entropy

MDPI

Article

A Spatially Correlated Model with Generalized Autoregressive Conditionally Heteroskedastic Structure for Counts of Crimes

Isabel Escudero [1,2], José M. Angulo [2,*] and Jorge Mateu [3]

[1] Estadística, Facultad de Ciencias, Escuela Superior Politécnica de Chimborazo, Riobamba EC 060155, Ecuador; aescudero@espoch.edu.es or aisabel@correo.ugr.es
[2] Department of Statistics and Operations Research, University of Granada, 18071 Granada, Spain
[3] Department of Mathematics, University Jaume I, 12071 Castellón, Spain; mateu@mat.uji.es
* Correspondence: jmangulo@ugr.es

Abstract: Crime is a negative phenomenon that affects the daily life of the population and its development. When modeling crime data, assumptions on either the spatial or the temporal relationship between observations are necessary if any statistical analysis is to be performed. In this paper, we structure space–time dependency for count data by considering a stochastic difference equation for the intensity of the space–time process rather than placing structure on a latent space–time process, as Cox processes would do. We introduce a class of spatially correlated self-exciting spatio-temporal models for count data that capture both dependence due to self-excitation, as well as dependence in an underlying spatial process. We follow the principles in Clark and Dixon (2021) but considering a generalized additive structure on spatio-temporal varying covariates. A Bayesian framework is proposed for inference of model parameters. We analyze three distinct crime datasets in the city of Riobamba (Ecuador). Our model fits the data well and provides better predictions than other alternatives.

Keywords: autoregressive structure; Bayesian inference; B-splines; crimes; MCMC; self-exciting models; spatio-temporal patterns

Citation: Escudero, I.; Angulo, J.M.; Mateu, J. A Spatially Correlated Model with Generalized Autoregressive Conditionally Heteroskedastic Structure for Counts of Crimes. *Entropy* **2022**, *24*, 892. https://doi.org/10.3390/e24070892

Academic Editor: Carlo Cattani

Received: 30 May 2022
Accepted: 24 June 2022
Published: 29 June 2022

1. Introduction

Modeling time series of counts has received important and growing attention since the 1950s [1–5] and over recent decades (see [6–10]). It is known that some well-known discrete distributions, such as Poisson and negative binomial (NB), can only deal with overdispersion; however, generalized Poisson (GP) and double Poisson (DP) distributions can treat both overdispersion and underdispersion. The latter two models have some shortcomings or limitations. Alternatively, the class of observation-driven models called integer-valued generalized autoregressive conditionally heteroskedastic (INGARCH) models [9,11] shows flexibility in modeling a wide range of overdispersion and underdispersion cases, while possessing properties that make it methodologically appealing and useful in practice.

Although a classical Poisson INGARCH model appears to provide an adequate framework for modeling count time series data and has been applied to several fields, Ref. [12] pointed out that it cannot be employed for modeling negative correlation amongst counts, and it can exclusively include covariates which result in a positive regression term, since otherwise the mean of the Poisson process becomes negative. In addition, the conditional mean is equal to the conditional variance, and this restriction can lead to poor performance of a Poisson INGARCH model with the existence of potential extreme observations.

To overcome these drawbacks, two INGARCH models have been proposed to represent overdispersion or underdispersion in the same framework. These are the DP model [13] and the GP model [14]; see also [15] for a proposal of a Conway–Maxwell (COM) Poisson INGARCH distribution. The reader is referred to the very latest literature in this field [8–10].

In the spatial statistics literature, Ref. [16] made an early attempt at structuring spatial relationships for count data by conditionally specifying the data model distribution given a fixed spatial region. This, however, leads to a statistical model that only allows for negative association. Refs. [17,18] demonstrated how the statistical model might be modified to allow for both negative and positive correlation. The important assumption in these models is that the observed count distribution may be conditionally determined from the observed count distribution at spatial neighbors, which is a Markov assumption in space. A good modern review on using spatial structure in econometric models is [19]. While count data in the spatial statistics literature have predominately been addressed through structure in a latent process, in the time series literature it has evolved quite differently. For example, the INGARCH model of [13,20] is a time series model for counts where the data model is Poisson with the expectation that is a function of both previous counts and previous expectations. Ref. [20] demonstrated how the INGARCH(1,1) is analogous to an ARMA(1,1) for counts.

Noting the link between the stationary distribution of the INGARCH(1,1) process and a stochastic process given in [3], often called a self-exciting point process, we have a number of possible point process models [21] that have been shown to be beneficial to representing the dynamics of earthquakes, epidemics, forest fires, traffic accidents, or crimes, which is the motivating problem in this paper. We can find a good number of papers in this latter context, see, as nice examples, [22–25].

This paper is motivated by the analysis of crime data in the city of Riobamba (Ecuador) provided by three different governmental agencies with the aim of understanding crime behavior and its interaction with society to further help public institutions to enhance proper actions. We note that there are some existing exploratory studies (see [26,27]) that show relevant characteristics of this crime phenomenon. In any case, they do not go further in proposing a spatio-temporal modeling framework.

Following the line of reasoning of [24], we take into account spatial variation by considering a spatial integer-valued generalized autoregressive conditionally heteroskedastic (SPINGARCH) model. This model shares the INGARCH properties while allowing spatial correlation by adding a latent spatially correlated log-Gaussian process [28]. In this framework, and paralleling [24], we formulate a stochastic difference equation for the intensity of the space–time process within a class of spatially correlated self-exciting spatio-temporal models that captures both dependence due to self-excitation, as well as dependence on an underlying spatial process. We indeed consider some extensions from [24] to adapt such methodology to our particular data context. We note that the model in [24] considers a linear regression structure in the covariates which are also constant in time. We structure space–time dependency for our count data through a combination of distance-based covariates that vary naturally in both space and time. We thus consider a B-splines procedure within a generalized additive model that permits it to handle space–time variation and non-linear dependencies. This is indeed another aspect that makes our model different from that of [24]. Our B-splines strategy also allows us to combine covariates that are only varying in space with others (such as the climatological ones) that vary only in time, and with those based on distances that are varying in both space and time. Altogether, our strategy is more flexible and adapts better to the case of our data.

The plan of the paper is the following. Section 2 presents the motivating crime datasets together with the corresponding spatial and temporal covariates. Section 3 introduces the methodological approach and the Bayesian inferential framework. The related computational aspects and corresponding results are described in Section 4. The paper ends with a discussion in Section 5 together with some open lines for future research.

2. Description

Citizen insecurity is one of the major problems that affects the development of the population in any country. Riobamba, an Ecuadorian city, is the head of the Riobamba canton and capital of the Chimborazo province (Figure 1). It is located in the inter-Andean

region, surrounded by several volcanoes such as Chimborazo, Tungurahua, Altar, and Carihuairazo. Located at 2754 m above sea level, it has a cold Andean climate with an average temperature of 12 °C. According to the 2010 census, this city had 234,170 inhabitants and a population growth of 1.06% until 2014.

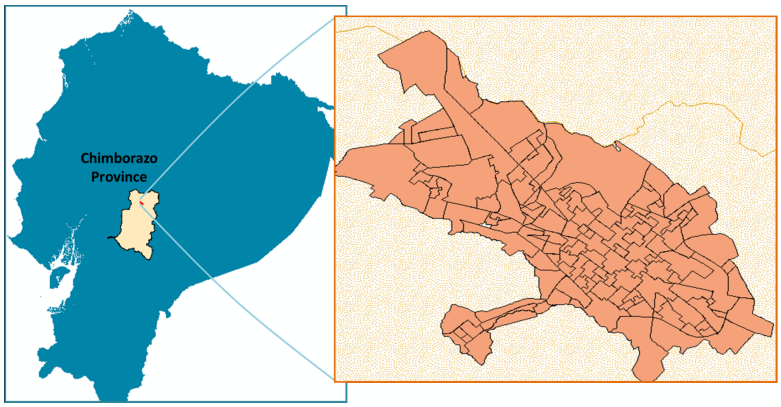


Figure 1. Riobamba within the Chimborazo province in Ecuador.

Commerce is a typical feature of the city, considered a center of business and employment, and it is the third city with higher education institutes in the country. However, one of the main problems that haunts the place is criminal acts such as assaults, robberies in homes and commercial premises, and intimidation, among others, that cause confusion, concern, and significant general losses to the population [29]. In the Ecuadorian national survey conducted in 2011, the province of Chimborazo ranked seventh with 16.9% of people having been victims of some crime, 73.4% of the population considering that the city is unsafe, and 38.0% having experienced a crime increase in their neighborhood.

