véase detalles en [Taylor et al., 2015](#)).

Valores extremos bivariantes

Sean Z_{1_s} y Z_{2_s} dos variables aleatorias de excesos con función de distribución conjunta H_s y marginales continuas F_{1_s} y F_{2_s} ; entonces

$$H_s = \mathcal{C}(F_{1_s}^{-1}(u), F_{2_s}^{-1}(v)), \quad (5.9)$$

en términos de una única función cópula $\mathcal{C} : [0, 1]^d \rightarrow [0, 1]$.

5.3.3. Estimación de las funciones de distribución acumulativas marginales

El enfoque POT utiliza los datos de forma eficiente, centrándose en las localizaciones \mathbf{s} cuyas realizaciones superan un umbral c . Según el teorema de Pickands (Pickands, 1975) $Z_{\mathbf{s}} = Z_{\mathbf{s}} - c$ ($Z_{\mathbf{s}}$ recuentos de delitos en cada localización \mathbf{s}) tienen aproximadamente distribución de Pareto generalizada (GPD) con parámetros de forma $\xi(\mathbf{s})$ y de escala $\sigma(\mathbf{s})$ para cada localización \mathbf{s} y por cada tipo de delitos.

La literatura menciona que existen varios métodos para estimar los parámetros de la GPD, siendo el MLE la más común; sin embargo, debido a sus deficiencias y limitaciones descritas en la introducción de este capítulo, se adopta un enfoque bayesiano.

Para cada localización \mathbf{s} de los dos tipos de eventos, se definen las a priori de Jeffreys (invariantes bajo reparametrización), dadas por

$$\pi(\xi(\mathbf{s}), \sigma(\mathbf{s})) \propto \sigma^{-1}(\mathbf{s})(1 + \xi(\mathbf{s}))^{-1}(1 + 2\xi(\mathbf{s}))^{-1/2}, \quad (5.10)$$

con $\xi(\mathbf{s}) > -0.5$, y $\sigma(\mathbf{s}) > 0$.

La expresión (5.10) se obtiene a partir de la matriz de covarianzas límite, dada por la inversa de la matriz de información de Fisher (Davison y Smith, 1990), de la forma

$$\frac{1 - \xi(\mathbf{s})}{N} \begin{pmatrix} 2\sigma^2(\mathbf{s}) & \sigma(\mathbf{s}) \\ \sigma(\mathbf{s}) & 1 - \xi(\mathbf{s}) \end{pmatrix}, \quad \xi(\mathbf{s}) < 0.5. \quad (5.11)$$

Se puede ver que para $\xi(\mathbf{s}) = 0$ le corresponde la a priori de Jeffreys $\pi(\xi(\mathbf{s}), \sigma(\mathbf{s}))$ para el parámetro de escala de la distribución exponencial, bajo el supuesto de que $\xi(\mathbf{s})$ y $\sigma(\mathbf{s})$ tienen a priori independientes y que sólo la distribución marginal de $\sigma(\mathbf{s})$ es impropia (véase ejemplos de simulación de a priori de Jeffreys en Castellanos y Cabras, 2007).

La densidad a posteriori no está disponible en forma fija para $n > 1$, por lo que se aproxima numéricamente. Para esto se utiliza MCMC, concretamente un algoritmo de Metrópolis-Hasting dentro de un muestreo de Gibbs; el núcleo está dado por

$$\begin{aligned} \pi(\xi(\mathbf{s}), \sigma(\mathbf{s}) | Z(\mathbf{s})) &\propto \sigma^{-(n+1)}(\mathbf{s})(1 + \xi(\mathbf{s}))^{-1}(1 + 2\xi(\mathbf{s}))^{-1/2} \\ &\times \prod_{i=1}^n \left(1 + \xi(\mathbf{s}) \frac{z_i(\mathbf{s})}{\sigma(\mathbf{s})} \right)^{-\frac{1+\xi(\mathbf{s})}{\xi(\mathbf{s})}}, \end{aligned} \quad (5.12)$$

con $\xi(\mathbf{s}) > -0.5$, $\frac{\xi(\mathbf{s})}{\sigma(\mathbf{s})} > -1/\max(z_1(\mathbf{s}), \dots, z_n(\mathbf{s}))$ y $\sigma(\mathbf{s}) > 0$.

5.3.4. Estimación del parámetro de dependencia

Se toma la familia de cópulas de Gumbel-Hougaard (GHC), puesto que satisfacen la definición de cópula de valor extremo (EVC) (véase [Salvadori et al., 2007](#)); además es específica para colas derechas, que es lo que se requiere para los extremos delictivos. Está parametrizada por un $\theta \geq 1$ real y se define como

$$\mathcal{C}_{\mathbf{s}}^{GH}(\mathbf{u}) = \exp\left(-\left[\sum_{i=1}^d \{-\ln(u_i)\}^\theta\right]^{1/\theta}\right), \quad \mathbf{u} \in [0, 1]^d. \quad (5.13)$$

El primer paso práctico en la construcción de un modelo para cópulas es comprobar si los \mathcal{Z} son mutuamente independientes, es decir, si pueden considerarse como una muestra aleatoria simple de $H(\mathbf{s})$. Como el procedimiento de la simulación realizada ya garantiza la condición de que \mathcal{Z} sean i.i.d., no se realizan las comprobaciones; sin embargo, puede consultarse sobre los test de independencia y bondad de ajuste de cópulas en [Kojadinovic y Yan \(2010\)](#).

Siguiendo los razonamientos de [Kojadinovic y Yan \(2010\)](#), [Kong et al. \(2015\)](#), si $H(\mathbf{s})$ puede ser interpretada por una distribución exponencial bivalente, se puede definir un parámetro $\theta \geq 0$ en cada localización \mathbf{s} para modelizar la estructura de dependencia entre $\mathcal{Z}_1(\mathbf{s})$ y $\mathcal{Z}_2(\mathbf{s})$. Se utiliza el teorema de la transformación integral para obtener $u = F_{\mathcal{Z}_s}(z)$ y $v = F_{\mathcal{Z}_s}(z)$, las marginales normalizadas, de forma que u y v de cada localización \mathbf{s} oscilen en el intervalo $[0, 1]$. Con las variables transformadas, aplicando el teorema de Sklar, es posible expresar la distribución conjunta de $\mathcal{Z}_1(\mathbf{s})$ y $\mathcal{Z}_2(\mathbf{s})$ como una función de GHC de la siguiente forma:

$$\mathcal{C}_{\mathbf{s}}^{GH}(u, v) = \exp\left(-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{1/\theta}\right) \quad (5.14)$$

y función de densidad de probabilidad

$$\frac{\partial \mathcal{C}_{\mathbf{s}}^{GH}}{\partial u \partial v} = \frac{(-\ln u \ln v)^{\theta-1}}{uv} \exp(\wp^{1/\theta}) \left[\wp^{2/(\theta-2)} + (\theta-1)\wp^{1/(\theta-2)} \right], \quad (5.15)$$

donde $\wp = (-\ln u)^\theta + (-\ln v)^\theta$, siendo θ el parámetro desconocido a estimar en cada localización \mathbf{s} .

Considerando un conjunto de observaciones (z_{1i}, z_{2j}) , $i, j = 1, 2, \dots, n$, en cada localización \mathbf{s} , con distribuciones marginales normalizadas $u = F_{\mathcal{Z}_1}(z)$, $v = F_{\mathcal{Z}_2}(z)$ y una estructura de dependencia que se puede modelizar mediante una función de GHC a través de un parámetro θ , las funciones de log-verosimilitud vienen dadas por

$$\begin{aligned} \ln \mathcal{C}_{\mathbf{s}}^{GH}(\theta|u, v) &= -\ln uv + (\theta + 1) \ln(\ln u \ln v) - \wp^{1/\theta} + \\ &+ \ln \left[\wp^{2/(\theta-2)} + (\theta-1)\wp^{1/(\theta-2)} \right], \end{aligned} \quad (5.16)$$

siendo

$$\hat{\theta} = \operatorname{argmax} \sum_{i=1}^n \ln \mathcal{C}_{\mathbf{s}}^{GH}(\theta|u, v). \quad (5.17)$$

Periodo de retorno

El periodo de retorno de un evento proporciona un medio sencillo para el análisis de riesgos. Normalmente, se define como *el tiempo medio que transcurre entre dos realizaciones sucesivas del evento dado* (véase más detalles en [Salvadori, 2004](#)). Para el caso bivalente, se considera para cada localización \mathbf{s} una secuencia $\varepsilon = E_1, E_2, \dots$ de eventos independientes. Cada suceso E_i se caracteriza por el comportamiento conjunto de las variables aleatorias \mathcal{Z}_1 y \mathcal{Z}_2 , teniendo cópula $\mathcal{C}_{\mathbf{s}}^{GH}$ y ley conjunta $F = \mathcal{C}_{\mathbf{s}}^{GH}(F_1, F_2)$. Se puede suponer que los sucesos ocurren en los instantes $t_1 < t_2 < \dots$, es decir, se utiliza como modelo un proceso temporal de puntos marcados. Una definición muy general del periodo de retorno T del suceso (medible) E es la siguiente:

$$T = \vartheta(\mathbb{P}(E)) = \frac{\mu}{\mathbb{P}(E)}, \quad (5.18)$$

donde μ es el tiempo medio de llegada de los sucesos de la secuencia ε ; por ejemplo, considerando los máximos anuales, $\mu = 1$ año. El periodo de retorno T es una función $\vartheta : \mathcal{F} \rightarrow [\mu, +\infty)$ que depende de la medida de probabilidad inducida por F , donde \mathcal{F} es la σ -álgebra de los sucesos medibles del espacio de probabilidad considerado. En el caso univariante, el suceso de interés es $E = [\mathcal{Z} > c]$, para algún umbral c dado.

Para el caso bivalente, dada la cópula $\mathcal{C}_{\mathbf{s}}^{GH}$ y un nivel de probabilidad $t \in I$, un enfoque directo implica el uso de la medida de Kendall $K_{\mathcal{C}_{\mathbf{s}}^{GH}}(t)$ de la región

$$\mathbf{B}_{\mathcal{C}_{\mathbf{s}}^{GH}}(t) = \{(u, v) \in I^2 : \mathcal{C}_{\mathbf{s}}^{GH}(u, v) \leq t\}, \quad (5.19)$$

en donde $\mathbf{B}_{\mathcal{C}_{\mathbf{s}}^{GH}}(t)$ es la parte del cuadrado unitario que se encuentra sobre la isolínea de $\mathcal{C}_{\mathbf{s}}^{GH}$ con un nivel de t , o por debajo y a la izquierda de ésta. Si $\mathcal{C}_{\mathbf{s}}^{GH}$ es MEV, la medida de Kendall de $\mathbf{B}_{\mathcal{C}_{\mathbf{s}}^{GH}}(t)$ viene dada por

$$K_{\mathcal{C}_{\mathbf{s}}^{GH}}(t) = t - (1 - \tau_{\mathcal{C}_{\mathbf{s}}^{GH}})t \ln t, \quad (5.20)$$

donde $\tau_{\mathcal{C}_{\mathbf{s}}^{GH}}$ es el valor de la τ de Kendall asociada a la cópula $\mathcal{C}_{\mathbf{s}}^{GH}$. Claramente, las cópulas MEV con el mismo valor de τ comparten la misma función $K_{\mathcal{C}_{\mathbf{s}}^{GH}}$. Ahora, si $\tilde{p} \in I$ ($\tilde{p} = F_{\mathcal{Z}}^{-1}(\tilde{c})$ y \tilde{c} umbral crítico fijado) es un nivel de probabilidad crítico arbitrario, entonces la isolínea de $\mathcal{C}_{\mathbf{s}}^{GH}$ que tiene el nivel \tilde{p} puede utilizarse para definir dos regiones complementarias, la de los sucesos subcríticos dada por $\mathbf{B}_{\mathcal{C}_{\mathbf{s}}^{GH}}(\tilde{p})$, y la de los supercríticos dada por $\tilde{\mathbf{B}}_{\mathcal{C}_{\mathbf{s}}^{GH}}(\tilde{p}) = I^2 \setminus \mathbf{B}_{\mathcal{C}_{\mathbf{s}}^{GH}}(\tilde{p})$. Al igual que en el caso univariante,

5.4. RESULTADOS

puede introducirse un periodo de retorno \tilde{T} para estos últimos sucesos (potencialmente peligrosos), asociado al nivel crítico \tilde{p} , como

$$\tilde{T} = \vartheta(\bar{K}_{\mathcal{C}_s^{GH}}(\tilde{p})) = \frac{\mu}{1 - \bar{K}_{\mathcal{C}_s^{GH}}(\tilde{p})}. \quad (5.21)$$

Nótese que $K_{\mathcal{C}_s^{GH}}(\tilde{p})$ es la medida de Kendall de la región de los eventos supercríticos, un análogo multivariante del evento univariante $Z > \tilde{p}$. En [Durante y Salvadori \(2010\)](#), \tilde{T} se denomina periodo de retorno secundario, para diferenciarlo del *estándar*,

$$T_p = \frac{1}{1 - \mathcal{C}_s^{GH}(u_p, v_p)}, \quad (5.22)$$

con $w_p = (u_p, v_p)$ un punto situado en la isolínea de \mathcal{C}_s^{GH} que tiene un nivel p . Entonces, el evento de interés corresponde a la región de I^2 en la que $U > u_p$ o $V > v_p$, o bien ambos (véase más detalles en [Durante y Salvadori, 2010](#)).