According to [30], the Ecuadorian government promoted a set of new policies to reduce crime between 2010 and 2014. These policies involved organized civil society and competent entities. At the end of 2014, the victimization rate, homicide, and robberies decreased but with an increase in societal complaints, a sign of greater confidence in the competent institutions.

We use data from three governmental agencies whose mission and vision are to guarantee citizen security and social coexistence (Unidades Policiales Comunitarias (UPC), Consejo de la Judicatura de Chimborazo (CJCH), and Ministerio del Interior (MI)). The ideal registration of information dictates that MI saves all reports from the other two institutions, as shown in Figure 2. However, this is far from being true and analyzing the three datasets will prove this anomaly. Figure 3a shows the criminal acts reported from MI for 2010–2014. Figure 3b depicts the flagrant criminal acts recorded by the CJCH for the period 2015–2019, that is, crimes committed with the arrest of the aggressor within 24 h, and finally, Figure 3c shows the crimes registered by the UPC for 2015–2017. The information used here was provided under a confidentiality contract and is not directly available on any website; however, one can consult the criminal data of the Ministerio del Interior from 2015 onwards at http://cifras.ministeriodegobierno.gob.ec/comisioncifras/inicio.php.

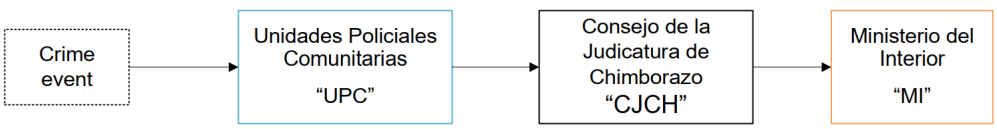


Figure 2. Hierarchical structure for the registration of crimes in Riobamba city.

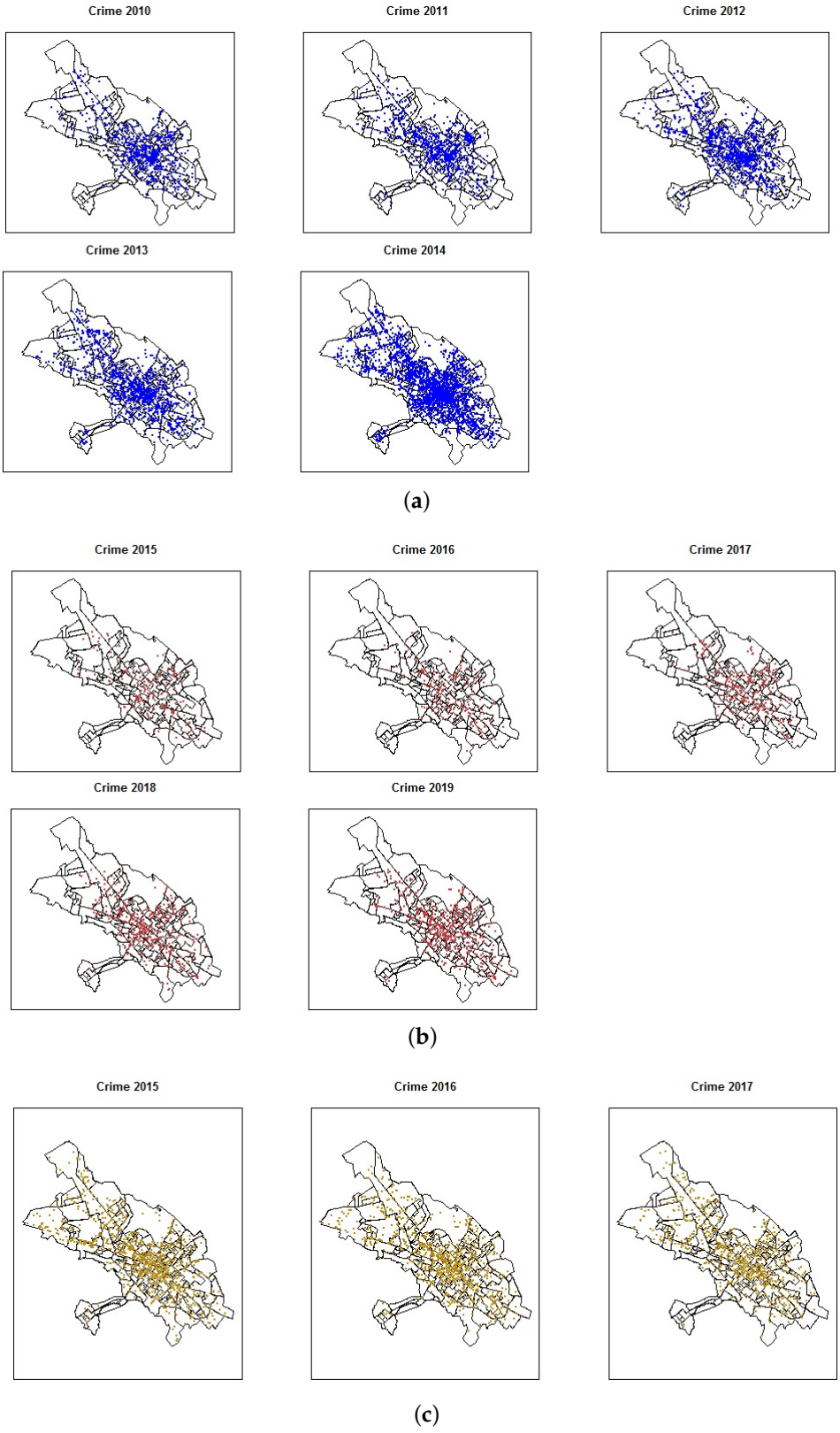


Figure 3. Reported crimes from the three governmental agencies in Riobamba. (**a**) Crimes recorded by MI (2010–2014). (**b**) Flagrant crimes recorded by CJCH (2015–2019). (**c**) Crimes recorded by UPC (2015–2017).

The city of Riobamba is divided into 141 administrative zones as can be seen in Figure 4. The ECU911 (Servicio Integrado de Seguridad) provides the locations in space

and time of the crimes, the location of community police units (upc), and the surveillance cameras installed in strategic locations throughout the city (cam). We also consider some important city landmarks describing areas with a higher pedestrian traffic, such as locations of parks (par), hospitals (hos), and markets (cc), that include squares, shopping malls, and supermarkets, and population density at the administrative zone level. In terms of temporal-varying covariates, we consider monthly averages of temperature and precipitation in the city of Riobamba (see Figure 5); these data are available at http://ceaa.espoch.edu.ec:8080/redEma/. These climatological variables are taken into account because there are previous studies (see [25]) that relate them with theft-based crimes.

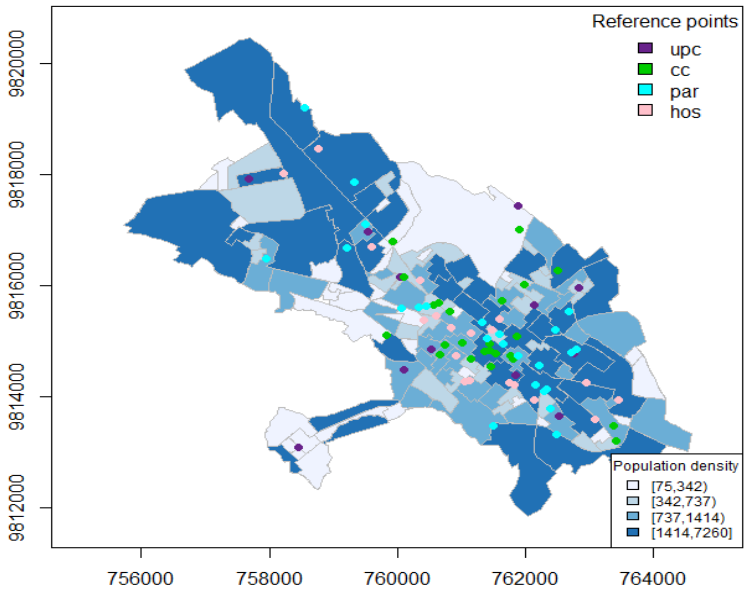


Figure 4. Population density (blue scale), and locations of some landmarks, such as community police units (upc), markets (cc), parks (par), and hospitals (hos).

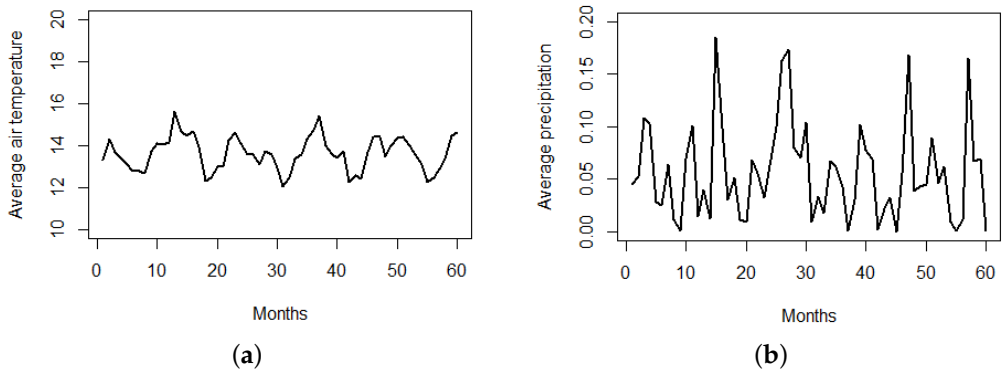


Figure 5. Monthly averages of some climatological variables in Riobamba. (**a**) Temperature. (**b**) Precipitation.

A first exploratory analysis by month highlights that the highest numbers of crimes for general records are found in January, June, and October, flagrant crimes are highest in February, September, and October, and crimes recorded by police increased in January, April, and May (see Figure 6a). This is an indication both that the three types of crime datasets behave differently, and that the month of year plays an important role.

When we look at the data by weekday (see Figure 6b), we find the highest numbers of crimes on Fridays and Saturdays, and according to their spatial location (see Figure 3), there is a high level of crime cases in the downtown area.

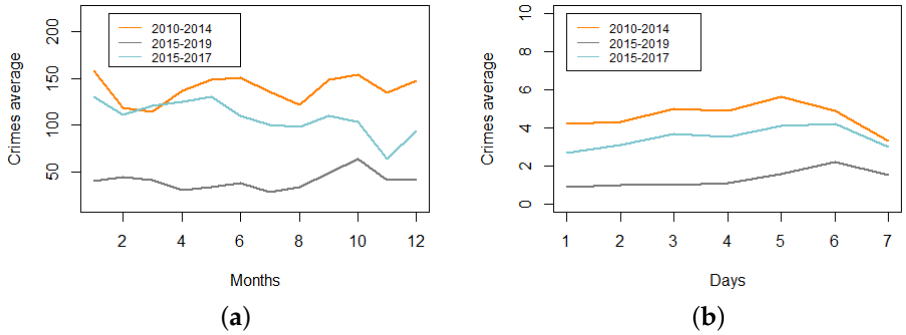


Figure 6. Crime average counts by months (**a**) and days (**b**) for crimes recorded from MI 2010–2014, CJCH 2015–2019, and UPC 2015–2017.

The locations of the landmarks are taken into account as nearest-neighbor distances between any crime event and the corresponding landmark location. These distances inform about the link between a particular crime and how close one of these landmarks is, and so they inform if landmarks act as attractors or repulsors of crimes. The distributions of these distances are shown in Figure 7, noting how small distances between crime events and landmarks are much more frequent than larger ones, indicating naively that these landmarks could be sources or attractors of crimes.

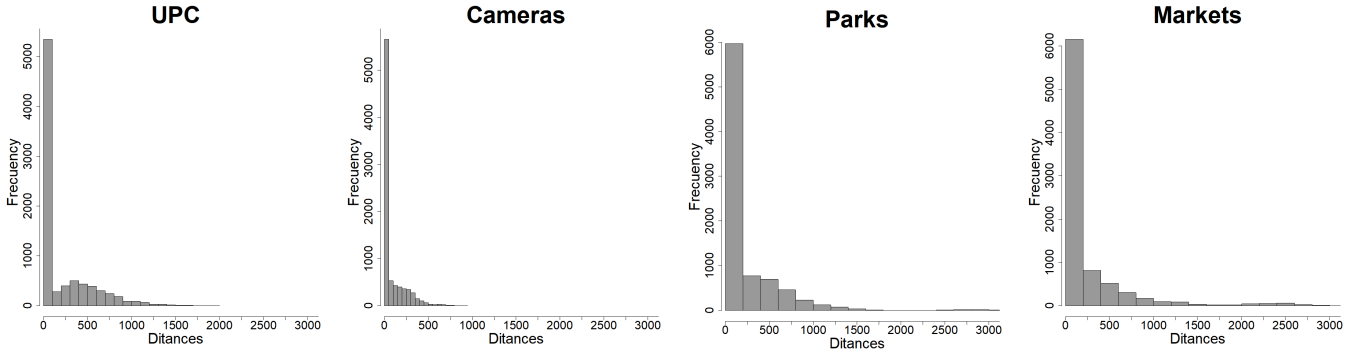


Figure 7. Distance distributions: upc (community police units), cameras (surveillance cameras), parks, and markets (squares, shopping malls, supermarkets).

3. Methodology

The overall methodological approach in this paper is to structure space–time dependency for count data through a combination of spatial dependence in a latent process model and temporal dependence in a data model, with exogenous factors that vary over space and/or time. Following [24], we consider a stochastic difference equation for the intensity of the space–time process within a class of spatially correlated self-exciting spatio-temporal models for count data that capture both data model dependence as well as dependence in a latent spatial process. In particular, we focus on a SPINGARCH(1,1) model that overall allows the modeler to define the autocorrelation present in the data and the mean–variance ratio with greater flexibility.

We use the following notation throughout this manuscript. We denote by $(s_1, s_2, \ldots, s_n)$ a vector of spatial (lattice) locations that remain fixed in time, and let t be a discrete time period. We denote by $N|s_i|$ the spatial neighborhood of lattice location s_i. Finally, $Y(s_i, t)$ is the observed process at spatial location s_i and time t, and $X(s_i, t)$ is the unobserved latent process. We use a conditional Poisson distribution and place spatio-temporal structure on the covariance of the latent Gaussian process. The data model $Y(s_i, t)$ can be defined conditionally on the process model $X(s_i, t)$. As a result, the process model is a function of both observable spatial or temporal covariates and unobservable latent spatial errors.

In our case, the spatio-temporal intensity $\lambda(s,t)$ provides the process model, and our full model is a stochastic difference equation operating directly on the intensity function. Thus, crime counts in space and time, $Y(s_i,t)$, are conditionally distributed Poisson random variables for $i = 1,\ldots,n$, i.e., $Y(s_i,t)|\lambda(s_i,t) \sim \text{Pois}(\lambda(s_i,t))$, with $\lambda(s_i,t)$ representing the rate at location s_i in time t. Hence, $\text{E}[Y(s_i,t)|\lambda(s_i,t)] = \lambda(s_i,t)$.

We can assume that a change in crime rate at a specific location and in a specific period is a function of particular geographic features of the location given by $\alpha_t = (\alpha(s_1,t), \alpha(s_2,t), \cdots, \alpha(s_n,t))^T$ (also known as reference baseline tension and is simply a function of potentially variable exogenous factors), together with two other factors, a natural deterioration χ, and repeated victimization η.

We propose a SPINGARCH(1,1) model, with $Y(s_i,t)$ defined conditionally on the intensity $\lambda(s_i,t)$ which can be modeled using observable spatial and temporal covariates $\alpha(s_i,t)$, as well as non-observable latent errors ϵ_t. Thus, the final model is defined through the following hierarchical structure:

$$Y(s_i,t)|\lambda(s_i,t) \sim \text{Pois}(\lambda(s_i,t)) \tag{1}$$

with

$$\lambda_t = \exp(X_t + \epsilon_t) + \eta Y_{t-1} + \kappa\lambda_{t-1}$$
$$X_t \sim \text{Gau}\left(\alpha_t, (I_{n,n} - \zeta C)^{-1}\sigma^2\right)$$
$$\epsilon_t \sim \text{Gau}\left(0, I_{n,n}\sigma_\epsilon^2\right),$$

where $\kappa = 1 - \chi$ represents stress in the absence of repeated victimization, η captures the expected change due to repeated or nearly repeated actions [24], $\lambda_t = (\lambda(s_1,t), \lambda(s_2,t), \ldots, \lambda(s_n,t))^T$ is a Markov chain in $(\mathbb{R}^+)^n$, and the same notation applies for Y_t and X_t. Note that $I_{n,n}$ is the identity matrix, σ^2 is the conditional variance, and ζ controls the amount of spatial dependence in the model not captured by the covariates in α_t. Large scale spatial structure is accounted for in the latent process X_t by the spatial regression parameter α_t, whereas small scale spatial structure is accounted for by conditionally defining X_t. For the latter, a conditionally autoregressive (CAR) model is used (through spatially adjacent neighbors):

$$X(s_i,t)|X(s_j,t), s_j \in N|s_i| \sim N(\mu(s_i,t), \sigma^2) \tag{2}$$

with

$$\mu(s_i,t) = \alpha(s_i,t) + \zeta \sum_{s_j \in N|s_i|} [X(s_j,t) - \alpha(s_j,t)].$$

If locations s_i and s_j are neighbors, the entry (i,j) of C will be one. Note that by adding space–time noise $\epsilon(s_i,t) \overset{iid}{\sim} N(0,\sigma_\epsilon^2)$, further variation in the spatio-temporal process is allowed; see some alternative approaches to modeling spatial effects in count data in [19].