5.4. Resultados

Se utiliza el modelo lineal propuesto por [Taylor et al. \(2015\)](#), que para este estudio corresponde a $X \sim X_{UPC} + X_{cameras} + X_{marks} + X_{parks} + X_{hospitals} + X_{ATM}$. Las covariables se interpolan mediante la media ponderada por área, debido a que son numéricas (véase la Figura 5.3).

Se utiliza $\mu_\beta = 7$, que corresponde al número de covariables y la constante, para $\Sigma_\beta = 10^6$, para $\mu_\eta = (1, 500)$ y para $\Sigma_\eta = 0.15$ (matriz diagonal dimensión dos). Los valores iniciales para η pueden ser cambiados según convenga. Y se define una función de covarianza exponencial para Y_1 e Y_2 y Matérn para Y_3 .

El MCMC con MALA se ejecuta con 100000000 iteraciones, burn-in inicial de 10000000 y una toma de muestras cada 90000 iteraciones de las 90000000 establecidas. Una vez finalizado el proceso computacional, se realizan las comprobaciones de diagnóstico mediante el análisis de convergencia de la a posteriori logarítmica.

En la Figura 5.4 se puede observar que, el logaritmo de la a posteriori (log-target) de los parámetros se inició aproximadamente en -13.000 , pero rápidamente parece estabilizarse alrededor de -35.000 y -36.000 .

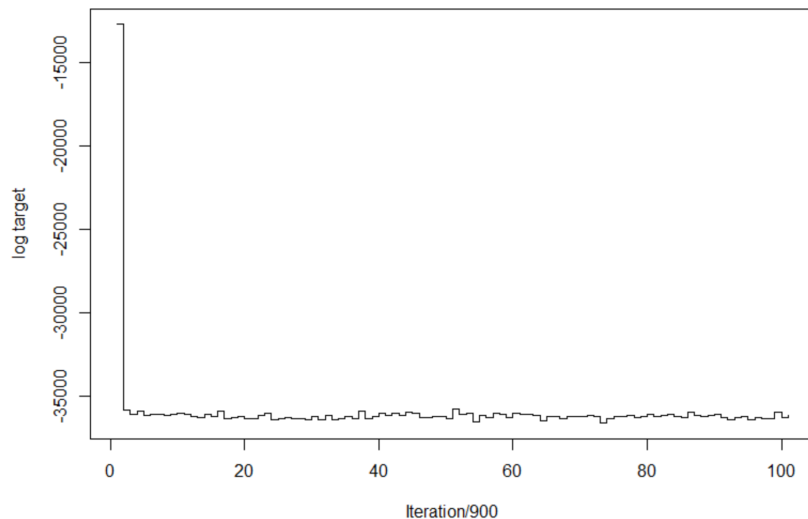


Figura 5.4: Diagnóstico de la convergencia de la distribución a posteriori de los parámetros (log-target)

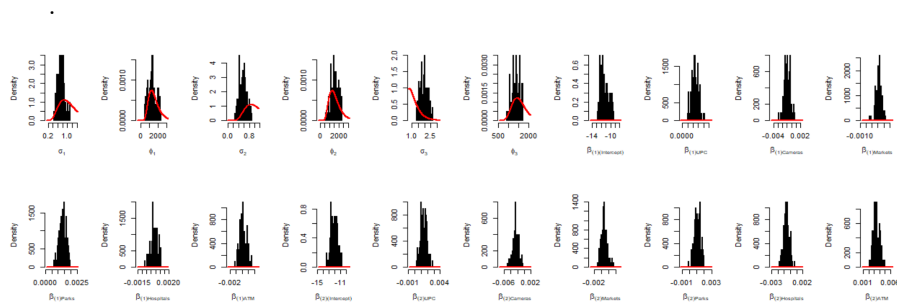


Figura 5.5: A priori y a posteriori de cada parámetro

El parámetro σ muestra una mayor desviación de la a priori en comparación con el parámetro ϕ ; por lo tanto, se debe tener cuidado al hacer declaraciones inferenciales fuertemente probabilísticas sobre estos parámetros, ya que parecen estar influenciados por la a priori.

La Figura 5.6 muestra las a posteriori de las funciones de covarianza para cada tipo de delito y su comportamiento conjunto.

Una vez establecida la convergencia satisfactoria de la cadena, se pueden hacer inferencias sobre el modelo. En primer lugar, se obtiene la tabla de estimaciones de los parámetros (véase la Tabla 5.1).

5.4. RESULTADOS

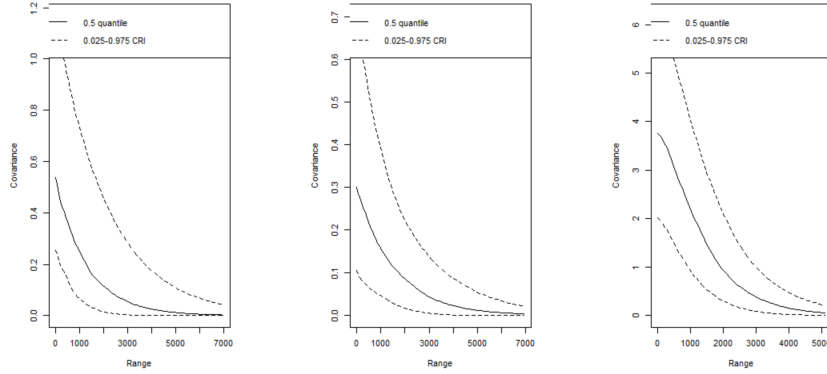


Figura 5.6: Covarianzas a posteriori para Y_1 , Y_2 e Y_3

Como muestra la Tabla 5.1, los parámetros del campo latente σ_k son significativos a excepción de σ_1 que corresponde a hurtos. Todos los parámetros de escala espacial ϕ_k son significativos con medianas de 1361m, 1584m y 1382m, respectivamente.

Tabla 5.1: Valores de los parámetros estimados

Parámetros	Mediana	Inf. 95 %	Sup. 95 %	Signif.
σ_1	7.33e-01	5.05e-01	1.08	no sig.
ϕ_1	1.36e+03	6.14e+02	2.27e+03	
σ_2	5.47e-01	3.25e-01	8.38e-01	
ϕ_2	1.58e+03	7.78e+02	2.29e+03	
σ_3	1.93e+00	1.41e+00	2.48e+00	
ϕ_3	1.38e+03	9.98e+02	1.69e+03	
exp[β_1 (<i>Intercept</i>)]	1.04e-05	3.29e-06	7.57e-05	
exp[β_1 (<i>UPC</i>)]	1.00e+00	1.00e+00	1.00e+00	
exp[β_1 (<i>Cameras</i>)]	9.99e-01	9.97e-01	1.00e+00	no sig.
exp[β_1 (<i>Markets</i>)]	1.00e+00	1.00e+00	1.00e+00	
exp[β_1 (<i>Parks</i>)]	1.00e+00	1.00e+00	1.00e+00	
exp[β_1 (<i>ATM</i>)]	1.00e+00	9.99e-01	1.00e+00	no sig.
exp[β_2 (<i>Intercept</i>)]	5.89e-06	1.63e-06	1.96e-05	
exp[β_2 (<i>UPC</i>)]	1.00e+00	1.00e+00	1.00e+00	
exp[β_2 (<i>Cameras</i>)]	9.97e-01	9.95e-01	9.99e-01	
exp[β_2 (<i>Markets</i>)]	9.99e-01	9.98e-01	1.00e+00	no sig.
exp[β_2 (<i>Parks</i>)]	1.00e+00	1.00e+00	1.00e+00	
exp[β_2 (<i>ATM</i>)]	1.00e+00	1.00e+00	1.00e+00	

El efecto de las distancias a hospitales no es significativo para ninguno de los dos tipos de delitos, por lo que esta covariable fue retirada. Las distancias a las cámaras de vigilancia y ATM no son significativas para hurtos, lo que indica menor probabilidad de que un evento ocurrido cercano a estos puntos de referencia sea hurto. Del mismo modo, si se hace referencia a la covariable mercados, ésta no es significativa para robo a personas, es decir, es poco probable que un evento ocurrido cercano a los mercados se trate de un robo con uso de violencia o amedrentamiento.

Con respecto a las demás covariables, por cada metro de distancia a los diferentes puntos de referencia considerados, se tiene un aumento en la intensidad delictiva por hurto y robo a personas, según se recoge en las columnas 3 y 4 de la Tabla 5.1, con una mediana aproximada de 1.00 y 0.99 para los tipos de delitos, respectivamente.

Vale la pena destacar que las covariables no significativas posiblemente no son atractores significativos para este tipo de delitos.

La Figura 5.7 muestra las intensidades resultantes para Y_1 , Y_2 e Y_3 . Se identifica que Y_3 concentra las intensidades en el centro de la ciudad. Para facilitar la descripción de los resultados, cada vez que se haga referencia al centro de la ciudad, se sobreentenderá que se trata de las regiones que contienen los límites adyacentes a los sectores denominados *La Estación*, *La Panadería* y *Loma de Quito*.

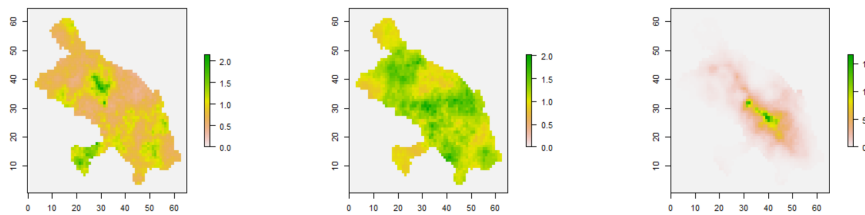


Figura 5.7: Intensidad estimada de hurtos (Λ_1), robos a personas (Λ_2) y conjunta (Λ_3)

La Figura 5.8 muestra la probabilidad condicionada de que un punto en cada ubicación sea de un hurto (lado derecho) o robo a personas (lado izquierdo). Es más probable que se cometan más hurtos que robos a personas independientemente del lugar, puesto que los hurtos se marcan con mayor frecuencia en gran parte de la ciudad.

5.4. RESULTADOS

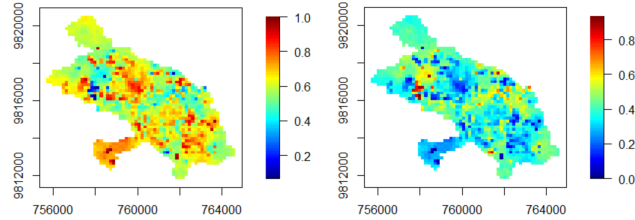


Figura 5.8: Probabilidades condicionadas

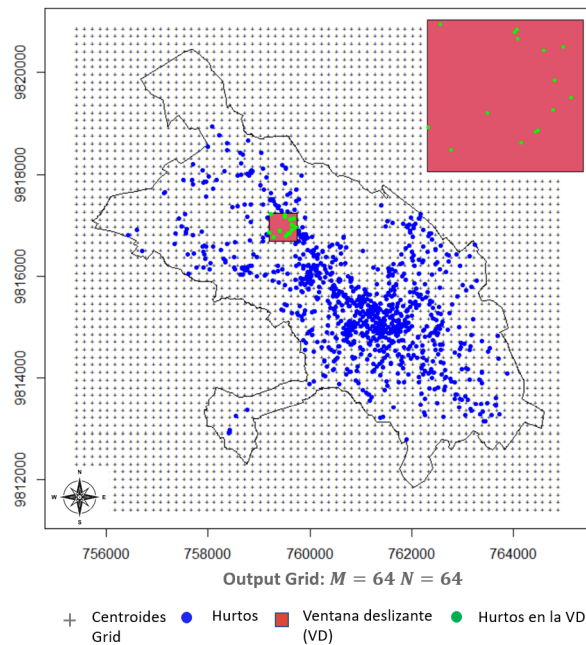


Figura 5.9: Discretización de delitos mediante una ventana deslizante

A partir de la intensidad conjunta hallada mediante el LGCP bivalente y considerando las inconsistencias en la estimación de los parámetros de la GPD en muestras pequeñas ($n < 500$) descrita en la literatura (véase [Salvadori et al., 2007](#); [Malevergne y Sornette, 2006](#); [Northrop et al., 2017](#), entre otros), se simulan 1000 patrones puntuales mediante la distribución de Poisson para cada tipo de delito. Luego se realiza la discretización mediante el uso de una ventana deslizante (VD) para regularizar, suavizar y contrarrestar el efecto de exceso de ceros, es decir, una ventana con movimientos de aproximadamente 139m en horizontal y 135m en vertical. El área de la VD cuadruplica las distancias de

los deslizamientos y recorre toda el área de la malla definida (véase la Figura 5.9).