3.1. Bayesian Inference

The hierarchical model defined above depends on a set of parameters in the final level of the hierarchy given by $\theta = (\eta, \kappa, \sigma, \sigma_\epsilon, \alpha_t, \zeta)$, similarly to a classical Besag–York–Mollié (BYM) model [31] which defines a fully Bayesian spatial model (see [32]).

Thus, following [24] we use a Bayesian inferential framework consisting in updating beliefs about θ according to the available data through an a priori density $\pi(\theta)$ and a conditional density or likelihood $\pi(data|\theta)$ to obtain $\pi(\theta|data)$, a posterior density of θ given the data. The a priori joint distribution of the parameters in the model can be expressed as $\pi(\theta) = \pi(\eta|\kappa)\pi(\kappa)\pi(\sigma)\pi(\sigma_\epsilon)\pi(\alpha_t)\pi(\zeta)$, where independence is assumed in the background except for η and κ due to the condition $\eta + \kappa < 1$. Letting $U(s_i,t) = X(s_i,t) + \epsilon(s_i,t)$, the full conditional distribution for θ is given by

$$\pi(\theta|Y,U) \propto \prod_{t=1}^{T} \pi(Y_t|\lambda_t)\pi(\lambda_t|\lambda_{t-1},Y_{t-1},\theta,U_t)\pi(U_t,\theta)\pi(\lambda_0|\theta)\pi(Y_0|\lambda_0)\pi(\theta), \quad (3)$$

and for U we have

$$\pi(U|Y,\theta) \propto \prod_{t=1}^{T} \pi(Y_t|\lambda_t)\pi(\lambda_t|\lambda_{t-1},Y_{t-1},\theta,U_t)\pi(U_t|\theta)\pi(\lambda_0|\theta)\pi(Y_0|\lambda_0). \quad (4)$$

For any inference on the parameters, Markov chain Monte Carlo (MCMC) must take samples of the full latent state density U, which requires evaluation of

$$\log(U|\alpha_t,\sigma,\sigma_\epsilon,\zeta) \propto \frac{-T\times n}{2}\log(2\pi) + \frac{1}{2}\log|\Sigma_f^{-1}(\theta)| - \frac{1}{2}(X_t-\alpha_t)^T\Sigma_f^{-1}(\theta)(X_t-\alpha_t). \quad (5)$$

Since we can assume that the neighborhood structure is constant for all periods, we can interpret $\Sigma_f(\theta)$ as the full space–time covariance matrix $(I_{T,T}\otimes I_{n,n}-C)^{-1}\sigma^2 + I_{n\times T}\sigma_\epsilon^2$.

The covariance structure's sparsity means that the only computations of $\frac{1}{2}(X_t-\alpha_t)_t^T\Sigma_f^{-1}(\theta)(X_t-\alpha_t)$ that need to occur are for spatial neighbors. Thus, the computation of the determinant $\log|\Sigma_f^{-1}(\theta)|$ is the most challenging element. The specific structure of $\Sigma^{-1}(\theta) = (I_{n,n}-C)(1/\sigma^2)$ makes $\log|\Sigma^{-1}(\theta)| = \frac{n}{2\log\sigma^2} + \log|I_{n,n}-\zeta N|$, where N is the neighborhood or adjacency matrix. This in turn can be rewritten as $\log|\Sigma_f^{-1}(\theta)| = \frac{n\times T}{\log\sigma^2} + T\log|\Sigma^{-1}(\theta)|$, finally resulting in

$$\log|\Sigma_f^{-1}(\theta)| = T\log|\Sigma^{-1}(\theta)| \propto \frac{n\times T}{\log\sigma^2} + T\sum_{j=1}^{n}(1-\zeta\chi_j), \quad (6)$$

with χ_j being the eigenvalues of the neighborhood matrix which depend solely on its structure and not on the parameters.

3.2. Generalized Additive Models with B-Splines

The action of the deterministic covariates depending on space or space–time is highly non-linear on the responses. Thus, we have used a generalized additive model (GAM) that supports integrated smoothness estimation addressing the lack of linearity [33]. The GAM results in a more efficient analytical method than the more classical linear models. The relationship between each predictor x_i and the mean of the response variable, $g(u)$, is indirect because it is calculated using the smooth (usually splines with polynomial bases [33]) function $f(x_i)$:

$$g(u) = \beta_0 + \sum_{i=1}^{p} f_i(x_i). \quad (7)$$

We can also have multivariate versions of the smooth functions per temporal instant. For example,

$$g(u) = \beta_0 + f_t(x_1,x_2), \quad (8)$$

with f_t being a smooth spatial surface in the t-th time. This smooth surface for each t can be written as $f_t(x_1,x_2) = \mathbf{B}_s\psi^{(t)}$, where $\mathbf{B}_s = \mathbf{B}_1\otimes\mathbf{B}_2$ is a B-spline k-dimensional basis of dimension $I\times k_1k_2$ arising from the Kronecker product per row of the marginal B-spline bases for $\mathbf{B}_1$, $\mathbf{B}_2$, and $\psi^{(t)} = \left(\psi_1^{(t)},...,\psi_{k_1k_2}^{(t)}\right)^T$. Note that k_1 and k_2 are the number of columns of the marginal bases $\mathbf{B}_1$ and $\mathbf{B}_2$, respectively, and depend on the number of nodes and degree of polynomials used to generate these bases (see [34]). The generalized cross-validation (GCV) criterion is used to estimate the smoothing parameters, which provide the degree of smoothness. To define the version of smoothing that best fits the data, we test the joint interactions of the spatial covariates with crime.

4. Results

The city is divided into $m = 141$ administrative zones (see Figure 4), whose centroids are denoted by $\{s_1, s_2, \ldots, s_{141}\}$. We compute the adjacency or neighborhood matrix needed in determining the spatial latent process; this is a sparse matrix that permits optimizing the computational costs [35]. Our temporal unit is month, so we consider the number of crimes per zone per month. As the three datasets have different time periods (recall we have for general records 2010–2014, for flagrant crimes 2015–2019, and for police records 2015–2017), we set the time instants (n) for the first two cases $t \in \{1, 2, \ldots, 60\}$, and for the latter case $t \in \{1, 2, \ldots, 36\}$.

We refer to Section 2 for a number of covariates considered in our model. In particular, we compute nearest-neighbor distances from each crime to community police units (upc), to surveillance cameras (cam), to markets (cc), to parks (par), and to hospitals (hos). These distances are averaged per administrative zone providing matrices of $m \times n$. The population density (pob) enters the model as a spatial-only covariate of dimension $m \times 1$. Although we initially considered two climatological variables (see Figure 5), in an exploratory analysis we noted they were not significant in this particular city, with monthly average temperatures ranging within 12–15 °C and precipitation ranging within 0.00–0.15 mm. Thus, although they are considered in other studies, in our particular region they are not influential on crime.

Although we tested all possible combinations of a multivariate GAM, we found that univariate GAMs provide the best fits. Therefore, we use a univariate generalized additive model with cubic B-splines (denoted by $\widehat{f}_i^{[3]}$) which allows the incorporation of non-linear relationships between each covariate and the response variable. Our complete GAM model is as follows:

$$\begin{aligned} \alpha_t = & \widehat{\beta}_0 + \widehat{f}_1^{[3]}(upc) + \widehat{f}_2^{[3]}(cam) + \\ & + \widehat{f}_3^{[3]}(cc) + \widehat{f}_4^{[3]}(par) + \widehat{f}_5^{[3]}(hos) + \widehat{f}_6^{[3]}(pob). \end{aligned} \tag{9}$$

In particular, the final significant models for each of the three datasets are the following:

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

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

and

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

Figure 8 depicts for each dataset the corresponding fitted model with B-splines. We observe how the model fits the real data delineating its behavior well.

Once the parameter α_t is estimated depending on the covariates, and keeping $\zeta = 0.99$ fixed near the edge of the parameter space [24], the remaining parameters $\theta = (\eta, \kappa, \sigma, \sigma_\epsilon, \alpha_t)$ are estimated using a Bayesian framework as previously explained. We use informative beta distributions as priors for η and κ, and Cauchy for σ and σ_ϵ that minimize the impact on the posterior densities (see also [24]). We run three Markov chains of 70,000 iterations each per parameter, and for each of the three datasets. The first 10,000 iterations are discarded as a burn-in period, and we take samples every 100 iterations to remove any possible sample autocorrelation. Figures 9a–11a depict the MCMC chains for the four parameters and for the three datasets. We can see visually the convergence and stability of these chains. The posterior distributions of each of the parameters are shown in Figures 9b–11b. We also show the autocorrelation of the parameters as sampled from the posterior distribution, reconfirming the absence of autocorrelation (Figures 9c–11c).

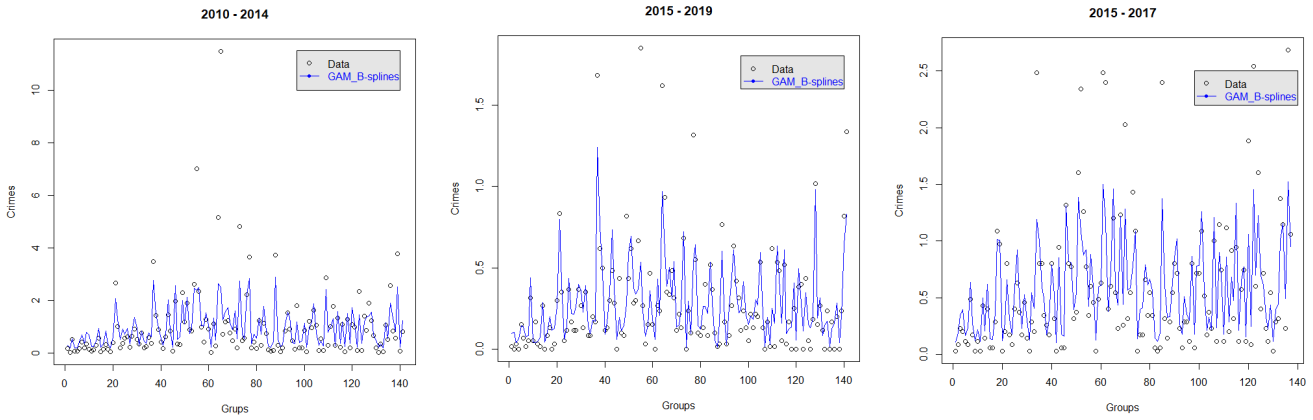


Figure 8. Real crime data and fitted model using B-splines over the exogenous covariates for the three datasets. (**a**) Crimes recorded by the MI. (**b**) Flagrant crimes recorded by the CJCH. (**c**) Crimes recorded by the UPC.

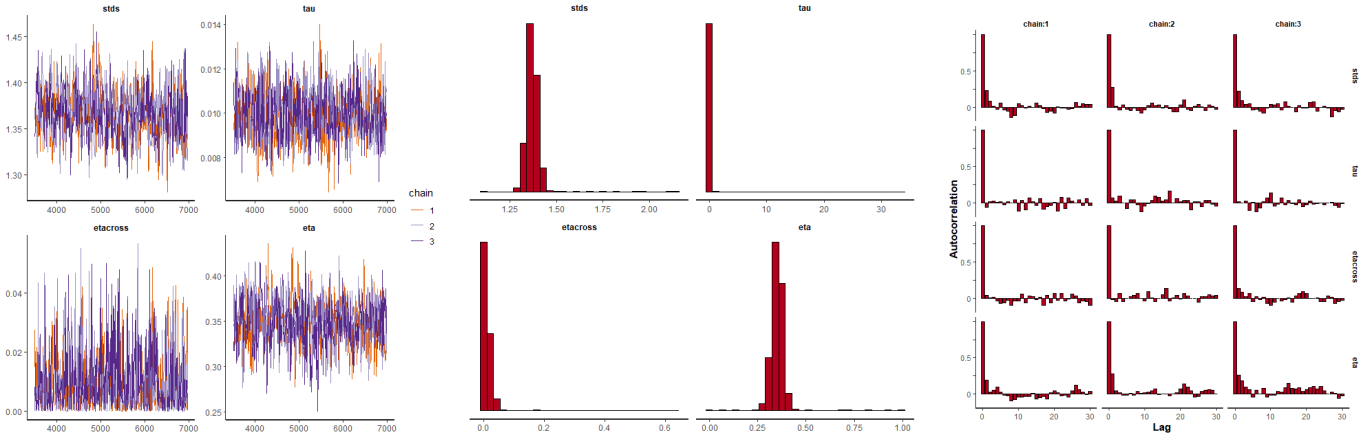


Figure 9. Bayesian inference for the MI data (2010–2014), where: $\eta = eta$, $\kappa = etacross$, $1/\sigma^2 = tau$, $\sigma_\epsilon = stds$. (**a**) Markov chain convergence. (**b**) Parameter distributions. (**c**) Parameter autocorrelations.

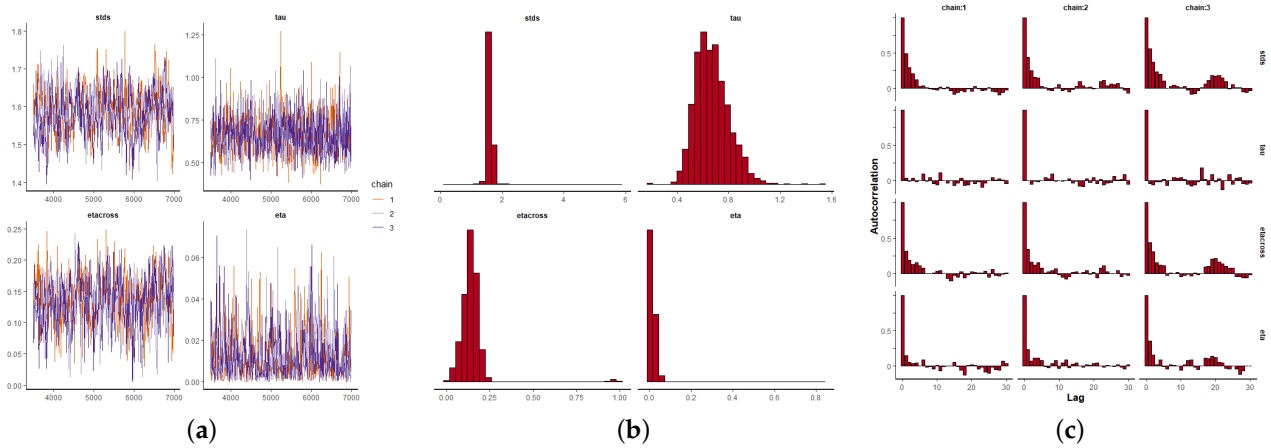


Figure 10. Bayesian inference for the CJCH data (2015–2019), where: $\eta = eta$, $\kappa = etacross$, $1/\sigma^2 = tau$, $\sigma_\epsilon = stds$. (**a**) Markov chain convergence. (**b**) Parameter distributions. (**c**) Parameter autocorrelations.