Tomando la sugerencia de la representación de Sklar, en primer lugar se describe la estimación de las funciones de distribución acumulativas marginales F_1, \dots, F_d .

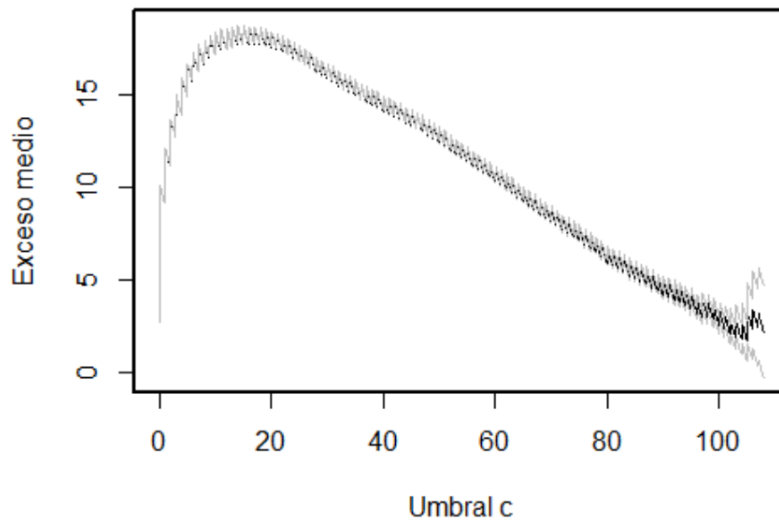


Figura 5.10: Vida residual media de exceso de delitos

La literatura explica varias formas de elegir el umbral. Sin embargo, se usa el método basado en la esperanza de la GPD, de forma que si $Z(\mathbf{s})$ sigue esta distribución con parámetros $\sigma(\mathbf{s})$ y $\xi(\mathbf{s})$, el umbral puede ser el valor donde el gráfico de vida residual media muestre una tendencia lineal, por lo que se analizan los posibles valores contenidos dentro del intervalo donde se visualiza dicha tendencia. Para los datos de hurtos y robos a personas que se analizan, la elección del valor del umbral podría tomarse entre aproximadamente 14 y 80 (véase la Figura 5.10). Para fijar el umbral se realizan análisis previos en función de las estimaciones de los parámetros de forma $\xi(\mathbf{s})$ obtenidas con MCMC, para $c = 15$, $c = 30$ y $c = 45$ (véase la Figura 5.11). Los colores verdes intensos muestran las ubicaciones donde es más probable que se presenten eventos delictivos. Además, muestran tendencia a concentrarse aproximadamente en la zona antes mencionada como centro de la ciudad.

Se identifica que los valores de los parámetros de forma para hurtos son distintos a los obtenidos para robos a personas. Sin embargo, ambos tipos de delitos tienden a centrarse en similares localizaciones \mathbf{s} .

5.4. RESULTADOS

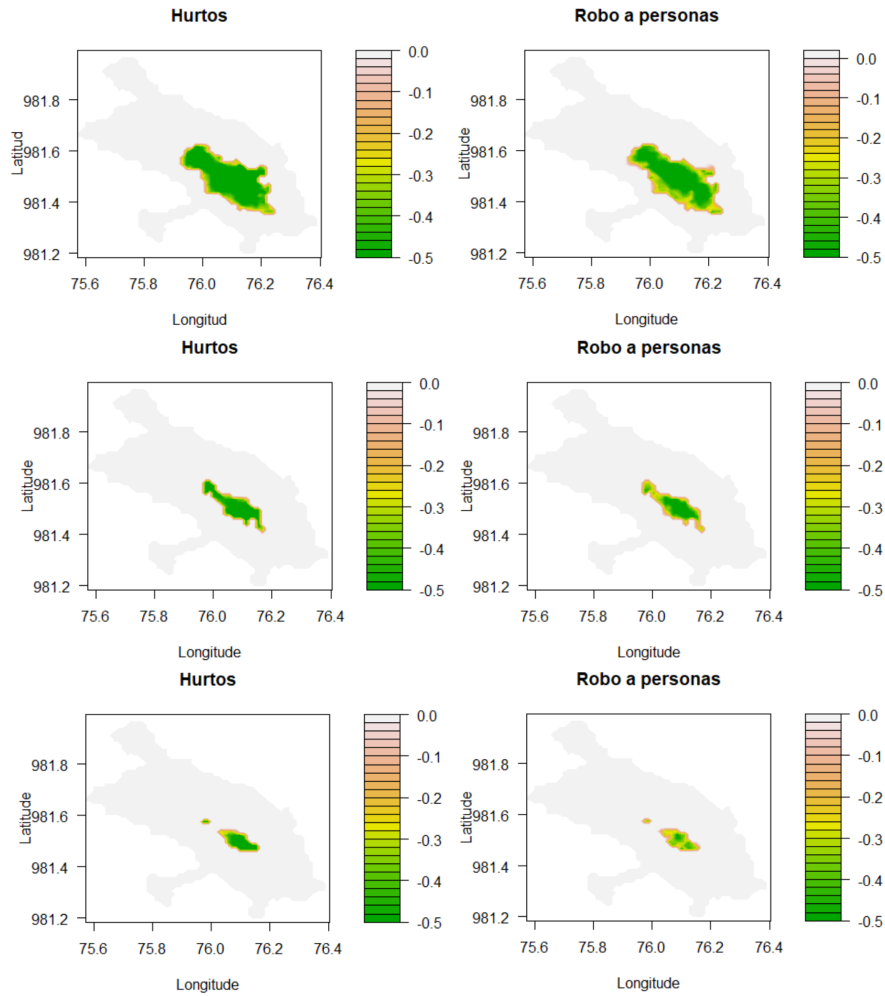


Figura 5.11: Parámetro de forma ξ de la GPD para $c = 15$, $c = 30$ y $c = 45$, respectivamente

Se fija el umbral en $c = 15$ con el propósito de incluir la mayoría de localizaciones con excedencias en el cálculo de las medidas de riesgo; además, al tomar un umbral más alto se dificulta la formación de pares necesarios para el uso de las cópulas. Por el teorema de Pickands, los excesos sobre el umbral están distribuidos según una GPD, por lo que se realiza estimación bayesiana de los parámetros de escala $\sigma(\mathbf{s})$ y de forma $\xi(\mathbf{s})$ de la GPD para cada localización \mathbf{s} , con ayuda de MCMC (véase la Figura 5.12). Las estimaciones de los parámetros de forma para todas las localizaciones \mathbf{s} muestran valores $\xi(\mathbf{s}) < 0$, lo que indica que pertenecen al dominio de atracción de Weibull.

En las Figuras 5.12, 5.13 y 5.14 se muestran los resultados del ajuste de los parámetros de forma $\xi(\mathbf{s})$ y de escala $\sigma(\mathbf{s})$ para algunas localizaciones \mathbf{s} .

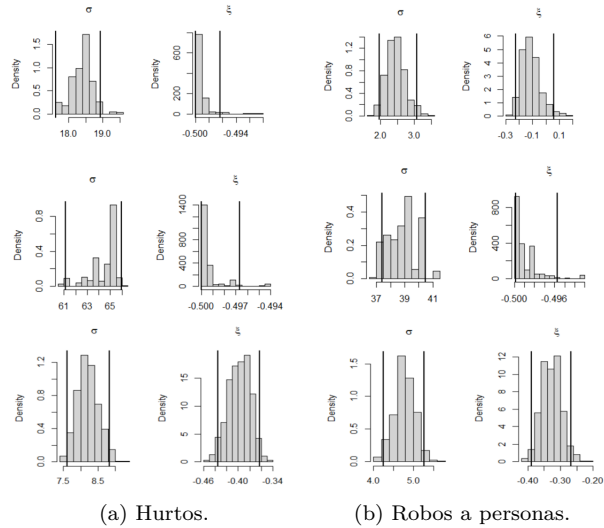


Figura 5.12: Histogramas de los parámetros de escala $\sigma(\mathbf{s})$ y de forma $\xi(\mathbf{s})$ de la GPD, para tres de las localizaciones \mathbf{s}

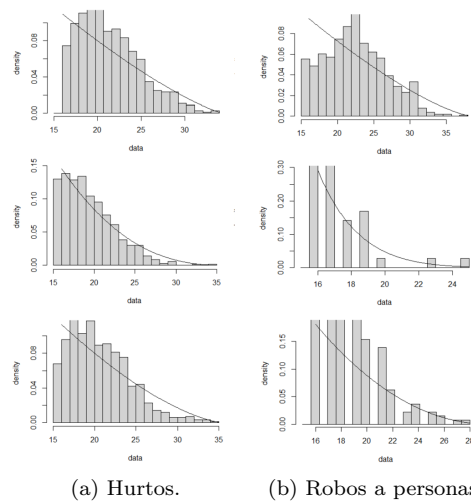


Figura 5.13: Densidades predictivas de las distribuciones a posteriori de los parámetros de la GPD para algunas de las localizaciones \mathbf{s}

5.4. RESULTADOS

Los parámetros de escala $\sigma(\mathbf{s})$, tienden a tomar valores grandes, lo que permite asumir la existencia de la distribución marginal de los excesos sobre el umbral c . La Figura 5.13 muestra las densidades predictivas resultantes de la distribución a posteriori de los parámetros de la GPD para algunas localizaciones \mathbf{s} . Como se observa, las distribuciones muestran comportamientos distintos, pero en general, todas tienden a la GPD.

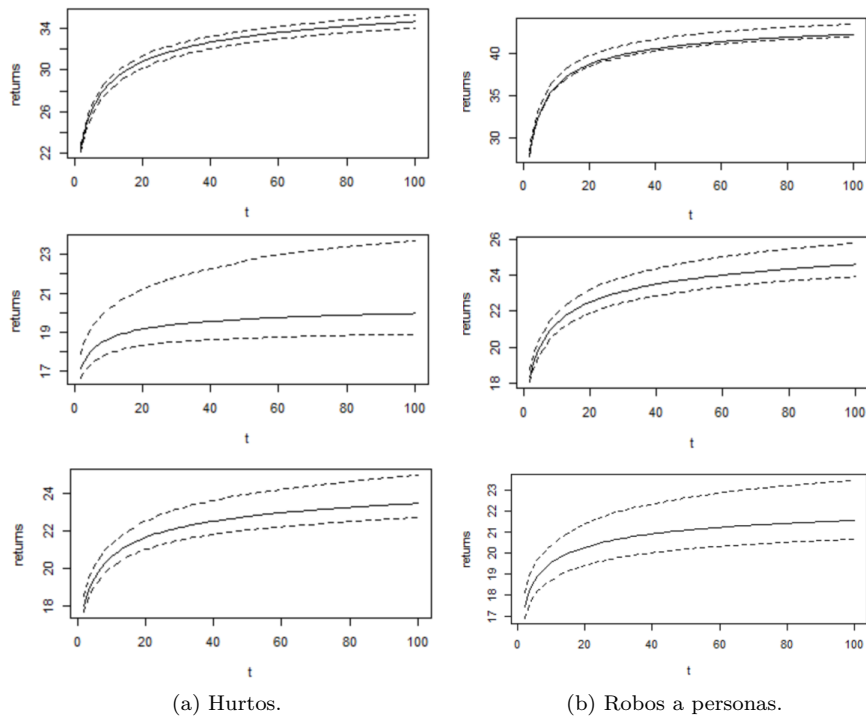


Figura 5.14: Niveles de retorno de la GPD de hurtos y robos a personas para tres de las localizaciones \mathbf{s}

Una de las gráficas que permiten identificar el ajuste de los parámetros de la GPD suele ser la curva de nivel de retorno, por lo que se muestra en la Figura 5.14 para tres localizaciones \mathbf{s} . Se observan diferencias visibles, tanto para hurto como para robo a personas, especialmente presencia de ruido para ambos tipos de delitos. En forma general, las excedencias son GPD y los parámetros de forma $\xi(\mathbf{s})$ indican colas ligeras. Los valores de $\xi(\mathbf{s})$ para hurtos son más bajos que en robos a personas, y estos últimos presentan cierta tendencia a formar colas pesadas en al menos una localización (véase la Figura 5.15).

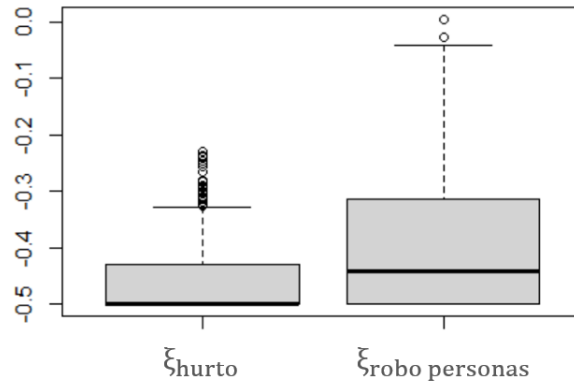


Figura 5.15: Resumen de los valores de los parámetros de forma $\xi(\mathbf{s})$ resultantes

Una vez estimadas las distribuciones marginales acumulativas, se realiza la estimación de las cópulas de Gumbel-Hougaard denotadas como \mathcal{C}_s^{GH} , siendo una alternativa para medir los riesgos bivariantes.

En primer lugar, se estima el parámetro de dependencia θ de la cópula \mathcal{C}_s^{GH} por máxima verosimilitud para cada localización \mathbf{s} y se representan los resultados a través de curvas de nivel (véase la Figura 5.16) y su valor restringido a la región de estudio (véase la Figura 5.17).

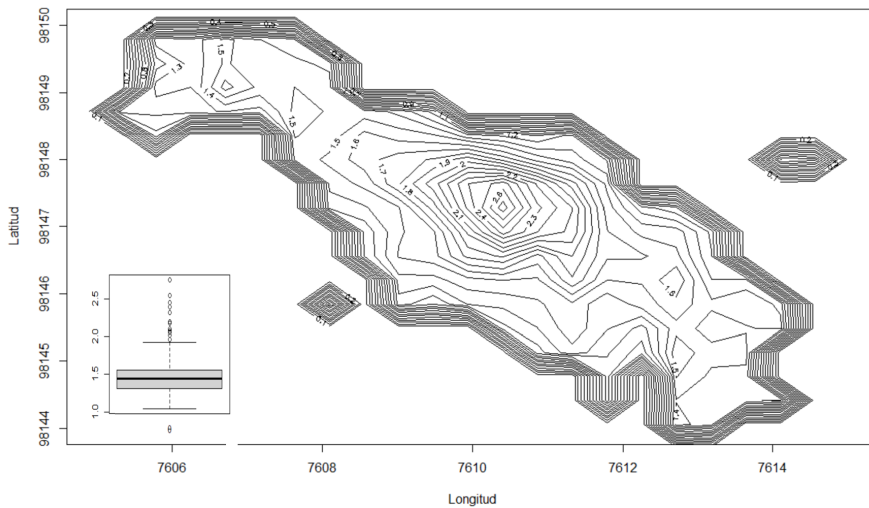


Figura 5.16: Curvas de nivel de los valores del parámetro de dependencia (θ) de las cópulas de Gumbel-Hougaard

5.4. RESULTADOS

La Figura 5.16 muestra las curvas de nivel de los valores del parámetro de dependencia de \mathcal{C}_s^{CH} en cada localización s , en donde se identifica mayor dependencia en la región del centro de la ciudad.

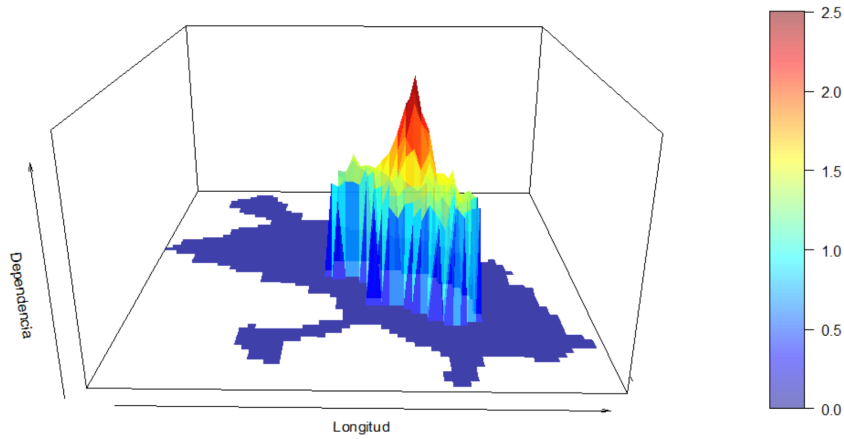


Figura 5.17: Valores del parámetro de dependencia (θ) de la cópula de Gumbel-Hougaard en la región de estudio

En la Figura 5.17 se observa el comportamiento de la dependencia considerando su representatividad con respecto a la región de estudio; en donde, a partir del color amarillo, naranja y hasta el color rojo representan niveles crecientes de la dependencia, alcanzando valores de θ hasta 2.76, con una media aproximada de 1.40. Los valores de θ tienden a 1 en las localizaciones que bordean la región central de la ciudad; esto indica que los hurtos y los robos a personas tienen comportamientos independientes (color celeste) mientras más se alejan del centro de la ciudad.

Estos resultados son de interés para la oficina de policía, debido a que permiten asociar los tipos de delitos en cada zona geográfica y la posible ocurrencia de nuevos eventos, además de proveer información que coadyuve a la optimización de los recursos en la planificación de estrategias para el control y/o prevención de ocurrencias, y de algún modo contrarrestar este fenómeno.

En la Figura 5.18 se muestran las cópulas \mathcal{C}_s^{GH} resultantes para tres localizaciones s , dado el parámetro de dependencia $\theta(s)$.

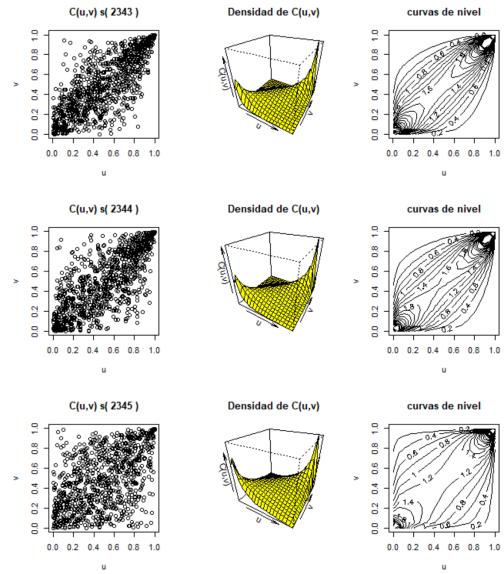


Figura 5.18: Cópulas extremas de Gumbel-Hougaard para tres de las localizaciones s

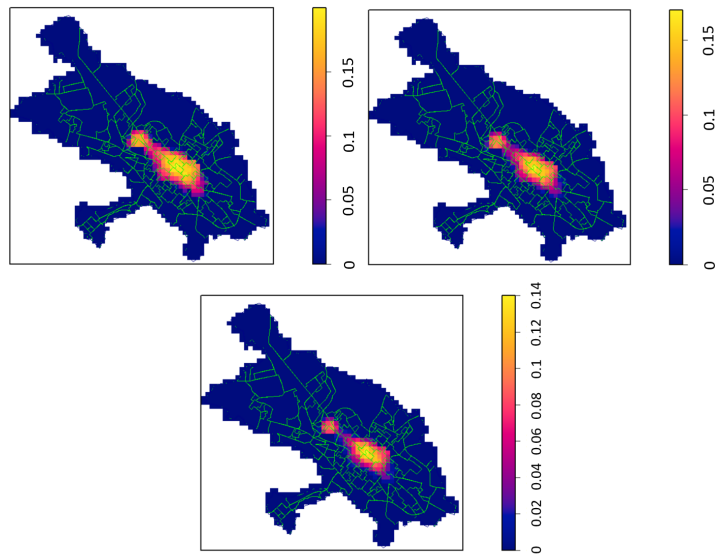


Figura 5.19: Distribución bivalente (\mathcal{C}_s^{GH}) de delitos sobre los cuantiles 50 %, 75 % y 90 %, respectivamente

5.4. RESULTADOS

En la Figura 5.19 se muestra la distribución conjunta (H) obtenida con la \mathcal{C}_s^{GH} sobre el 50%, 75% y 90%, respectivamente. Aproximadamente 26 sectores o barrios se encuentran marcados, de los cuales, los valores más altos (localizaciones de color amarillo) están en alrededor de 12, 9 y 6 regiones para los porcentajes analizados, respectivamente. Esta información es relevante, tanto para la policía como para la sociedad en general, puesto que indica un foco de alerta para el respectivo seguimiento.

A continuación, se analizan las distribuciones de u y v estandarizadas (véase la Figura 5.20), que en general para todas las localizaciones s tienen cierta aproximación a la normal en concordancia con el teorema central de límite (véase la Figura 5.20).

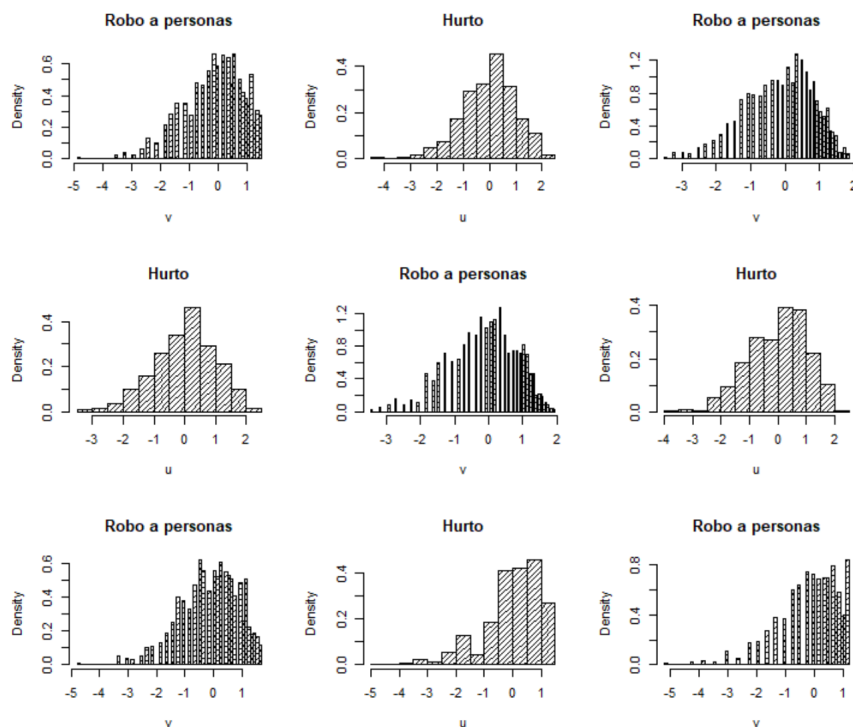


Figura 5.20: Distribuciones de u y v normalizadas

Se estiman las medias y las varianzas de los excesos en cada s por máxima verosimilitud. Se realizan 5000 simulaciones de eventos delictivos y se incluyen las curvas de contorno a 1.25, 2, 5, 10, 20, 50, 100, 200 y 500 niveles de retorno. En la Figura 5.21 se muestran los resultados obtenidos para algunas de las localizaciones. Se identifica mayor dispersión en las localizaciones con menor

número de eventos delictivos, y mayor aglomeración en aquellas con mayor número de eventos.