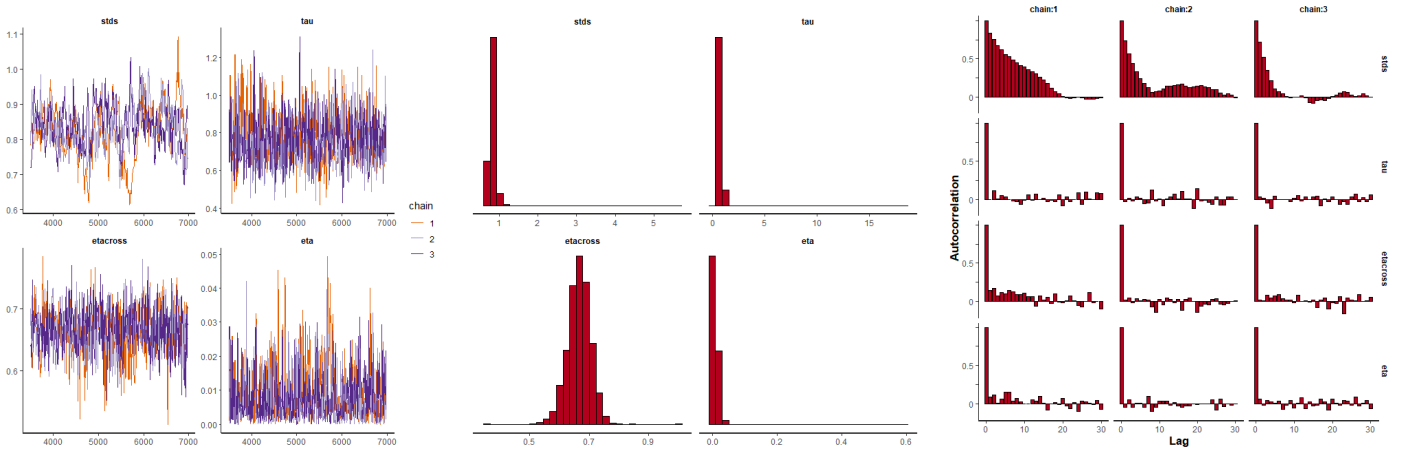


Figure 11. Bayesian inference for the UPC data (2015–2017), where: $\eta = eta$, $\kappa = etacross$, $1/\sigma^2 = tau$, $\sigma_\epsilon = stds$. (**a**) Markov chain convergence. (**b**) Parameter distributions. (**c**) Parameter autocorrelations.

Tables 1–3 show summary statistics of the estimates and diagnostic statistics for the posterior distributions. Noting that κ (coefficient of deterioration) weights the expected value (i.e., the intensity) while η (coefficient of victimization) weights the data or observations themselves, it is expected that η is larger for MI crimes as from 2015 police interventions were increased in response to an increased number of complaints. In addition, κ is larger for UPC indicating that the model weights the expected intensity more, giving more importance to what is expected than to real data. Recall that the effective sample size (n_eff) and a measure of chain equilibrium (Rhat) are the number of independent draws in the sample and diagnostic statistics on chain convergence, respectively. Rhat values equal or close to 1 are indicative of convergence [36]. For completeness, we also calculate Shannon entropy for each of the parameters associated with each of the three datasets. Taking advantage of the output of the MCMC for each parameter, for which we have a posterior sample of size 1050, we discretize its range length into a number of bins equal to the integer value closest to the square root of the sample size, and calculate the entropy H based on these bins. The value of H is shown in Tables 1–3, and reflects the uncertainty associated with each parameter. A small H indicates a small uncertainty in the estimation of the parameter, and thus a larger confidence on its value. Indeed, the estimated parameters with the lowest uncertainty are η and $1/\sigma^2$ for the period 2015–2019, κ for 2010–2014, and σ_ϵ^2 for 2015–2017.

Table 1. Posterior distribution of the parameters for the MI crimes 2010–2014.

Posterior Parameters	Mean	Sd	2.5%	25.0%	50.0%	75.0%	97.5%	n_eff	Rhat	H
η	0.35	0.03	0.29	0.33	0.35	0.37	0.40	685	1.00	2.99
κ	0.01	0.01	0.00	0.00	0.01	0.01	0.04	920	1.00	3.01
$1/\sigma^2$	0.01	0.00	0.01	0.01	0.01	0.01	0.01	1028	1.00	2.69
σ_ϵ	1.37	0.03	1.31	0.35	0.037	1.39	1.42	618	1.01	2.95

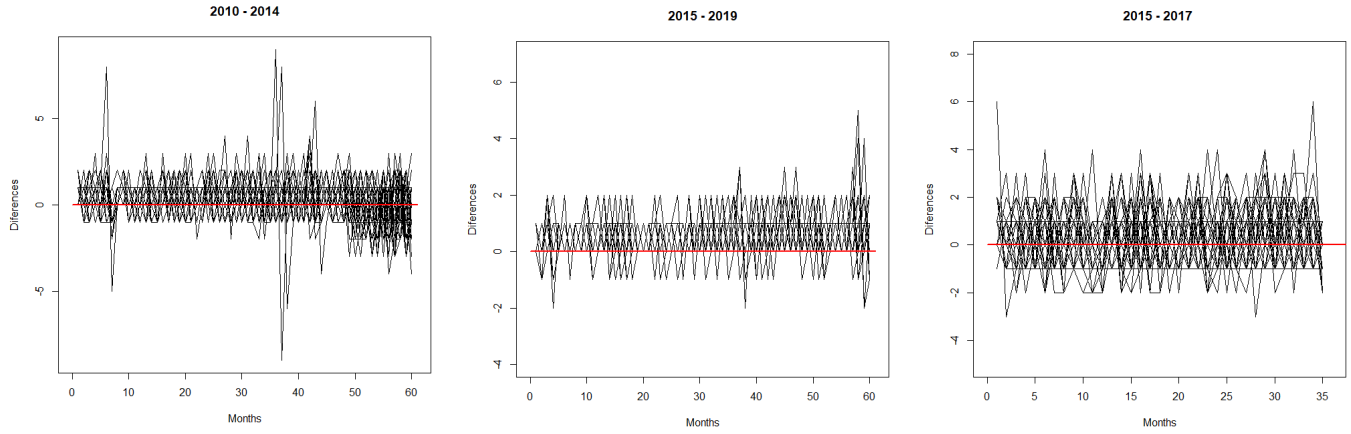
Table 2. Posterior distribution of the parameters for the CJCH crimes 2015–2019.

Posterior Parameters	Mean	Sd	2.5%	25.0%	50.0%	75.0%	97.5%	n_eff	Rhat	H
η	0.01	0.01	0.00	0.00	0.01	0.02	0.50	530	1.00	2.98
κ	0.13	0.04	0.04	0.11	0.14	0.16	0.21	342	1.01	2.86
$1/\sigma^2$	0.66	0.12	0.45	0.57	0.65	0.74	0.93	1057	1.00	3.07
σ_ϵ	1.59	0.06	1.47	1.55	1.59	1.63	1.70	293	1.01	2.68

Table 3. Posterior distribution of the parameters for the UPC crimes 2015–2017.

Posterior Parameters	Mean	Sd	2.5%	25.0%	50.0%	75.0%	97.5%	n_eff	Rhat	*H*
η	0.01	0.01	0.00	0.00	0.01	0.01	0.03	635	1.00	2.89
κ	0.67	0.04	0.59	0.64	0.67	0.69	0.74	505	1.00	2.96
$1/\sigma^2$	0.77	0.14	0.53	0.67	0.75	0.85	1.08	881	1.00	2.91
σ_ϵ	0.83	0.07	0.68	0.79	0.83	0.88	0.96	79	1.02	2.70

As a goodness-of-fit tool, we compute temporal mean square prediction errors (MSPEs) (see Table 4), which report a measure of differences between predicted and real values, noting that the SPINGARCH with cubic B-splines shows the best MSPE values. Additionally, we compute differences between predicted and real values in space–time (see Figure 12), with the corresponding MSPEs being 0.45 (MI), 0.20 (CJCH), and 0.41 (UPC), keeping small in general terms.



(**a**) (**b**) (**c**)

Figure 12. Differences between predicted and real values in space–time. (**a**) Crimes recorded by MI. (**b**) Flagrant crimes recorded by CJCH. (**c**) Crimes recorded by UPC.

Table 4. Temporal mean square prediction errors (MSPEs).

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

As a final illustration, we compare the temporal prediction of an INGARCH(1,1) model in which there is no spatial effect, of a SPINGARCH(1,1) with exogenous factors entering linearly in a regression fashion, and of our SPINGARCH(1,1) with exogenous factors that vary on space and time and modeled with cubic B-splines. The comparative results are depicted in Figure 13, noting that SPINGARCH(1,1) with smoothed covariates with B-splines provides the best predicted results as they are closer to the real crime data.