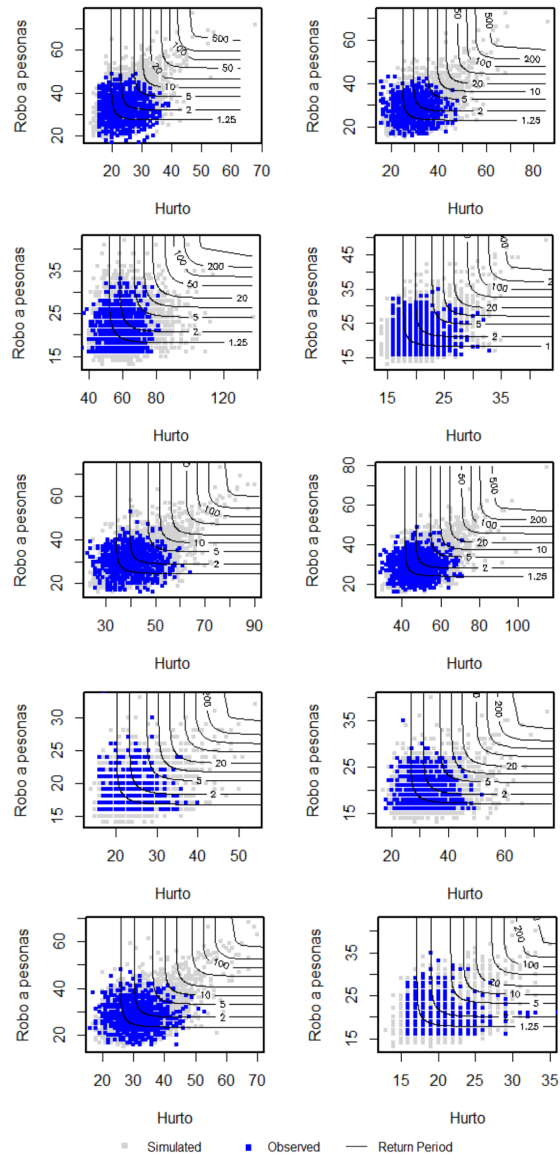


Figura 5.21: Simulación de hurtos y robos a personas para algunas de las localizaciones s mediante \mathcal{C}_s^{CH} y sus contornos

Los resultados muestran que las cópulas modelizan la dependencia de los

5.4. RESULTADOS

hurtos y robos a personas razonablemente (véase la Figura 5.21), lo que se corrobora con el análisis de errores cuadráticos medios (véase la Figura 5.22). Para hallar el RMSE (root-mean-square), se simula el número de excesos que corresponda a cada localización. Se observa que los RMSE más altos son 13 para hurtos y 11 para robos a personas, como mínimo 5 para ambos tipos de delitos, y una media de 9 y 7, respectivamente, por unidad de área (VD). Existe diferencia significativa entre los RMSE para ambos tipos de delitos.

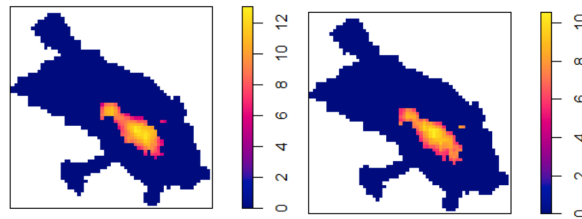


Figura 5.22: Error cuadrático medio (RMSE) para hurtos y robos a personas, respectivamente

Una vez analizado el ajuste de las cópulas de Gumbel-Hougaard, se hallan los $VaR_\alpha(\mathbf{s})$ para cada localización \mathbf{s} con $\alpha = 0.95$, tanto de forma univariante como bivalente. Para el caso bivalente, aproximadamente 35 sectores cubren las localizaciones marcadas, presentando mayor número de delitos (color amarillo) en aquellas que coinciden con los límites adyacentes de los sectores denominados *La Estación*, *La Panadería* y *Loma de Quito* (véase la Figura 5.23).

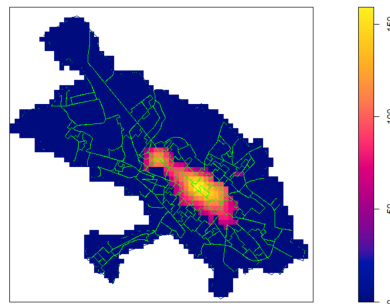
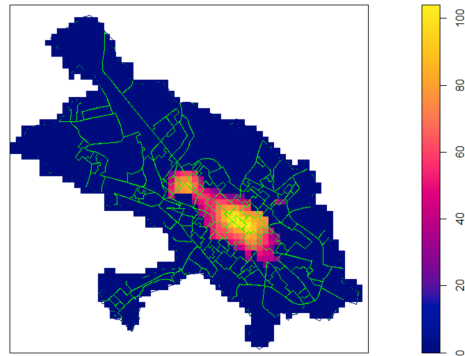


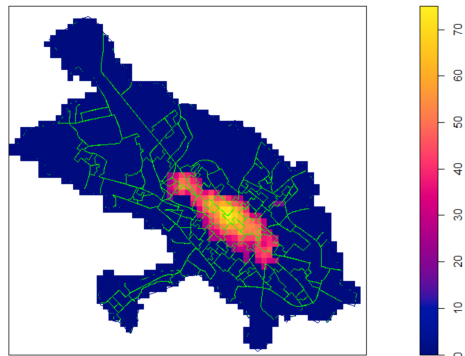
Figura 5.23: $VaR_{0.95}(\mathbf{s})$ bivalente

Para el caso univariante, los hurtos se identifican con mayor frecuencia en el sector denominado *La Concepción* (véase la Figura 5.24a), y el robo a personas se concentra aproximadamente al Sur-Este del sector denominado *Loma de Quito* y los sectores vecinos *La Estación* y *La Panadería* (véase la

Figura 5.24b). Se identifica que aproximadamente suceden 22 hurtos y 19 robos a personas por cada $4 \times (138.62m \times 134.52m)$ de área.



(a) Hurto



(b) Robo a personas

Figura 5.24: $VaR_{0.95}(\mathbf{s})$ univariante

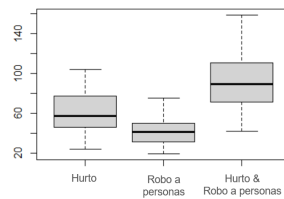


Figura 5.25: Resumen de los $VaR_{0.95}(\mathbf{s})$ univariante y bivalente

Es relevante mencionar que en el centro de la ciudad se concentran los comercios, la circulación vehicular y peatonal, así como también el mayor número de cámaras de vigilancia.

5.4. RESULTADOS

La Figura 5.25 muestra mayor peligrosidad (riesgo) en hurtos por cada unidad de área, y diferencia significativa con respecto a la peligrosidad por robos a personas.

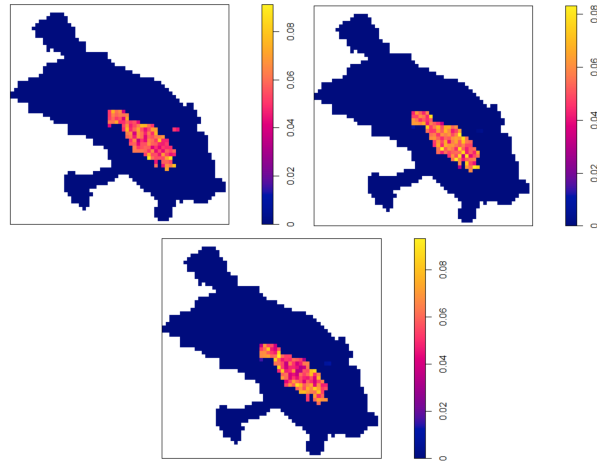


Figura 5.26: Ratio de excepciones para $VaR_{0.95}(s)$ univariante y bivalente

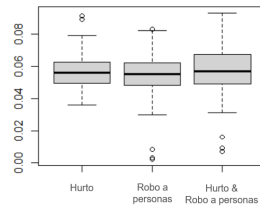
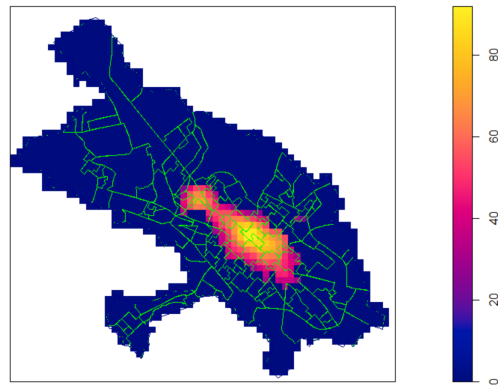
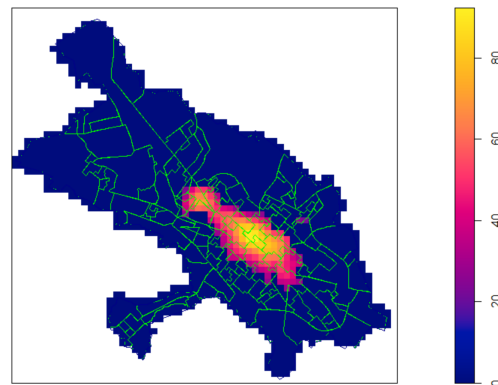


Figura 5.27: Resumen del ratio de excepciones del $VaR_{0.95}(s)$ univariante y bivalente

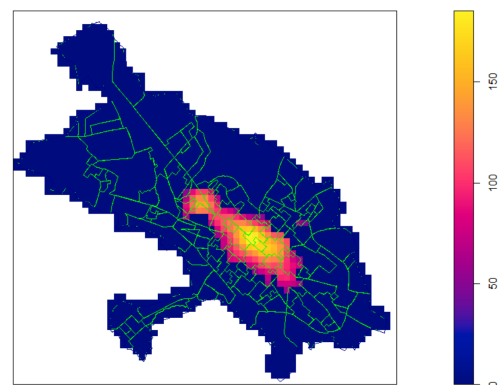
Se evalúan las medidas de riesgo mediante el ratio de excepciones para cada localización s , tanto para el caso univariante como para el bivalente (véase la Figura 5.26). En forma general, la Figura 5.27 resume los valores, en donde se observa una media de 0.057 para hurto, de 0.054 para robo a personas y de 0.057 para el caso bivalente. Los valores se aproximan a $1 - \alpha$ para cada caso, por lo que se pueden aceptar como válidas, puesto que no subestiman ni sobrestiman la peligrosidad. Sin embargo, al realizar el análisis detallado, se observa que varias localizaciones presentan tendencia a la subestimación o sobrestimación.



(a) Hurto



(b) Robo a personas



(c) Hurto y robo a personas

Figura 5.28: $ES_{0.95}(s)$ univariante y bivalente

5.5. DISCUSIÓN Y CONCLUSIONES

En cuanto al $ES_{0.95}(\mathbf{s})$, la Figura 5.28 muestra los resultados para el caso univariante y bivalente.

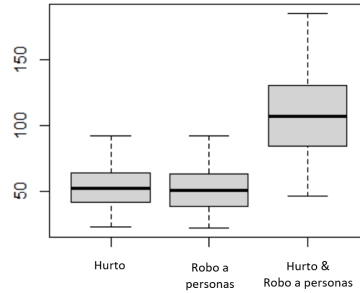


Figura 5.29: Resumen estadístico de los valores del $ES_{0.95}(\mathbf{s})$, univariante y bivalente

En la Figura 5.29 se observa que tanto los hurtos como los robos a personas se presentan con mayor frecuencia en el centro de la ciudad, con promedios de 53 hurtos y 52 robos a personas, por cada $4 \times (138.62m \times 134.52m)$ de área. No existe diferencia significativa entre el $ES_{0.95}(\mathbf{s})$ de ambos tipos de delitos. Con respecto al $ES_{0.95}(\mathbf{s})$ bivalente, el promedio resultante es de 109 eventos.

5.5. Discusión y conclusiones

En múltiples aplicaciones se plantea como objetivo predecir las excedencias espaciales y sus regiones asociadas, debido a que los sucesos inusuales (anomalías) tienden a afectar fuertemente el medio en donde se desarrolla el fenómeno estudiado. [Zhang et al. \(2008\)](#) proponen un enfoque de función de pérdida de Baddeley dado un umbral. Utilizan métricas de imagen para imágenes binarias de cambios de temperatura. [Moser y Ghosh \(2021\)](#) identifican localizaciones de excedencias en un patrón puntual sobre inventario forestal, utilizando regresión no paramétrica con suavizado de núcleo. El estudio más cercano a procesos puntuales LGCP lo realizan [Diggle et al. \(2005\)](#), pero de forma univariante. Como se puede observar, los tres enfoques se basan en patrones puntuales univariantes y no asocian las medidas de riesgo VaR ni ES.

En este capítulo se formula un procedimiento para identificar las excedencias espaciales bivariantes, sus regiones asociadas y una cuantificación de la peligrosidad en términos de probabilidad o riesgo. Se utilizan patrones puntuales bivariantes modelizados con procesos de Cox log-gaussianos (LGCP). Se sigue un enfoque bayesiano tanto en la estimación de los parámetros para

la modelización con LGCP como para los parámetros de escala y de forma de la distribución de Pareto generalizada (GPD).

Para lograr resultados confiables, se simulan como mínimo 1000 patrones puntuales LGCP de cada tipo de eventos. Se realiza la discretización de los delitos mediante una ventana deslizante, estrategia que también sirve como método de regularización y reducción de la aparición de ceros. El proceso de simulación de patrones puntuales garantiza distribuciones marginales i.i.d., por lo que directamente se estiman por máxima verosimilitud los parámetros de dependencia de cada una de las cópulas extremas de Gumbel-Hougaard utilizadas. Una vez obtenidas las distribuciones conjuntas resultantes de las cópulas, se hallan los excesos del LGCP bivalente sobre los cuantiles 50 %, 75 % y 90 %, y sus regiones asociadas, así como también las simulaciones de los delitos estimados. Se hallan las medidas de riesgo VaR y ES para hurtos y robos a personas en forma marginal y conjunta, para las distintas localizaciones.

Se identifica que las localizaciones marcadas con alta probabilidad se asocian a 4 sectores o barrios, calificándolos como altamente peligrosos. Esto corresponde a un aproximado del 3 % (144 sectores son considerados dentro del perímetro estudiado).

Esta información es de relevancia para fortalecer y sustentar los planes de acción de la oficina de policía, debido a su interés particular en identificar y dar seguimiento a las localizaciones con porcentajes altos de eventos delictivos, además de mantener un foco de alerta en aquellas localizaciones que superan un cierto umbral, según VaR y ES, así como también identificar los tipos de delitos.

A pesar de que se muestra un procedimiento para eventos delictivos, este enfoque sería aplicable a cualquier tipo de eventos en un contexto bivalente o, por extensión, multivalente.

6. Conclusiones e ideas abiertas

6.1. Conclusiones

Desde el punto de vista científico es habitual querer predecir fenómenos en función de su evolución espacial, temporal y espacio-temporal. Los procesos estocásticos de Poisson dan origen a los procesos de Cox log-gaussianos (LGCP) y auto-excitados de Hawkes. Sus estructuras estocásticas para la intensidad son flexibles, lo que permite la incorporación de información de primer orden mediante componentes deterministas espaciales y temporales, y de segundo orden mediante componentes estocásticas.

Se describen a continuación aspectos relativos a la implementación práctica de los enfoques de modelización abordados en este trabajo de tesis.

Los modelos auto-excitados como el SPINGARCH trabajan con tiempos discretos y unidades espaciales de referencia tipo áreas (manzanas, barrios, distritos, ciudades, países, etc.). Esto hace que el investigador realice estimaciones sobre lugares previamente delimitados. La estructura en forma de ecuación en diferencias estocástica permite capturar la dependencia tanto del modelo de datos como del proceso espacial latente. La acción de la función de factores exógenos es altamente no lineal, de modo que el uso de B-splines hizo que las predicciones sean más ajustadas a la realidad.

Los modelos LGCP pueden trabajar con tiempos discretos o continuos y sobre estructuras espaciales regulares o irregulares. Lógicamente, la complejidad computacional de estos modelos dependerá del nivel de desagregación espacio-temporal. La estructura separable de primer orden facilita la incorporación de efectos de variables meteorológicas en la componente temporal, y B-splines en las covariables basadas en distancias para la componente espacial. Esto proporcionó predicciones robustas.

Los modelos auto-excitados de Hawkes se trabajan con tiempo continuo y referencias puntuales. Su estructura separable facilita modelizar los términos de las componentes de background y triggering. La estimación de sus coeficientes de relajación permite identificar el efecto desencadenante (reincidencia)

en tiempo y espacio. Se muestra este modelo como complemento del LGCP, puesto que la estructura de background se toma como base sobre la cual se simulan eventos a partir de las intensidades predichas con el LGCP, dando una visión general de los posibles eventos sobre el background.

Los modelos LGCP espaciales bivariantes asocian la interacción entre dos tipos de eventos, lo que permite estimaciones de cada evento condicionados a un efecto común entre los dos. Se toma el efecto común para la simulación de patrones puntuales, y con ello se formula un procedimiento basado en cópulas extremas de Gumbel-Hougaard, que identifica las excedencias espaciales bivariantes y sus regiones asociadas. Las medidas de riesgo VaR y ES cuantificaron la peligrosidad delictiva localizada, con una media del ratio general equilibrada, es decir, en general no se subestima ni sobrestima la peligrosidad.

Los modelos descritos tienen numerosas ventajas en cuanto a flexibilidad en su estructura y resultados. Y especialmente en el ámbito delictivo, porque han permitido cuantificar en su totalidad lo que postula la *teoría de patrones delictivos* descrita en [Higgins y Swartz \(2018\)](#).

Se menciona como posible desventaja el alto coste computacional, puesto que se requiere un tiempo considerable para lograr la convergencia en la estimación de los parámetros.

Mediante la redacción de esta memoria, se facilita al público en general el acceso a las metodologías que se emplean en investigaciones actuales sobre fenómenos de recuentos, con interacción del espacio y el tiempo, especialmente en el contexto de los patrones puntuales y medidas de riesgo.

6.2. Ideas abiertas

En el contexto de los temas abordados en cada uno de los capítulos son muchas las ideas abiertas, especialmente en lo referente al campo de análisis delictivo. Sin embargo, podrían ser de mayor interés las siguientes:

- Incorporar covariables con características categóricas, como lo mencionan [Zhu y Xie \(2022\)](#), de forma que, al combinar la información y la coincidencia de palabras clave de los informes policiales, los incidentes relacionados podrían capturarse automáticamente y ayudar a los investigadores policiales a identificar el modus operandi de la delincuencia.
- Incorporar efectos de covariables temporales y espaciales en el algoritmo interactivo de [Zhuang y Mateu \(2019\)](#), para identificar influencia temporal y/o espacial en las componentes de background y triggering. Como caso particular, contemplar en una periodicidad diaria para variables meteorológicas, especialmente si las variaciones entre los promedios diarios, semanales o mensuales no son significativas.

6.2. IDEAS ABIERTAS

- Identificar los delitos que ocurren solo en la red de calles de una ciudad, ya que mejora la tarea de modelización. En tal caso, el plano euclídeo tiene que ser sustituido por el soporte de la red (ver por ejemplo [Gilardi et al., 2022](#)).
- Plantear modelos para predicciones de ubicación de delitos en serie usando el método Next Hit Predictor (NHP), que adopta el marco de procesos auto-excitantes específicos creados para caracterizar las correlaciones entre delitos cometidos por un mismo delincuente (ver [Li y Wang, 2018](#)).
- Formular una estructura no lineal para la inferencia de las covariables espaciales en la malla computacional, y buscar estrategias para incorporar dicha estructura en la modelización con LGCP espacial bivalente, puesto que se cree que la estructura propuesta por [Taylor et al. \(2015\)](#) requiere un alto nivel de iteraciones para alcanzar la convergencia de los parámetros estimados. Además, de considerar la posibilidad de incorporar covariables espaciales específicas para cada tipo de eventos.
- Modelizar delitos espacio-temporales bivariantes con estructura de covarianza no separable, ya que para el caso univariante [Siino et al. \(2018\)](#) mencionan que una estructura no separable parece describir mejor los datos. Y de este modo, analizar el efecto de la no separabilidad espacio-tiempo en las medidas de riesgo.
- Al ser la GPD sensible tanto al parámetro de forma como al de escala, sería interesante profundizar en el planteamiento de algún método que permita estimar los umbrales variables acordes al comportamiento localizado de los delitos extremos, debido a que para umbrales demasiado bajos las colas no alcanzan el criterio de convergencia, lo que provoca un gran sesgo y, por defecto, resultados incorrectos; sumado a esto, las características específicas que presenta cada localización. Por otro lado, si los umbrales son demasiado altos, se dificulta la formación de pares que respeten su realización. Una alternativa pudiera ser el planteamiento de [Northrop et al. \(2017\)](#), que utilizan la validación cruzada bayesiana para abordar el equilibrio comparando umbrales basados en la capacidad de predicción en niveles extremos.
- Sería interesante profundizar en la estimación bayesiana para hallar los parámetros de dependencia de las cópulas, sobre todo de cópulas extremas, puesto que las implementaciones en software son limitadas, así como también para obtener simulaciones de las medidas VaR y ES, y disponer de distribuciones futuras del riesgo.

Bibliografía

- Abad, P., Benito, S., y López, C. (2014). A comprehensive review of value at risk methodologies. *The Spanish Review of Financial Economics*, 12(1):15–32.
- Adelfio, G. y Chiodi, M. (2015). Alternated estimation in semi-parametric space-time branching-type point processes with application to seismic catalogs. *Stochastic Environmental Research and Risk Assessment*, 29(2):443–450.
- Aït-Sahalia, Y., Cacho-Diaz, J., y Laeven, R. J. (2015). Modeling financial contagion using mutually exciting jump processes. *Journal of Financial Economics*, 117(3):585–606.
- Andresen, M. y Malleon, N. (2015). Intra-week spatial-temporal patterns of crime. *Crime Science*, 4(12):1–11.
- Augustin, N. H., McNicol, J., y Marriott, C. A. (2006). Using the truncated auto-Poisson model for spatially correlated counts of vegetation. *Journal of Agricultural, Biological, and Environmental Statistics*, 11(1):1–23.
- Baddeley, A., Coeurjolly, J. F., Rubak, E., y Waagepetersen, R. (2014). Logistic regression for spatial Gibbs point processes. *Biometrika*, 101:377–392.
- Baddeley, A., Rubak, E., y Turner, R. (2015). *Spatial Point Patterns: Methodology and Applications with R*. Chapman & Hall.
- Baddeley, A., Turner, R., Møller, J., y Hazelton, M. (2005). Residual analysis for spatial point processes (with discussion). *Journal of the Royal Statistical Society, Series B*, 67(5):617–666.
- Baddeley, A. J., Møller, J., y Waagepetersen, R. (2000). Non- and semi-parametric estimation of interaction in inhomogeneous point patterns. *Statistica Neerlandica*, 54:329–350.
- Bartlett, M. S. (1963). The spectral analysis of point processes. *Journal of the Royal Statistical Society, Series B*, 25(2):264–281.
- Beirlant, J., Teugels, J. L., y Vynckier, P. (1996). *Practical Analysis of Extreme values*, volume 50. Leuven University Press Leuven.
- Beneš, V., Bodlák, K., Møller, J., y Waagepetersen, R. (2002). Bayesian analysis of log-Gaussian Cox processes for disease mapping. Technical report, Department Mathematical Sciences, Aalborg University.
- Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society, Series B*, 36(2):192–225.
- Boulieri, A., Liverani, S., de Hoogh, K., y Blangiardo, M. (2017). A space-time multivariate Bayesian model to analyse road traffic accidents by severity. *Journal of the Royal Statistical Society, Series A*, 180(1):119–139.
- Brix, A. y Diggle, P. J. (2001). Spatio-temporal prediction for log-Gaussian Cox processes. *Journal of the Royal Statistical Society, Series B*, 63(4):823–841.
- Castellanos, M. E. y Cabras, S. (2007). A default Bayesian procedure for the generalized Pareto distribution. *Journal of Statistical Planning and Inference*, 137(2):473–483.