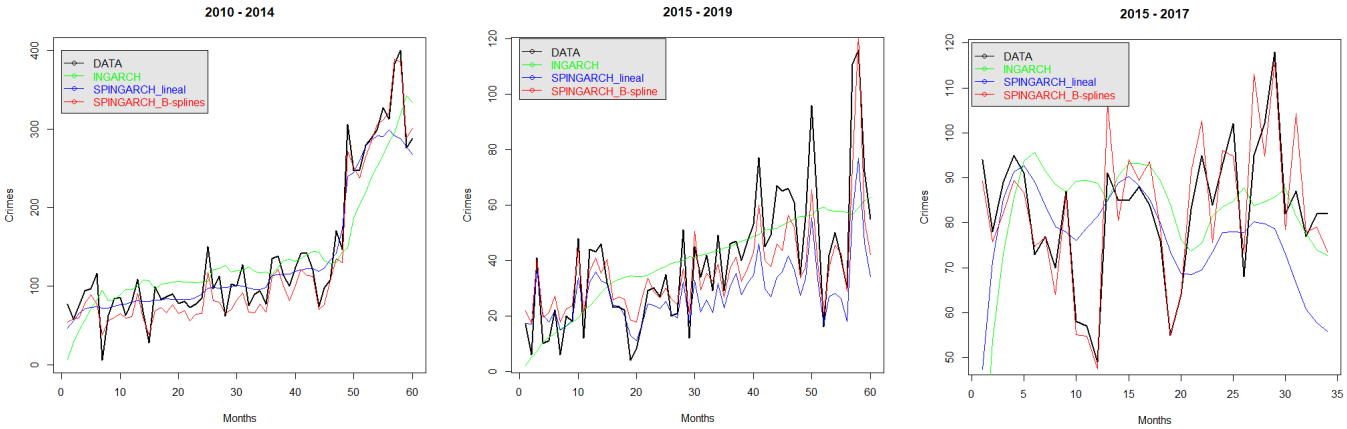


Figure 13. Real and predicted crimes using three competing models: INGARCH(1,1), SPINGARCH(1,1) with exogenous factors being constant in t, and SPINGARCH(1,1) with exogenous factors varying on t and modeled with cubic B-splines. (**a**) Crimes recorded by MI. (**b**) Flagrant crimes recorded by CJCH. (**c**) Crimes recorded by UPC.

For our model, we also report the spatial predictions for the three crime datasets in the city of Riobamba, illustrating that our spatio-temporal model is flexible enough to provide accurate temporal predictions and also spatial predictions.

5. Discussion and Conclusions

This manuscript formulates a statistical model that contains both latent spatial and temporal dependencies in the form of a stochastic difference equation for the spatio-temporal intensity. This model is consistent with common beliefs about how violence and crime evolve in space and time. Indeed, the proposed model is a spatially and temporally correlated self-exciting spatio-temporal model that captures both data dependence and dependence on a latent spatial process along the line INGARCH(1,1) models do. Another aspect of our model is that the effect of exogenous covariates is added using non-linear B-splines which improves previous models with only linear forms on the covariates.

We have followed a Bayesian inferential framework as it is flexible and can handle estimation of a large number of parameters with complex structures, such as those considered here in space and time. We are able to estimate neighborhood structure in space and temporal autoregression behavior in time.

In analyzing crime data in the city of Riobamba, we were able to detect, by an extensive preliminary search and inspection, which distance-based covariates were most influential and how they entered the prediction model. We highlighted some differences amongst the three types of datasets. For the general registries (dataset for 2010–2014), the minimum distances to surveillance cameras and parks were important because through the monitoring of these cameras, a criminal event was foreseen or taken for granted, and in places such as parks, there is a greater police protection, especially on weekends. For flagrant crimes (2015–2019), the relevant covariates were distances to squares, shopping malls, and supermarkets, and the population density, as having greater population movement contributed to the immediate warning and denunciation of criminal events. Finally, for the police records (2015–2017), distances to cameras and upc had a more representative influence because most of the victims go to the police in the first instance requesting help, regardless of whether the registered criminal event is legally reported or not. The estimation results showed a higher number of crimes in area 65, called San Alfonso, because in the period 2010–2014 the largest market in the city was located there. However, for registered flagrant crimes (2015–2019), we found more predicted cases in zones 37, 55, 76, and 141 (La Dolorosa, La Merced, La Station, and Tubasec, respectively) while for the police files (2015–2017) nine

other zones were highlighted (see Figure 14, 2015–2017). These results provide valuable information to governmental entities in charge of citizen security to optimize resources by improving planning, deployment of police units, or patrolling and random verification.

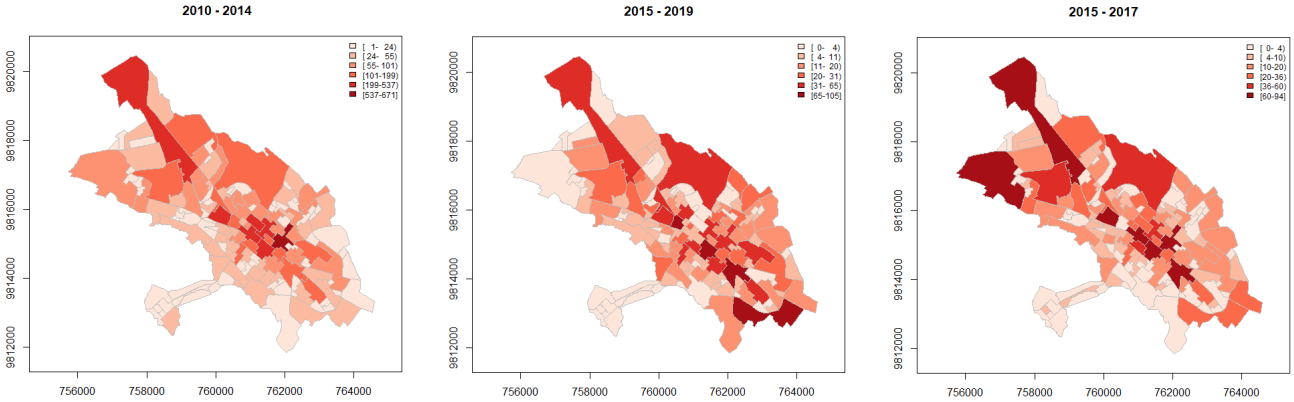


Figure 14. Spatial predictions of crimes from the posterior predictive distributions. (**a**) Crimes recorded by MI. (**b**) Flagrant crimes recorded by CJCH. (**c**) Crimes recorded by UPC.

Our data, although in the form of spatio-temporal coordinates, have some limitations. One is that the data provided by governmental entities do not have detailed information on the crime and the characteristics of the events. If they had some additional information, more complex models using mark information could have been used. Another aspect is that INEC (a governmental institution providing and making the data available) has as a minimum unit of study, the parishes, and does not keep statistics by district or by administrative division of the cities. This forced us to randomly disaggregate the data, causing crude approximations of the population.

Open ideas in the context of modeling crime data are many, but identifying crimes happening only in the network of streets in a city enhances the modeling task. In such a case, the Euclidean plane has to be substituted by the network support and this makes things different (see, for example, [37]). We can also think of models for location predictions of the following serial crime using the next hit predictor (NHP) method which adopts the framework of specific self-exciting processes created to characterize the correlations between crimes committed by the same criminal (see [38]).

We finally note that, in [39], the authors studied, by simulation and under different scenarios, the information/complexity transfer from intensity realizations to generated point patterns in spatial log-Gaussian Cox processes. As further research under the model structure proposed in the present paper, we aim at investigating the use of information-complexity measures for assessment of the influential significance of random covariates, involved in the specification of the unobservable latent process, for the response observed patterns. This represents an important challenge due to the intrinsically complex nature of the self-excitation mechanism, and would be particularly useful for comparing different scenarios (type of crimes, urban specificities, etc.), as well as for identification of the relevant explanatory covariates.

Author Contributions: Conceptualization, I.E., J.M.A. and J.M.; Data curation, I.E.; Formal analysis, J.M.A. and J.M.; Funding acquisition, J.M.A. and J.M.; Investigation, I.E., J.M.A. and J.M.; Methodology, I.E., J.M.A. and J.M.; Software, I.E.; Supervision, J.M.A. and J.M.; Visualization, I.E. All authors have read and agreed to the published version of the manuscript.