-
- Castro, D., Jácomey, J. C., y Mancero, J. (2015). Seguridad ciudadana en Ecuador: Política ministerial y evaluación de impacto, años 2010–2014. *Nova Criminis*, 9(1):111–148.
- Cepa, C., Zabala, R., y López, M. (2018). Proyecto seguridad barrial con involucramiento de los vecinos en Riobamba - Ecuador. *Revista Caribeña de Ciencias Sociales*, 1(1):1–21.
- Chávez, Y., Cortez, P., y Medina, P. (2013). Cuantificación de las pérdidas inesperadas ocasionadas por la delincuencia en Ecuador. *Analitika: Revista de Análisis Estadístico*, 5(1):51–64.
- Chiang, W.-H., Liu, X., y Mohler, G. (2022). Hawkes process modeling of COVID-19 with mobility leading indicators and spatial covariates. *International Journal of Forecasting*, 38(2):505–520.
- Chiodi, M. y Adelfio, G. (2011). Forward likelihood-based predictive approach for space–time point processes. *Environmetrics*, 22(6):749–757.
- Choi, E. y Hall, P. (2001). Nonparametric analysis of earthquake point-process data. *Lecture Notes-Monograph Series*, 36:324–344.
- Ciuperca, G. y Mercadier, C. (2010). Semiparametric estimation for heavy tailed distributions. *Extremes*, 13(1):55–87.
- Clark, N. y Dixon, P. (2018). Modeling and estimation for self-exciting spatio-temporal models of terrorist activity. *The Annals of Applied Statistics*, 12(1):633–653.
- Clark, N. J. y Dixon, P. M. (2021). A class of spatially correlated self-exciting statistical models. *Spatial Statistics*, 43:100493.
- Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer Series in Statistics.
- Contreras Roza, D. L. (2018). Spatio-temporal point process modeling of urban-crime data in bogota through conditional intensity techniques. Master’s thesis, Universidad Nacional de Colombia.
- Cox, D. R. (1955). Some statistical methods connected with series of events. *Journal of the Royal Statistical Society, Series B*, 17(2):129–157.
- Cox, D. R. y Lewis, P. A. (1966). *The Statistical Analysis of Series of Events*. Springer.
- Cressie, N. y Moores, M. T. (2021). Spatial statistics. *arXiv preprint arXiv:2105.07216*.
- Cui, L., Li, Z., y Yi, H. (2019). Partial self-exciting point processes and their parameter estimations. *Communications in Statistics - Simulation and Computation*, 48(10):2913–2935.
- Davies, T. M. y Hazelton, M. L. (2013). Assessing minimum contrast parameter estimation for spatial and spatio-temporal log-Gaussian Cox processes. *Statistica Neerlandica*, 67:355–389.
- Davis, R. A., Fokianos, K., Holan, S. H., Joe, H., Livsey, J., Lund, R., Pipiras, V., y Ravishanker, N. (2021). Count time series: A methodological review. *Journal of the American Statistical Association*, 116(535):1533–1547.
- Davison, A. C. y Smith, R. L. (1990). Models for exceedances over high thresholds. *Journal of the Royal Statistical Society, Series B*, 52(3):393–425.
- De Haan, L. (1970). *On Regular Variation and Its Application to the Weak Convergence of Sample Extremes*. Mathematical Centre Tracts. Mathematical Centre Tracts.
- Diggle, P., Rowlingson, B., y Li Su, T. (2005). Point process methodology for on-line spatio-temporal disease surveillance. *Environmetrics*, 16:423–434.
- Diggle, P. J., Moraga, P., Rowlingson, B., y Taylor, B. M. (2013). Spatial and spatio-temporal log-Gaussian Cox processes: Extending the geostatistical paradigm. *Statistical Science*, 28(4):542–563.
- Durante, F. y Salvadori, G. (2010). On the construction of multivariate extreme value models via copulas. *Environmetrics*, 21(2):143–161.

BIBLIOGRAFÍA

- Durbán, M. (2015). *Modelos Aditivos Generalizados con P-splines*, page 92. Universidad Carlos III de Madrid.
- ECU911 (2020). Servicio integrado de seguridad. Technical Report 1, Servicio Integrado de Seguridad, <https://www.ecu911.gob.ec>.
- Escudero, I., Angulo, J. M., y Mateu, J. (2022). A spatially correlated model with generalized auto-regressive conditionally heteroskedastic structure for counts of crimes. *Entropy*, 24(7):892.
- Ferland, R., Latour, A., y Oraichi, D. (2006). Integer-valued GARCH process. *Journal of Time Series Analysis*, 27(6):923–942.
- Fisher, R. A. y Tippett, L. H. C. (1928). Limiting forms of the frequency distribution of the largest or smallest member of a sample. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 24, pages 180–190. Cambridge University Press.
- Flagg, K. y Hoegh, A. (2022). The integrated nested Laplace approximation applied to spatial log-Gaussian Cox process models. *Journal of Applied Statistics*, pages 1–24.
- Fokianos, K. y Tjøstheim, D. (2011). Log-linear Poisson autoregression. *Journal of Multivariate Analysis*, 102(3):563–578.
- Fréchet, M. (1927). Sur la loi de probabilité de l'écart maximum. *Annales de la Société Polonaise de Mathématique*, 6:93–116.
- Gabriel, E. (2014). Estimating second-order characteristics of inhomogeneous spatio-temporal point processes. *Methodology and Computing in Applied Probability*, 16(2):411–431.
- Gasparrini, A., Scheipl, F., Armstrong, B., y Kenward, M. (2017). A penalized framework for distributed lag non-linear models. *Biometrics*, 73(3):938–948.
- Gilardi, A., Mateu, J., Borgoni, R., y Lovelace, R. (2022). Multivariate hierarchical analysis of car crashes data considering a spatial network lattice. *Journal of the Royal Statistical Society, Series A*, pages 1–28.
- Gnedenko, B. (1943). Sur la distribution limite du terme maximum d'une série aléatoire. *Annals of Mathematics*, 44:423–453.
- Goetz, J., Guthrie, R., y Brenning, A. (2011). Integrating physical and empirical landslide susceptibility models using generalized additive models. *Geomorphology*, 129(3–4):376–386.
- Gómez, R. S. (2008). Estimación Bayesiana en la familia Pareto generalizada. *Revista de Matemática: Teoría y Aplicaciones*, 15(1):71–82.
- González, J. A. y Mateu, J. (2021). A double stochastic point process approach for spatio-temporal dynamics and prediction of crime data. Master's thesis, Universitat Jaume I.
- González, J. A., Rodríguez-Cortés, F. J., Cronie, O., y Mateu, J. (2016). Spatio-temporal point process statistics: A review. *Spatial Statistics*, 18:505–544.
- Grimshaw, S. D. (1993). Computing maximum likelihood estimates for the generalized Pareto distribution. *Technometrics*, 35(2):185–191.
- Gumbel, E. J. (1958). *Statistics of Extremes*. Columbia University Press, New York Chichester, West Sussex.
- Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109.
- Hawkes, A. G. (1971a). Point spectra of some mutually exciting point processes. *Journal of the Royal Statistical Society, Series B*, 33(3):438–443.
- Hawkes, A. G. (1971b). Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1):83–90.
- Hawkes, A. G. (2018). Hawkes processes and their applications to finance: A review. *Quantitative Finance*, 18:193–198.
- Hawkes, A. G. y Oakes, D. (1974). A cluster process representation of a self-exciting process. *Journal of Applied Probability*, 11:493–503.

-
- Heinen, A. (2003). Modeling time series count data: An auto-regressive conditional Poisson model. *CORE Discussion Paper 2003/62*. Available at SSRN: <https://ssrn.com/abstract=1117187> or <http://dx.doi.org/10.2139/ssrn.1117187>.
- Henderson, S. B., Shahirinia, A. H., y Tavakoli Bina, M. (2021). Bayesian estimation of copula parameters for wind speed models of dependence. *IET Renewable Power Generation*, 15(16):3823–3831.
- Hessellund, K. B., Xu, G., Guan, Y., y Waagepetersen, R. (2020). Second order semi-parametric inference for multivariate log-Gaussian Cox processes. *arXiv preprint arXiv:2012.02155*.
- Higgins, E. M. y Swartz, K. (2018). Edgeways as a theoretical extension: connecting crime pattern theory and new urbanism. *Crime Prevention and Community Safety*, 20(1):1–15.
- Hosking, J. R. y Wallis, J. R. (1987). Parameter and quantile estimation for the generalized Pareto distribution. *Technometrics*, 29(3):339–349.
- Hsu, C.-P., Huang, C.-W., y Chiou, W.-J. P. (2012). Effectiveness of copula-extreme value theory in estimating value-at-risk: Empirical evidence from Asian emerging markets. *Review of Quantitative Finance and Accounting*, 39(4):447–468.
- Hu, T., Zhu, X., Duan, L., y Guo, W. (2018). Urban crime prediction based on spatio-temporal Bayesian model. *PLoS ONE*, 13(10):e0206215.
- Ilhan, F. y Kozat, S. S. (2020). Modeling of spatio-temporal Hawkes processes with randomized kernels. *IEEE Transactions on Signal Processing*, 68:4946–4958.
- Illian, J. B., Sørbye, S. H., y Rue, H. (2012). A toolbox for fitting complex spatial point process models using integrated nested Laplace approximation (INLA). *The Annals of Applied Statistics*, 6.
- INEC (2012). Encuesta de victimización y percepción de inseguridad 2011. Techreport, Instituto Nacional de Estadísticas y Censos. [Consulted in June 2021].
- Joe, H. (2015). *Dependence Modeling with Copulas (1st ed.)*. Chapman and Hall/CRC.
- Johnson, O., Diggle, P., y Giorgi, E. (2019). A spatially discrete approximation to log-Gaussian Cox processes for modelling aggregated disease count data. *Statistics in Medicine*, 38:4871–4887.
- Jun, M. y Cook, S. (2022). Flexible multivariate spatio-temporal Hawkes process models of terrorism. *arXiv preprint arXiv:2202.12346*.
- Jung, R. C. y Tremayne, A. R. (2011). Useful models for time series of counts or simply wrong ones? *AStA Advances in Statistical Analysis*, 95:59–91.
- Kaiser, M. S. y Cressie, N. (1996). Modeling Poisson variables with positive spatial dependence. *Statistics & Probability Letter*, 35:423–432.
- Karmakar, M. (2017). Dependence structure and portfolio risk in Indian foreign exchange market: A GARCH-EVT-copula approach. *The Quarterly Review of Economics and Finance*, 64:275–291.
- Kedemo, B. y Fokianos, K. (2002). *Regression Models for Time Series Analysis*. Wiley-Interscience.
- Kojadinovic, I. y Yan, J. (2010). Modeling multivariate distributions with continuous margins using the copula R package. *Journal of Statistical Software*, 34:1–20.
- Kong, X., Huang, G., Fan, Y., y Li, Y. (2015). Maximum entropy-Gumbel-Hougaard copula method for simulation of monthly streamflow in Xiangxi river, China. *Stochastic Environmental Research and Risk Assessment*, 29(3):833–846.
- Kriele, M. y Wolf, J. (2014). *Value-oriented risk management of insurance companies*. Springer Science & Business Media.
- Kuronen, M., Särkkä, A., Vihola, M., y Myllymäki, M. (2022). Hierarchical log-Gaussian Cox process for regeneration in uneven-aged forests. *Environmental and Ecological Statistics*, 29(1):185–205.

BIBLIOGRAFÍA

- Lewis, T. y Govier, L. (1964). Some properties of counts of events for certain types of point process. *Journal of the Royal Statistical Society, Series B*, 26(2):325–337.
- Li, Y. y Wang, T. (2018). Next hit predictor-self-exciting risk modeling for predicting next locations of serial crimes. *arXiv preprint arXiv:1812.05224*, pages 1–4.
- Lieshout, V., Nicolette, M., y Stein, A. (2012). Earthquake modelling at the country level using aggregated spatio-temporal point processes. *Mathematical Geosciences*, 44(3):309–326.
- Liu, J. y Vanhatalo, J. (2020). Bayesian model based spatio-temporal survey designs and partially observed log-Gaussian Cox process. *Spatial Statistics*, 35:100392.
- Malevergne, Y. y Sornette, D. (2006). *Extreme Financial Risks from Dependence to Risk Management*. Springer.
- Manosalva, B. H. (2020). Application of extreme value theory and multivariate copula for VaR measurement of a currencies portfolio of latin american countries. *Odeon*.
- Markose, S. y Alentorn, A. (2005). Option pricing and the implied tail index with the generalized extreme value (GEV) distribution. In *Working paper, University of Essex*.
- Marsan, D. y Lengliné, O. (2010). A new estimation of the decay of aftershock density with distance to the mainshock. *Journal of Geophysical Research: Solid Earth*, 115(B9).
- McNeil, A. J. y Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. *Journal of Empirical Finance*, 7(3-4):271–300.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., y Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092.
- Mises, R. v. (1936). La distribution de la plus grande de n valeurs. *Rev. Math. Union Interbalcanique*, 1:141–160.
- Mohler, G., Short, M., Brantingham, J., Schoenberg, F., y Tita, G. (2011). Self-exciting point process modeling of crime. *Journal of the American Statistical Association*, 106(493):100–108.
- Møller, J. y Díaz-Avalos, C. (2010). Structured spatio-temporal shot-noise Cox point process models, with a view to modelling forest fires. *Scandinavian Journal of Statistics*, 37:2–25.
- Moller, J. y Ghorbani, M. (2012). Aspects of second-order analysis of structured inhomogeneous spatio-temporal point processes. *Statistica Neerlandica*, 66(4):472–491.
- Moller, J., Syversveen, A. R., y Waagepetersen, R. P. (1998). Log-Gaussian Cox processes. *Statistics*, 25:451–482.
- Moller, J. y Waagepetersen, R. P. (2003). *Statistical Inference and Simulation for Spatial Point Processes*. Chapman and Hall/CRC, Boca Raton.
- Møller, J. y Waagepetersen, R. P. (2007). Modern statistics for spatial point processes. *Scandinavian Journal of Statistics*, 34(4):643–684.
- Mora Valencia, A. (2010). Estimadores del índice de cola y el valor en riesgo. *Cuadernos de Administración (Universidad del Valle)*, 44:71–88.
- Moradi, M. M. y Mateu, J. (2020). First and second-order characteristics of spatio-temporal point processes on linear networks. *Journal of Computational and Graphical Statistics*, 29(3):432–443.
- Moreno Trujillo, J. F. (2019). Stochastic model for risky assets price using Hawkes processes. *Munich Personal RePEc Archive*, 101327:1–17.
- Morris, M., Wheeler, K., Simpson, D., Mooney, S., Gelman, A., y DiMaggio, C. (2019). Bayesian hierarchical spatial models: Implementing the Besag York Mollie model in stan. *Spatial and Spatio-temporal Epidemiology*, 31:100301.
- Moser, G. E. y Ghosh, S. (2021). Finding exceedance locations in a large spatial database using nonparametric regression. *Ecological Complexity*, 45:100905.

-
- Myllymäki, M. y Penttinen, A. (2009). Conditionally heteroscedastic intensity-dependent marking of log-Gaussian Cox processes. *Statistica Neerlandica*, 63(4):450–473.
- Nasirzadeh, F., Shishebor, Z., y Mateu, J. (2021). On new families of anisotropic spatial log-Gaussian Cox processes. *Stochastic Environmental Research and Risk Assessment*, 35(2):183–213.
- Northrop, P. J., Attalides, N., y Jonathan, P. (2017). Cross-validators extreme value threshold selection and uncertainty with application to ocean storm severity. *Journal of the Royal Statistical Society, Series C*, 66(1):93–120.
- Ogata, Y. (1988). Statistical models for earthquake occurrences and residual analysis for point processes. *Journal of the American Statistical Association*, 83(401):9–27.
- Okawa, M., Iwata, T., Tanaka, Y., Toda, H., Kurashima, T., y Kashima, H. (2021). Dynamic Hawkes processes for discovering time-evolving communities states behind diffusion processes. *arXiv preprint arXiv:2105.11152*.
- Ozaki, T. (1979). Maximum likelihood estimation of Hawkes self-exciting point processes. *Annals of the Institute of Statistical Mathematics*, 31:145–155.
- Park, J., Schoenberg, F. P., Bertozzi, A. L., y Brantingham, P. J. (2021). Investigating clustering and violence interruption in Gang-related violent crime data using spatial-temporal point processes with covariates. *Journal of the American Statistical Association*, pages 1–14.
- Pickands, J. (1975). Statistical inference using extreme order statistics. *The Annals of Statistics*, 3(1):119–131.
- Pishro-Nik, H. (2016). *Introduction to Probability, Statistics and Random Processes*. Kappa Research.
- Pratiwi, H., Slamet, I., Saputro, D. R. S., y Respatiwan (2017). Self-exciting point process in modeling earthquake occurrences. *Journal of Physics: Conference Series*, 855(1):012033.
- Rathbun, S. L. y Cressie, N. (1994). A space-time survival point process for a longleaf pine forest in Southern Georgia. *Journal of the American Statistical Association*, 89(428):1164–1174.
- Ravindra, K., Rattan, P., Mor, S., y Aggarwal, A. (2019). Generalized additive models: Building evidence of air pollution, climate change and human health. *Environment International*, 132:104987.
- Reinhart, A. (2018). Rejoinder: A review of self-exciting spatio-temporal point processes and their applications. *Statistical Science*, 33:330–333.
- Reinhart, A. y Greenhouse, J. (2017). Self-exciting point processes with spatial covariates: Modeling the dynamics of crime. *Journal of the Royal Statistical Society, Series C*, 67:1305–1329.
- Roberts, G. O. y Tweedie, R. L. (1996). Exponential convergence of Langevin distributions and their discrete approximations. *Bernoulli*, 2(4):341–363.
- Rodrigues, A. y Diggle, P. J. (2012). Bayesian estimation and prediction for inhomogeneous spatio-temporal log-Gaussian Cox processes using low-rank models, with application to criminal surveillance. *Journal of the American Statistical Association*, 107:93–101.
- Rodríguez Cortés, F. J. (2014). *Modelling, estimation and applications of second-order spatio-temporal characteristics of point processes*. PhD thesis, Universitat Jaume I.
- Rostami, M., Mohammadi, Y., Jalilian, A., y Nazparvar, B. (2017). Modeling spatio-temporal variations of substance abuse mortality in Iran using a log-Gaussian Cox point process. *Spat Spatio-Temporal Epidemiol*, 22:15–25.
- Sahu, S. K. (2022). *Bayesian Modeling of Spatio-Temporal Data with R*. Chapman and Hall/CRC.
- Salvadori, G. (2004). Bivariate return periods via 2-copulas. *Statistical Methodology*, 1(1-2):129–144.

BIBLIOGRAFÍA

- Salvadori, G., Michele, C. D., Kottegoda, N. T., y Rosso, R. (2007). *Extremes in Nature: An Approach Using Copulas*. Water Science and Technology Library.
- Schoenberg, F. P. (2002). On rescaled Poisson processes and the Brownian bridge. *Annals of the Institute of Statistical Mathematics*, 54(2):445–457.
- Schoenberg, F. P., Hoffmann, M., y Harrigan, R. J. (2019). A recursive point process model for infectious diseases. *Annals of the Institute of Statistical Mathematics*, 71(5):1271–1287.
- Serra, L., Saez, M., Mateu, J., Varga, D., Juan, P., Díaz-Ávalos, C., y Rue, H. (2014). Spatio-temporal log-Gaussian Cox processes for modelling wildfire occurrence: The case of Catalonia, 1994–2008. *Stochastic Environmental Research and Risk Assessment*, 21:531–563.
- Shirotta, S. y Banerjee, S. (2019). Scalable inference for space-time Gaussian Cox processes. *Journal of Time Series Analysis*, 40:269–287.
- Shirotta, S. y Gelfand, A. E. (2017). Space and circular time log-Gaussian Cox processes with application to crime event data. *The Annals of Applied Statistics*, 11(2):481–503.
- Siino, M., Adelfio, G., y Mateu, J. (2018). Joint second-order parameter estimation for spatio-temporal log-Gaussian Cox processes. *Stochastic Environmental Research and Risk Assessment*, 32:3525–3539.
- Simpson, D., Illian, J. B., Lindgren, F., Sørbye, S. H., y Rue, H. (2016). Going off grid: Computationally efficient inference for log-Gaussian Cox processes. *Biometrika*, 103(1):49–70.
- Sklar, M. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, 8:229–231.
- Smith, R. L. (1985). Maximum likelihood estimation in a class of nonregular cases. *Biometrika*, 72(1):67–90.
- Solarte, G., Soto, J., y Muñoz, L. (2013). Matrices dispersas descripción y aplicaciones. *Scientia et Technica*, 18(1):171–177.
- Sørbye, S. H., Illian, J. B., Simpson, D. P., Burslem, D., y Rue, H. (2019). Careful prior specification avoids incautious inference for log-Gaussian Cox point processes. *Journal of the Royal Statistical Society, Series C*, 68(3):543–564.
- Tang, J., Zhao, C., Liu, F., Hao, W., y Gao, F. (2022). Analyzing travel destinations distribution using large-scaled GPS trajectories: A spatio-temporal log-Gaussian Cox process. *Physica A: Statistical Mechanics and its Applications*, page 127305.
- Taylan, P., Weber, G.-W., Liu, L., y Yerlikaya-Özkurt, F. (2010). On the foundations of parameter estimation for generalized partial linear models with B-splines and continuous optimization. *Computers & Mathematics with Applications*, 60(1):134–143.
- Taylor, B. M., Davies, T. M., Rowlingson, B. S., y Diggle, P. J. (2013). lgcp: An R package for inference with spatial and spatio-temporal log-Gaussian Cox processes. *Journal of Statistical Software*, 52:1–40.
- Taylor, B. M., Davies, T. M., Rowlingson, B. S., y Diggle, P. J. (2015). Bayesian inference and data augmentation schemes for spatial, spatio-temporal and multivariate log-Gaussian Cox processes in R. *Journal of Statistical Software*, 63:1–48.
- Thamrin, S. y Alimun (2019). Geographical mapping of dengue fever incidence 2012–2016 in Makassar, Indonesia. In *IOP Conference Series: Earth and Environmental Science*, volume 279, pages 1–8.
- Trejo, C., Cisneros, J., et al. (2013). La delincuencia en la ciudad de Guayaquil, un análisis espacial de su distribución por delito. *Revista Caribeña de Ciencias Sociales*, 2013(12):1–17.
- Van Lieshout, M. y Baddeley, A. (1996). A nonparametric measure of spatial interaction in point patterns. *Statistica Neerlandica*, 50(3):344–361.

-
- Vicente, G., Goicoa, T., y Ugarte, M. (2021). A novel multivariate spatio-temporal approach based on splines to analyse different crimes against women. *International Statistical Review*, Epub ahead of print:1–29. ISR-OA-079-20.
- Vuong, Q.-H., La, V.-P., Nguyen, M.-H., Ho, M.-T., Tran, T., y Ho, M.-T. (2020). Bayesian analysis for social data: A step-by-step protocol and interpretation. *MethodsX*, 7:1–17.
- Waagepetersen, R. (2004). Convergence of posteriors for discretized log-Gaussian Cox processes. *Statistics & Probability Letters*, 66(3):229–235.
- Wei, C. H., Zhu, F., y Hoshiyar, A. (2022). Softplus INGARCH models. *Statistica Sinica*, 32(2):1099–1120.
- Xu, Y. y Zhu, F. (2022). A new GJR-GARCH model for \mathbb{Z} -valued time series. *Journal of Time Series Analysis*, 43(3):490500.
- Yang, X., Zhang, J., y Ren, W.-X. (2018). Threshold selection for extreme value estimation of vehicle load effect on bridges. *International Journal of Distributed Sensor Networks*, 14(2):1550147718757698.
- Yu, W., Yang, K., Wei, Y., y Lei, L. (2018). Measuring value-at-risk and expected shortfall of crude oil portfolio using extreme value theory and vine copula. *Physica A: Statistical Mechanics and its Applications*, 490:1423–1433.
- Yuan, B., Schoenberg, F. P., y Bertozzi, A. L. (2021). Fast estimation of multivariate spatio-temporal Hawkes processes and network reconstruction. *SIAM Journal on Mathematics of Data Science*, 73:1127–1152.
- Zhai, C. y Lafferty, J. (2017). A study of smoothing methods for language models applied to ad hoc information retrieval. *ACM SIGIR Forum*, 51(2):268–276.
- Zhang, J., Craigmile, P. F., y Cressie, N. (2008). Loss function approaches to predict a spatial quantile and its exceedance region. *Technometrics*, 50(2):216–227.
- Zhu, F. (2012a). Modeling overdispersed or underdispersed count data with generalized Poisson integer-valued GARCH models. *Journal of Mathematical Analysis and Applications*, 389:58–71.
- Zhu, F. (2012b). Modeling time series of counts with COM-Poisson INGARCH models. *Mathematical and Computer Modelling*, 56:191–203.
- Zhu, S. y Xie, Y. (2022). Spatiotemporal-textual point processes for crime linkage detection. *The Annals of Applied Statistics*, 16(2):1151 – 1170.
- Zhuang, J. (2006). Second-order residual analysis of spatio-temporal point processes and applications in model evaluation. *Journal of the Royal Statistical Society, Series B*, 68:635–653.
- Zhuang, J. y Mateu, J. (2019). A semi-parametric spatio-temporal Hawkes-type point process model with periodic background for crime data. *Journal of the Royal Statistical Society, Series A*, 182(3):919–942.
- Zhuang, J., Ogata, Y., y Vere-Jones, D. (2002). Stochastic declustering of space-time earthquake occurrences. *Journal of the American Statistical Association*, 97(458):369–380.