Funding: J.M. Angulo was partially supported by MCIU/AEI/ERDF, UE grant PGC2018-098860-B-I00, grant A-FQM-345-UGR18 cofinanced by ERDF Operational Programme 2014–2020 and the Economy and Knowledge Council of the Regional Government of Andalusia, Spain, and grant CEX2020-001105-M MCIN/AEI/10.13039/501100011033. J. Mateu was partially supported by grant PID2019-107392RB-I00/AEI/10.13039/501100011033 from the Spanish Ministry of Science and Innovation and grant UJI-B2018-04 from University Jaume I, Spain.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Cox, D.R. Some statistical methods connected with series of events. *J. R. Stat. Soc. Ser. B* **1955**, *17*, 129–157. [CrossRef]
2. Bartlett, M.S. The spectral analysis of point processes. *J. R. Stat. Soc. Ser. B* **1963**, *25*, 264–281. [CrossRef]
3. Hawkes, A.G. Spectra of some self-exciting and mutually exciting point processes. *Biometrika* **1971**, *58*, 83–90. [CrossRef]
4. Hawkes, A.G. Point spectra of some mutually exciting point processes. *J. R. Stat. Soc. Ser. B* **1971**, *33*, 438–443. [CrossRef]
5. Hawkes, A.G.; Oakes, D. A cluster process representation of a self-exciting process. *J. Appl. Probab.* **1974**, *11*, 493–503. [CrossRef]
6. Kedem, B.; Fokianos, K. *Regression Models for Time Series Analysis*; Wiley-Interscience: Hoboken, NJ, USA, 2002.
7. Jung, R.C.; Tremayne, A.R. Useful models for time series of counts or simply wrong ones? *Adv. Stat. Anal.* **2011**, *95*, 59–91. [CrossRef]
8. Davis, R.A.; Fokianos, K.; Holan, S.H.; Joe, H.; Livsey, J.; Lund, R.; Pipiras, V.; Ravishanker, N. Count time series: A methodological review. *J. Am. Stat. Assoc.* **2021**, *116*, 1533–1547. [CrossRef]
9. Xu, Y.; Zhu, F. A new GJR-GARCH model for $\mathbb{Z}$-valued time series. *J. Time Ser. Anal.* **2022**, *43*, 490–500. [CrossRef]
10. Weiß, C.H.; Zhu, F.; Hoshiyar, A. Softplus INGARCH models. *Stat. Sin.* **2022**, *32*, 1099–1120. [CrossRef]
11. Li, Q.; Chen, H.; Zhu, F. Robust estimation for Poisson integer-valued GARCH models using a new hybrid loss. *J. Syst. Sci. Complex.* **2021**, *34*, 1578–1596. [CrossRef]
12. Fokianos, K.; Tjøstheim, D. Log-linear Poisson Autoregression. *J. Multivar. Anal.* **2011**, *102*, 563–578. [CrossRef]
13. Heinen, A. *Modeling Time Series Count Data: An Autoregressive Conditional Poisson Model*; CORE Discussion Paper 2003/62; Université Catholique de Louvain: Ottignies-Louvain-la-Neuve, Belgium, 2003.
14. Zhu, F. Modeling overdispersed or underdispersed count data with generalized Poisson integer-valued GARCH models. *J. Math. Anal. Appl.* **2012**, *389*, 58–71. [CrossRef]
15. Zhu, F. Modeling time series of counts with COM-Poisson INGARCH models. *Math. Comput. Model.* **2012**, *56*, 191–203. [CrossRef]
16. Besag, J. Spatial interaction and the statistical analysis of lattice systems. *J. R. Stat. Soc. Ser. B* **1974**, *36*, 192–225. [CrossRef]
17. Augustin, N.H.; McNicol, J.; Marriott, C.A. Using the truncated auto-Poisson model for spatially correlated counts of vegetation. *J. Agric. Biol. Environ. Stat.* **2006**, *11*, 1–23. [CrossRef]
18. Kaiser, M.S.; Cressie, N. Modeling Poisson variables with positive spatial dependence. *Stat. Probab. Lett.* **1996**, *35*, 423–432. [CrossRef]
19. Glaser, S. *A Review of Spatial Econometric Models for Count Data*; Working Paper; Universität Hohenheim, Fakultät Wirtschafts- und Sozialwissenschaften: Stuttgart, Germany, 2017.
20. Ferland, R.; Latour, A.; Oraichi, D. Integer-valued GARCH Process. *Time Ser. Anal.* **2006**, *27*, 923–942. [CrossRef]
21. Reinhart, A. Rejoinder: A review of self-exciting spatio-temporal point processes and their applications. *Stat. Sci.* **2018**, *33*, 330–333. [CrossRef]
22. Mohler, G.; Short, M.; Brantingham, J.; Schoenberg, F.; Tita, G. Self-exciting point process modeling of crime. *J. Am. Stat. Assoc.* **2011**, *106*, 100–108. [CrossRef]
23. Hu, T.; Zhu, X.; Duan, L.; Guo, W. Urban crime prediction based on spatio-temporal Bayesian model. *PLoS ONE* **2018**, *13*, e0206215. [CrossRef]
24. Clark, N.J.; Dixon, P.M. A class of spatially correlated self-exciting statistical models. *Spat. Stat.* **2021**, *43*, 100493. [CrossRef]
25. Andresen, M.; Malleson, N. Intra-week spatial-temporal patterns of crime. *Crime Sci.* **2015**, *4*, 12. [CrossRef]
26. Cepa, C.; Zabala, R.; López, M. Proyecto seguridad barrial con involucramiento de los vecinos en Riobamba-Ecuador. *Revista Caribeña de Ciencias Sociales* **2018**, 1–21.
27. Trejo, C.; Cisneros, J. La delincuencia en la ciudad de Guayaquil, un análisis espacial de su distribución por delito. *Revista Caribeña de Ciencias Sociales* **2013**, 1–17.
28. Clark, N.; Dixon, P. Modeling and estimation for self-exciting spatio-temporal models of terrorist activity. *Ann. Appl. Stat.* **2018**, *12*, 633–653. [CrossRef]
29. Chávez, Y.; Cortez, P.; Medina, P. Quantification of losses caused by delinquency in Ecuador. *Anal. Rev. Anál. Estad.* **2013**, *5*, 51–64.
30. Castro, D.; Jácomey, J.C.; Mancero, J. Seguridad ciudadana en Ecuador: Política ministerial y evaluación de impacto, años 2010–2014. *Nova Criminis* **2015**, *9*, 111–148.
31. Morris, M.; Wheeler, K.; Simpson, D.; Mooney, S.; Gelman, A.; DiMaggio, C. Bayesian hierarchical spatial models: Implementing the Besag York Mollié model in stan. *Spat. Spatio-Temporal Epidemiol.* **2019**, *31*, 100301. [CrossRef]

32. Thamrin, S.; Alimun. Geographical mapping of dengue fever incidence 2012–2016 in Makassar, Indonesia. *IOP Conf. Ser. Earth Environ. Sci.* **2019**, *279*, 012013. [CrossRef]
33. Taylan, P.; Weber, G.W.; Liu, L.; Yerlikaya-Özkurt, F. On the foundations of parameter estimation for generalized partial linear models with B-splines and continuous optimization. *Comput. Math. Appl.* **2010**, *60*, 134–143. [CrossRef]
34. Vicente, G.; Goicoa, T.; Ugarte, M.D. Multivariate Bayesian spatio-temporal P-spline models to analyze crimes against women. *Biostatistics* **2021**, kxab042. [CrossRef] [PubMed]
35. Solarte, G.; Soto, J.; Muñoz, L. Matrices dispersas descripción y aplicaciones. *Sci. Tech.* **2013**, *18*, 171–177.
36. Vuong, Q.H.; La, V.P.; Nguyen, M.H.; Ho, M.T.; Tran, T.; Ho, M.T. Bayesian analysis for social data: A step-by-step protocol and interpretation. *MethodsX* **2020**, *7*, 100924. [CrossRef]
37. Gilardi, A.; Mateu, J.; Borgoni, R.; Lovelace, R. Multivariate hierarchical analysis of car crashes data considering a spatial network lattice. *J. R. Stat. Soc. Ser. A* **2022**, 1–28. [CrossRef]
38. Li, Y.; Wang, T. Next hit predictor-self-exciting risk modeling for predicting next locations of serial crimes. *arXiv* **2018**, arXiv:1812.05224.
39. Medialdea, A.; Angulo, J.M.; Mateu, J. Structural complexity and informational transfer in spatial log-Gaussian Cox processes. *Entropy* **2021**, *23*, 1135. [CrossRef] [PubMed]